Stochastic Analysis of Bidding in Sequential Auctions and Related Problems.

S. Kartikeya Puranam

Rutgers Business School.
Thesis Overview

1. New auction models
   - Fixed demand model
   - Integrated auction-inventory model

2. Optimizing taboo criteria in Markov decision processes

3. Conclusions and further research
Types of auction procedures

We consider first price sealed bid. Other auction procedures are

- Second price sealed bid.
- English auction.
- Dutch auction.
Types of auction models studied in literature

Single item sold at a single auction
1. Many issues studied for this model.
2. Very important result is the revenue equivalence theorem due to Vickrey.

Multiple item auctions
1. Sequential vs simultaneous.
2. Simultaneous auction models study size of batches, price of individual items and batches (combinatorial auctions).
Sequential auctions

- First substantial study was by Weber.
- Empirical studies of real life sequential auctions, have noticed a downward decline in the winning bid prices.
- Most of the models assume that bidders have unit demands.
- Multi-item demand models are very few.
- Most of these models use game theoretic analysis.
- Rothkopf and Oren consider a decision theoretic model of sequential auctions.
Consider a firm ("the buyer"):  
- In a given time period there is a fixed demand $L$ for a certain item that must be met.  
- In each period there are $N$ auctions where these items are sold. ($N > L$)  
- It is assumed that there is a buy-it-now-price available at which the buyer can obtain the item outright.  
- After the buyer acquires all $L$ items he does not bid in any of the remaining auctions if any.  
- The objective of the buyer is to minimize his expected total cost for the period.  
- The bidders’ valuations derive from the strict demand fulfillment requirement.


**Auction procedure**

- Every bidder submits a sealed bid.
- The highest bidder wins the auction.
- At the end of each auction the winning bid is announced.
- In each auction every bidder can bid amounts \( \{a_1, a_2, \ldots, a_p\} \).
- We assume that \( a_0 = 0 \) and \( a_p \) denotes the *buy-it-now-price*.
- If there is a tie we assume that the buyer loses. (to simplify exposition)
Probability of winning

$p_m(a)$ corresponds to the probability that the buyer wins an auction when his bid is $a$ and there are $m$ opponents present.

$$p_m(a) = P(\text{all opponents' bids } < a).$$

(1)

$p_m(a_0) = 0$ and $p_m(a_p) = 1$. 
Number of opponents

- $Z_n$ represents the number of opponents in the $n^{th}$ auction.
- $Z_n$ is a discrete time Markov chain for $n = 1, 2, \cdots, N$.
- Transition probabilities are:
  
  $$q_{mm'}(n) = P(Z_{n+1} = m' | Z_n = m)$$

- The distribution of the number of opponents at the beginning of the time period:
  
  $$q_m(1) = P(Z_1 = m)$$
MDP model

1. The state space $S$ is the set of triplets $(n, m, l)$.
   - $n (1 \leq n \leq N)$ represents the number of auctions remaining.
   - $m$ represents the number of bidders participating in the auction.
   - $l (0 \leq l \leq L)$ represents the number of items already acquired by the buyer.

   with the restriction: $L \leq n + l \leq N$.

2. In any state $(n, m, l)$ the following action sets $A(n, m, l)$ are available.
   - $A(n, m, L) = \{a_0\}$,
   - $A(n, m, I) = \{a_p\}$, for $n + l = L$,
   - $A(n, m, I) = \{a_1, a_2, \ldots, a_p\}$.
When an action $a \in A(n, m, l)$ is taken in state $(n, m, l)$ the following transitions are possible.

- If $l = L$ the only possible transition is back to the same state $(n, m, L)$ with probability $1$.
- When $l < L$
  - if the buyer wins the auction the next state is $(n - 1, m', l + 1)$ with probability $p_m(a) q_{mm'}$
  - otherwise it is $(n - 1, m', l)$ with probability $(1 - p_m(a)) q_{mm'}$.

The following costs are incurred.

- In states $(n, m, L)$ there is no cost.
- In states $(n, m, l)$ with $L + 1 < n + l \leq N + L$, the expected cost when action $a$ is taken is $a p_m(a)$. 
Dynamic programming equations

For a given state \((n, m, l)\), let:

- \(a^*_{n,m,l}\) denote the optimal action,
- \(v(n, m, l)\) denote the value function
- \(w(n, m, l; a)\) denote the expected remaining cost when action \(a\) is taken in state \((n, m, l)\) and an optimal policy is followed thereafter.
- \(v(n, m, l) = w(n, m, l; a^*_{n,m,l})\).
Dynamic programming equations contd.

\[
\nu(n, m, l) = \min_{a \in A} \{ w(n, m, l; a) \} \tag{2}
\]

where,

\[
w(n, m, l; a) = a \ p_m(a) \\
+ \sum_{m' = 1}^{\infty} p_m(a) q_{mm'}(N - n) \nu(n - 1, m', l + 1) \\
+ \sum_{m' = 1}^{\infty} (1 - p_m(a)) q_{mm'}(N - n) \nu(n - 1, m', l),
\]

if \( n + l > L \),

\[= na_p, \text{ if } n + l = L, \ l < L\]

\[= 0, \text{ otherwise.}\]
Assumptions (needed for monotonicity proofs)

**Assumption A.** For any fixed $m$, $p_m(a)$ is an increasing function of $a$.

**Assumption B.** For any fixed $a$, $p_m(a)$ is a decreasing function of $m$.

**Assumption C.** There exists a function $G$ with $\sum_{i=\infty}^{\infty} G(i) = 1$ such that:

$$
q_{mm'}(n) = \begin{cases} 
G(m' - m) & \text{if } m' > 1, \\
\sum_{k=-\infty}^{-m+1} G(k) & \text{if } m' = 1.
\end{cases}
$$

(3)
The Single Item Case With a Constant Number of Opponents

- In all the auctions $L = 1$ and $m^0 \geq 1$.
- The state $n$ corresponds to the number of remaining auctions before the item is acquired.
- $A(n) = \{ a_1, \ldots, a_p \}$, $\forall n$.
- When action $a$ is taken in state $n$
  - the buyer either wins and leaves the auction with probability $p(a)$
  - or he loses and transitions to state $n - 1$ with probability $1 - p(a)$. 
Dynamic programming equations

\[ v(n) = \min_{a \in A} \{ w(n; a) \} \quad (4) \]

where,

\[
\begin{align*}
w(n; a) & = a p(a) + (1 - p(a)) v(n - 1), \text{ if } n > 1, \\
& = a_p, \text{ if } n = 1.
\end{align*}
\]
Theorem 2.2.1

Under assumption A the following relationships hold for all $n$.

\[
\begin{align*}
v(n+1) & \leq v(n), \\
a^*_{n+1} & \leq a^*_n.
\end{align*}
\]
Theorem 2.2.1

Proof:

- Both Eqs. (5) and (6) proved by induction on $n$.
- For $n = 1$, since $v(1) = a_p$ both inequalities are true.
- Induction assumption is
  \[ v(n) \leq v(n - 1), \]
  \[ a_n^* \leq a_{n-1}^*, \]
- Using the above we prove:
  \[ v(n + 1) \leq v(n), \]
  \[ a_{n+1}^* \leq a_n^*. \]
Theorem 2.2.1

Proof contd.
- We first prove that $a_{n+1}^* \leq v(n)$ by contradiction.
- Note that $v(n+1)$ is a convex combination of $v(n)$ and $a_{n+1}^*$.
- This implies that if $a_{n+1}^* > v(n)$, then $w(n+1, a_0) < v(n)$, which is a contradiction.
- If $a_{n+1}^* \leq v(n)$ then $v(n+1) \leq v(n)$. 
Theorem 2.2.1

Proof contd.

- We prove $a^*_{n+1} \leq a^*_n$ by contradiction.
- From the definitions of $v(\cdot)$ and $w(\cdot)$ we have $w(n+1; a^*_n) > v(n+1)$ and $w(n; a^*_{n+1}) > v(n)$. Simplifying the inequalities leads to

$$\frac{a^*_n p(a^*_{n+1}) - a^*_{n+1} p(a^*_n)}{p(a^*_{n+1}) - p(a^*_n)} < v(n) \quad (7)$$

and

$$\frac{a^*_n p(a^*_{n+1}) - a^*_{n+1} p(a^*_n)}{p(a^*_{n+1}) - p(a^*_n)} > v(n-1). \quad (8)$$

- Combining the results leads to $v(n) > v(n-1)$ which contradicts the induction assumption.
Computational example

- $N = 200$, $m^0 = L = 1$ and $A = \{1 \ldots, 50\}$.
- The winning probability $p(a)$ is calculated assuming the single opponent chooses bids from $A$ with equal probabilities.
Plot of $a_n^*$ versus $n$

Figure: $a_n^*$ vs $n$. 
Plot of $v(n)$ versus $n$

Figure: $v(n)$ vs $n$
Theorem 2.5.1

Under assumption A the following relationships hold true for all \( n, m \) and \( l \).

\[
\begin{align*}
v(n + 1, m, l) & \leq v(n, m, l) \quad (9) \\
v(n, m, l) & \leq v(n, m, l - 1) \quad (10) \\
v(n, m, l) & \geq v(n, m - 1, l) \quad (11)
\end{align*}
\]
Theorem 2.5.1

Proof:
- The proofs of all three inequalities is by induction on \( n \).
- For \( n = 1 \), all three inequalities hold since \( v(n, m, L) = 0 \) and \( v(n, m, L - n) = na_p \).
- Ineqs. 12 and 13 follow from the induction assumption as

\[
\begin{align*}
v(n, m, l) &= \min_{a \in A} w(n, m, l; a) \\
&\leq \min_{a \in A} w(n - 1, m, l; a) = v(n - 1, m, l).
\end{align*}
\]

and

\[
\begin{align*}
v(n, m, l) &= \min_{a \in A} w(n, m, l; a) \\
&\leq \min_{a \in A} w(n, m, l - 1; a) = v(n, m, l - 1).
\end{align*}
\]
Theorem 2.5.1

Proof contd.

- For Ineq. (11) assume $\nu(n - 1, m, l) \geq \nu(n - 1, m - 1, l)$ and prove $\nu(n, m, l) \geq \nu(n, m - 1, l)$.
- $\nu(n, m - 1, l) > \nu(n, m, l)$ produces a contradiction.
- This above implies $w(n, m - 1, l; a^*_{n,m,l}) > \nu(n, m, l)$.
- This implies

$$p_{m-1}(a^*)(a^* + E(n - 1, m - 1, l + 1)) + (1 - p_{m-1}(a^*))E(n - 1, m - 1, l)$$

$$> p_m(a^*)(a^* + E(n - 1, m, l + 1)) + (1 - p_m(a^*))E(n - 1, m, l).$$

where $a^* = a^*_{n,m,l}$. 
Using condition B the above inequality simplifies to

\[ p_{m-1}(a^*)(a^* + E(n - 1, m - 1, l + 1)) + (1 - p_{m-1}(a^*))E(n - 1, m - 1, l) >

\]

which contradicts the induction assumption.
Theorem 2.5.2

Theorem (2.5.2)

Under assumption A, B and C the following relationships hold true for all \( n, m \) and \( l \).

\[
\begin{align*}
a_n^{*},m,l+1 & \leq a_n^{*},m,l \quad (12) \\
a_n^{*},m,l & \leq a_n^{*},m+1,l \quad (13) \\
a_n^{*},m,l & \leq a_{n-1}^{*},m,l \quad (14)
\end{align*}
\]
Theorem 2.5.2

Proof:

- The proofs of all three inequalities is by induction on $n$.
- For $n = 1$, all three inequalities hold since $a_{1,m,L}^* = 0$ and $a_{1,m,L-1}^* = a_p$.
- For Ineq. (12) we assume that $a_{n-1,m,l}^* \leq a_{n-1,m,l-1}^*$ and prove $a_{n,m,l+1}^* \leq a_{n-1,m,l}^*$.
- We prove that $a_{n,m,l+1}^* > a_{n-1,m,l}^*$ leads to a contradiction.
- From $v(n, m, l) < w(n, m, l; a_{n,m,l+1}^*)$ and $v(n, m, l + 1) < w(n, m, l + 1; a_{n,m,l}^*)$ we have

$$E(n - 1, m, l) + E(n - 1, m, l + 2) > 2E(n - 1, m, l + 1).$$

- The above inequality when simplified using the induction assumption leads to a contradiction.
Theorem 2.5.2

Proof contd.

- For Ineq. (14) we assume that \( a^*_{n,m,l} \leq a^*_{n-1,m,l} \) and prove \( a^*_{n+1,m,l} \leq a^*_{n,m,l} \).
- We prove that \( a^*_{n+1,m,l} > a^*_{n,m,l} \) leads to a contradiction.
- From \( v(n, m, l) < w(n, m, l; a^*_{n+1,m,l}) \) and \( v(n + 1, m, l) < w(n + 1, m, l; a^*_{n,m,l}) \) we have
  \[
  E(n−1, m, l+1)−E(n, m, l+1) > E(n−1, m, l)−E(n, m, l).
  \]
- The above inequality when simplified using the induction assumption leads to a contradiction.
Theorem 2.5.2

Proof contd.

- For Ineq. (13) we assume that $a_{n-1,m,l}^* \leq a_{n-1,m+1,l}^*$ and prove $a_{n,m,l}^* \leq a_{n,m+1,l}^*$.
- We prove that $a_{n,m,l}^* > a_{n,m+1,l}^*$ leads to a contradiction.
- From $v(n, m, l) < w(n, m, l; a_{n+1,m,l}^*)$ and $v(n+1, m, l) < w(n+1, m, l; a_{n,m,l}^*)$ we get
  
  $$E(n-1, m, l) - E(n-1, m+1, l) < E(n-1, m, l+1) - E(n-1, m+1, l+1).$$

- The above inequality when simplified using the induction assumption and conditions B and C leads to a contradiction.
Computational example 1

- $N = 20$, $1 \leq m \leq 20$, $L = 1$ and $A = \{1 \ldots , 10\}$.
- $G(i) = 1/39$ for $i = -19 \ldots , 0, \ldots 19$.
- The winning probability $p_m(a)$ is calculated assuming that each of the opponents choose bids from $A$ with equal probabilities.
Plot of $a^*_{n,m,l}$ versus $n$

**Figure:** $a^*_{n,m,l}$ vs $n$ vs $m$. 
Computational example 2

- $N = 20$, $m^0 = 4$, $L = 10$ and $A = \{1 \ldots , 10\}$.
- The winning probability $p(a)$ is calculated assuming the four opponents choose bids from $A$ with equal probabilities.
Plot of $a^*_{n,m,l}$ versus $n$

Figure: $a^*_{n,4,l}$ vs $n$ vs $l$. 
Conclusions

The main results of this chapter are as follows.

1. We obtain bidding strategies for this problem by modeling it as a Markov decision process.

2. We prove that, under certain assumptions, the optimal value function \( v(n, m, l) \) and the optimal bid \( a^*_{n,m,l} \) are:
   - decreasing functions of \( n \), the number of remaining auctions
   - increasing functions of \( m \), the number of opponents
   - decreasing functions of \( l \), the inventory on hand.
Future research

The main results of this chapter are as follows.

- The bid distribution is not constant through all the auctions.
- A model with learning using a Bayesian framework.
- Batch sales with variable batch sizes.
We consider the following problem of a firm ("the buyer").

- In each time period multiple quantities of an item are bought via auctions and then sold in a secondary market.
- In each time period there are $N$ auctions of the item.
- The demand $D$ in the secondary market has a known distribution.
  
  \[
  p_D(d) = P(D = d), \quad P_D(d) = P(D \leq d), \quad \text{and} \quad \bar{P}_D(d) = 1 - P_D(d).
  \]
- The sales price $R$ also has a known distribution with $r = E(R) < \infty$.
- Excess demand is assumed to be lost and the penalty of losing sales of $x$ units is $\delta(x)$. 
Auction procedure

- Every bidder submits a sealed bid.
- The highest bidder wins the auction.
- At the end of each auction the winning bid is announced.
- In each auction every bidder can bid amounts \( \{a_1, a_2, \ldots a_p\} \).
- \( a_0 = 0 \) represents action of not bidding.
- If there is a tie the buyer loses. (to simplify exposition)
Probability of winning

\[ p_m(a) \] corresponds to the probability that the buyer wins an auction when his bid is \( a \) and there are \( m \) opponents present.

\[ p_m(a) = P(\text{all opponents' bids} < a). \quad (15) \]

\[ p_m(a_0) = 0 \]
Number of opponents

- $Z_n$ represents the number of opponents in the $n^{th}$ auction.
- $Z_n$ is a discrete time Markov chain for $n = 1, 2, \cdots, N$.
- Transition probabilities are:

$$q_{mm'}(n) = P(Z_{n+1} = m'|Z_n = m).$$

- The distribution of the number of opponents at the beginning of the time period:

$$q_m(1) = P(Z_1 = m).$$
Assumptions (needed for monotonicity theorems)

**Assumption A.** For any fixed \( m \), \( p_m(a) \) is an increasing function of \( a \).

**Assumption B.** For any fixed \( a \), \( p_m(a) \) is a decreasing function of \( m \).

**Assumption C.** There exists a function \( G \) with \( \sum_{i=-\infty}^{\infty} G(i) = 1 \) such that:

\[
q_{mm'}(n) = \begin{cases} 
G(m' - m) & \text{if } m' > 1, \\
\sum_{k=-\infty}^{-m+1} G(k) & \text{if } m' = 1.
\end{cases}
\]  

(16)

**Assumption D.** \( \delta(x) \) is an increasing convex function of \( x \) and \( \delta(x) = 0 \) if \( x \leq 0 \).
Two Cases

We consider two cases of this problem.

1. **Salvage Case:** Unsold items at the end of a time period are salvaged at $s$.

2. **Inventory Case:** Unsold at the end of a period are carried over as inventory; inventory carrying cost of $h$ per item per period.
The Markov Decision Process Model:

1. **State space** is the set \( \{(n, m, x)\} \)
   - \( n \) represents the number of auctions remaining during the current period
   - \( m \) represents the number of bidders in the current auction
   - \( x \) represents the inventory level at the beginning of the current auction.
   - If \( n = 0 \) then \( m = 0 \).
   - \((0, 0, x)\) represents the end of a period when all auctions are over.
In any state \((n, m, x)\) the following action sets \(A(n, m, x)\) are available.

- \(A(0, 0, x) = \{a_0\}\).
- \(A(n, m, x) = \{a_0, \ldots, a_p\}\) for \(n > 0\).

When an action \(a \in A(n, m, x)\) is taken in state \((n, m, x)\) the following transitions are possible.

- Starting from state \((0, 0, x)\) the next state is \((N, m, 0)\) with probability \(q_m(1)\).
- If \(n > 0\)
  - If the buyer wins the auction the next state is \((n - 1, m', x + 1)\) with probability \(p_m(a) q_{m'm}^\prime\).
  - otherwise it is \((n - 1, m', x)\) with probability \(\bar{p}_m(a) q_{m'm}(N - n)\).
When an action \( a \in A(n, m, x) \) is taken in state \((n, m, x)\) the expected reward \( r_a(n, m, x) \) is as follows.

\[
\begin{align*}
  r_a(0, 0, x) &= \sum_{d=0}^{x} (rd + s(x - d)) p_D(d) \\
  &\quad + \sum_{d=x+1}^{\infty} (rx - \delta(d-x)) p_D(d)
\end{align*}
\]

\( r_a(n, m, x) = -a \ p_m(a) \) if \( n > 0 \).
Dynamic programming equations

For a given state \((n, m, x)\), let:

- \(a^*_{n,m,x}\) denote the optimal action
- \(v(n, m, x)\) denote the value function
- \(w(n, m, x; a)\) denote the expected future reward when action \(a\) is taken in state \((n, m, x)\) and an optimal policy is followed thereafter.
- \(v(n, m, x) = w(n, m, x; a^*_{n,m,x})\).
Dynamic programming equations contd.

\[ v(n, m, x) = \max_{a \in A} \{ w(n, m, x; a) \} \tag{17} \]

where,

\[ w(n, m, x; a) = \begin{cases} r_a(0, 0, x) + \beta \sum_{m=1}^{\infty} q_m(1)v(N, m, 0) & \text{if } n = 0 \\ r_a(n, m, x) + p_m(a)E(n - 1, m, x + 1) + \bar{p}_m(a)E(n - 1, m, x) & \text{if } n > 0, \end{cases} \]

\[ E(n - 1, m, x) = \sum_{m'=1}^{\infty} q_{mm'}(N - n)v(n - 1, m', x) \] and \( \beta \) is the discount factor.
Lemma 3.2.1

Lemma (3.2.1)

The expected reward function in state \((0, 0, x)\), \(r_a(0, 0, x)\) is an increasing function of \(x\) i.e.

\[
r_a(0, 0, x) \leq r_a(0, 0, x + 1). \tag{18}
\]

Proof:

\[r_a(0, 0, x + 1) - r_a(0, 0, x)\]

can be simplified to

\[
\sum_{x+1}^{\infty} rp_D(d) + \sum_{d=0}^{x} sp_D(d) + \sum_{d=x+1}^{\infty} (\delta(d-x) - \delta(d-x-1))p_D(d).
\]
Theorem 3.2.3

Under assumptions A, B and C the following relationships hold.

\[ v(n, m, x) \leq v(n, m, x + 1) \quad \forall \ n \geq 0, \quad (19) \]
\[ v(n, m, x) \leq v(n + 1, m, x) \quad \forall \ n \geq 0, \quad (20) \]
\[ v(n, m, x) \geq v(n, m + 1, x) \quad \forall \ n > 0. \quad (21) \]
Theorem

Under assumption A, B and C the following relationships hold true for all \( n, m \) and \( l \).

\[
\begin{align*}
a^*_n, m, x & \geq a^*_n, m, x+1 \quad \text{for } n \geq 0, \\ a^*_n, m, x & \geq a^*_{n+1}, m, x \quad \text{for } n > 0, \\ a^*_n, m, x & \leq a^*_{n}, m+1, x \quad \text{for } n \geq 0.
\end{align*}
\]
The Markov Decision Process Model:

1. State space is the set \( \{(n, m, x)\} \)

- \( n \) represents the number of auctions remaining during the current period
- \( m \) represents the number of bidders in the current auction
- \( x \) represents the inventory level at the beginning of the current auction.
- If \( n = 0 \) then \( m = 0 \).
- \((0,0,x)\) represents the end of a period when all auctions are over.
In any state \((n, m, x)\) the following action sets \(A(n, m, x)\) are available.

- \(A(0, 0, x) = \{a_0\}\).
- \(A(n, m, x) = \{a_0, \ldots, a_p\}\) for \(n > 0\).

When an action \(a \in A(n, m, x)\) is taken in state \((n, m, x)\) the following transitions are possible.

- From state \((0, 0, x)\) the next state is \((N, m, (x - d)^+)\) with probability \(q_m(1)p_D(d)\), if \(x > d\) and \(q_m(1)\overline{P}_D(d)\), otherwise, where \(d = 0, 1, \ldots\).
- If \(n > 0\)
  - If the buyer wins the auction the next state is \((n - 1, m', x + 1)\) with probability \(p_m(a)q_{mm'}\).
  - otherwise it is \((n - 1, m', x)\) with probability \(\overline{p}_m(a)q_{mm'}(N - n)\).
When an action $a \in A(n, m, x)$ is taken in state $(n, m, x)$ the expected reward $r_a(n, m, x)$ is as follows.

$$r_a(n, m, x) = \begin{cases} \sum_{d=0}^{\infty} (r(d \land x) - h(x - d)^+ - \delta(d - x)^+) p_D \\ -ap_m(a) \end{cases}$$

where $d \land x = \min\{d, x\}$ and $d \lor x = \max\{d, x\}$. 
\[ v(n, m, x) = \max_{a \in A} \{ w(n, m, x; a) \} \] (25)

\[ w(n, m, x; a) = r(0, 0, x) + \beta \sum_{d=0}^{x} E_1(N, m, (x - d) \vee 0)p_D(d) \]
\[ = r_a(n, m, x) + p_m(a)E(n - 1, m, x + 1) \]
\[ + \bar{p}_m(a)E(n - 1, m, x) \]

\[ E(n - 1, m, x) = \sum_{m'=1}^{\infty} q_{mm'}(N - n)v(n - 1, m', x), \]
\[ E_1(N, m, x) = \sum_{m=1}^{\infty} q_m(1)v(N, m, x) \] and \( \beta \) is the discount factor.
**Lemma 3.3.1**

**Assumption E.** \( P(D \leq x_*) \leq \frac{r}{r+h} \)

**Lemma (3.3.1)**

*Under assumption E the expected reward function in state (0, 0, x), \( r_a(0, 0, x) \) is an increasing function of x i.e.*

\[
r_a(0, 0, x) \leq r_a(0, 0, x + 1).
\]

**Proof:**

\[
r(0, 0, x + 1) - r(0, 0, x) = rP_D(x + 1) - h\bar{P}_D(x + 1) + \sum_{d=x+1}^{\infty} (\delta(d-x) - \delta(d-x-1))p_D(d),
\]

which under assumption E is non-negative.
Theorem 3.3.3

Theorem (3.3.3)

Under assumptions A, B, C and D the following relationships hold.

\[ v(n, m, x) \leq v(n, m, x + 1) \quad \forall \ n \geq 0 \ \text{and} \ x < x^*, \quad (27) \]
\[ v(n, m, x) \leq v(n + 1, m, x) \quad \forall \ n \geq 0 \ \text{and} \ x \leq x^*, \quad (28) \]
\[ v(n, m, x) \geq v(n, m + 1, x) \quad \forall \ n \geq 0 \ \text{and} \ x \leq x^*. \quad (29) \]
Theorem 3.3.4

Under assumption A, B, C and D the following relationships hold true for all n, m and x.

\[
\begin{align*}
a_{n,m,x}^* & \geq a_{n,m,x+1}^* \quad \text{for } n \geq 0, \quad (30) \\
a_{n,m,x}^* & \geq a_{n+1,m,x}^* \quad \text{for } n > 0, \quad (31) \\
a_{n,m,x}^* & \leq a_{n,m+1,x}^* \quad \text{for } n \geq 0. \quad (32)
\end{align*}
\]
Successive approximation equations

At the $k^{th}$ step of the successive approximation process, let

1. $v^k(n, m, x)$ denote the value function in state $(n, m, x)$
2. $E^k(n, m, x) = \sum_{m'=1}^{\infty} q_{mm'} v^k(n, m', x)$
3. $E_1^k(N, 0, x) = \sum_{m'=1}^{\infty} q_{m'}(1) v^k(N, m', x)$
4. $a^*_{n,m,x}(k)$ denote the optimal action in state $(n, m, x)$
5. $\lim_{k \to \infty} v^k(n, m, x) = v(n, m, x)$
6. $\lim_{k \to \infty} a^*_{n,m,x}(k) = a^*_{n,m,x}$. 
Successive approximation equations

\[ \nu^k(n, m, x) = \max_{a \in A} \{w^k(n, m, x; a)\} \]

\[ w^k(n, m, x; a) \]

\[ = r(0, 0, x) + \beta \sum_{d=0}^{x} E_1^{k-1}(N, 0, (x - d) \lor 0) p_D(d) \]

\[ = r_a(n, m, x) + p_m(a) E^{k-1}(n - 1, m, x + 1) \]

\[ + \bar{p}_m(a) E^{k-1}(n - 1, m, x) \text{ if } x < x^*, \]

\[ = E^{k-1}(n - 1, m, x^*) \text{ if } x = x^*. \]
Future research

The main results of this chapter are as follows.

- This model can be extended to the case where the items are sold after each auction.
- The bid distribution is not constant through all the auctions.
- A model with learning using a Bayesian framework.
- Batch sales with variable batch sizes.
Previous work

- Derman avoided the costs:
  - Problem was to find the policy that maximizes the expected time between replacements.
  - Subject to the constraint that the probability of replacement through undesirable states is less than or equal to a fixed number.

- Nakagawa studied a complex system and calculated the average number of visits to a given state before system failure.

- Dynamic control of a system based on taboo measures has not been studied.
Introduction

- Optimization of Markovian Systems is based on costs/rewards which are difficult to determine.
- In such situations, one could consider maximizing taboo optimization measures.
- The taboo measures we consider are the taboo mean first passage reward and the taboo mean first passage time.
- Computing policies that maximize the taboo measures considered is in general a hard problem.
- Well known methods from MDP theory can not be applied.
1. Rewards and costs are known for some states only.
   - Taboo first passage reward to be used as an optimization criterion.
   - The first example shows how this can be achieved in the context of an inventory control problem.

2. None of the costs or rewards associated with the states are known.
   - Optimization criterion is to maximize the taboo mean return time to a fixed state.
   - The second example models such a case in the context of the classical replacement problem.
General notation

- $X_n$ denotes the state of the system at transition $n$.
- $S = \{i = 1, 2, \ldots, S\}$ denotes the finite state space of the system.
- The finite set $A_i$ denotes the actions available for each $i \in S$.
- $S = G \cup B \cup C$
- $B_j$ denotes the set $B \cup \{j\}$
- $BG$ denotes the set $B \cup G$. 
Definitions

For any given policy $\pi$, and fixed states $i, j$ where $i \in C \cup G$ and $j \in G$ let

$$
\tau_{ij} = \inf \{n : X_n = j, X_\nu \notin Bj, \text{ for } \nu \leq n - 1, X_0 = i\}.
$$

We will use the following notation.

- $B^n_{fi}(\pi) = P_\pi(\tau_{ij} = n \mid X_0 = i)$,
- $B_{fi}(\pi) = P_\pi(\tau_{ij} < \infty \mid X_0 = i)$,
- $B_{mi}(\pi) = E_\pi(\tau_{ij})$. 
Definitions contd.

\[ \tau_{iG} = \inf\{n : X_n \in G, X_\nu \notin B \cup G, \text{ for } \nu \leq n - 1, X_0 = i \} \]

and

\[
\begin{align*}
B_f^{iG}(\pi) &= P_\pi(\tau_{iG} = n \mid X_0 = i), \\
B_f^{iG}(\pi) &= P_\pi(\tau_{iG} < \infty \mid X_0 = i), \\
B_m^{iG}(\pi) &= E_\pi(\tau_{iG}).
\end{align*}
\]
Definitions contd.

\[
B_{ij}^n = \sum_{0}^{n-1} r_t \mathbf{1}_{\{X_t \notin B_j, 0 \leq \nu < n \text{ and } X_n = j\}},
\]

\[
B_{ij}^N = \mathbb{E}_\pi \left( \sum_{n=1}^{N} B_{ij}^n \right),
\]

\[
B_{ij} = \lim_{n \to \infty} B_{ij}^n.
\]
Definitions contd.

\[ BV_{iG}^n = \sum_{0}^{n-1} r_X t \mathbf{1}_{\{X_t \notin B \cup G, \ 0 \leq t < n \ and \ X_n \in G\}} \]

\[ BV_{iG}^N = \mathbb{E}_\pi \left( \sum_{n=1}^{N} BV_{iG}^n \right) \]

\[ BV_{iG} = \lim_{n \to \infty} BV_{iG}^N \]

where \( \mathbf{1}_B \) is the indicator function of the set \( B \).
Theorem 4.2.1

\[ BV_{iG}(\pi) = r_i BV_{iG}(\pi) + \sum_{l \notin BG} p_{il}(\pi) BV_{iG}(\pi) \]  

(33)

Proof:

\[ BV_{iG}^1 = r_i \sum_{j \notin B} Bp_{ij} = r_i \sum_{j \notin B} BF_{ij}^1 = r_i BF_{iG}^1 \]

\[ BV_{iG}^N = r_i \sum_{n=1}^{N} BF_{iG}^n + \sum_{k \notin BG} p_{ik} BV_{kG}^{N-1} \text{ for } N \geq 2. \]
Theorem 4.2.1

Proof contd.

- Taking the limit as $N$ goes to infinity on both the sides we have

\[ B^N_{V_iG}(\pi) = r_i B^0_{V_iG}(\pi) + \sum_{l \notin BG} p_{il}(\pi) B^N_{V_iG}(\pi). \] (34)

- To prove that the limit exists we show that
  - $B^N_{V_iG}$ increases with $n$.
  - $B^N_{V_iG}$ is bounded for every $n$. 
Discussion of optimal policy

- $r_B f_G(\pi)$, depend on the policy $\pi$, and not just on the state action pairs.
- Uniformly optimal policy does not exist for all $i \in C \cup G$.
- Example: $S = \{1, \ldots, 5\}$, $G = \{1\}$, $B = \{5\}$, $A_i = \{1, 2\}$ and $r_i = 1$ if $i \in C \cup G$, or 0 otherwise.
- Transition probability matrix for action 1

\[
\begin{pmatrix}
0.41 & 0.01 & 0.05 & 0.17 & 0.36 \\
0.27 & 0.20 & 0.01 & 0.05 & 0.47 \\
0.23 & 0.03 & 0.29 & 0.01 & 0.53 \\
0.06 & 0.09 & 0.03 & 0.20 & 0.62 \\
0.20 & 0.20 & 0.20 & 0.20 & 0.20
\end{pmatrix}
\]
Discussion of optimal policy

- Transition probability matrix for action 2

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

- Policy that maximizes $5v_{11}$ is $a = 1$ in states 1 and 2 and $a = 2$ in states 3, 4 and 5.

- Policy that maximizes $5v_{31}$ is $a = 1$ in states 1, 2 and 3 and $a = 2$ in states 4 and 5.
Inventory example.

- A sequence of ordering decisions is to be made at the beginning of a number of periods of equal duration.
- The state is the discrete amount of inventory available at the beginning of every time period: \( \{0, 1, \ldots, S\} \).
- \( a = 0 \) denotes the action “do not order” and \( a = 1 \) denotes the action “order”.
- The states \( \{0, 1, \ldots, L\} \), denote the set of taboo states where \( L \) denotes the minimum level of inventory that is acceptable.
- If the system enters a state \( i \leq L \) then an order must be placed.
Inventory example contd.

- The demand in each period is assumed to be an \textit{I.I.D} random variable with a known discrete distribution.
- If \( a = 0 \) then the next state will be state \( j \) with probability \( p_{ij} = P\{\text{Demand} = i - j\} \).
- This model has been studied with the assumption of a known penalty cost for lost sales.
- The criterion is maximizing the taboo first return reward \( H^{\mathcal{VSS}} \) to the “full capacity” state \( S \).
- Here the taboo set \( H \) the set \( \{0\} \) of zero inventory level.
Inventory example contd.

- A purchase or ordering cost $c(i)$, where $i$ is the amount purchased
- A holding cost $h(.)$, associated with the cumulative excess of supply over demand
- A shortage or penalty cost $p(.)$, associated with the cumulative excess of demand over supply
- We shall also consider a revenue $r$ which we shall assume to be linear in the number of units sold.
Inventory example contd.

- Optimal deterministic policy is a “control limit” policy or a \((s, S)\) policy.
- Condition A: For each \(k = 1, \ldots, S\), the function \(\rho_k(i) = \sum_{j=k}^{S} p_{ij}\) \(i = 0, 1, \ldots S\) is increasing in \(i\).
- Condition B: For any increasing function \(h\) on \(S = \{0, 1, \ldots, S\}\), the function \(\xi(i) = \sum_{j=1}^{S} p_{ij}h(j)\) is an increasing function of \(i\).
Related References


