Stochastic Knapsacks and Dynamic Pricing *

David D. Yao
http://www.columbia.edu/~yao

*Joint work with Grace Y. Lin and Yingdong Lu of IBM Research.

Lunteren, January 2005

Outline

- Motivation and dynamic programming formulation
- Structure of the optimal policy
- The switch-over policy: convex programming
- Asmptotic optimality
- Optimal pricing
- Numerical examples
- Concluding remarks

Motivation: Reverse Logistics

- PC and laptop's returned from lease to be liquidated
  - involves de-manufacturing and re-selling decisions.
- Focus here:
  - re-selling: pricing decisions
  - different pricing strategies for different channels:
    - catalog sales,
    - negotiation (broker/trader),
    - auction.

Dynamic Programming Formulation

- N periods: $n = 1, ..., N$; m prices: $p_1, \ldots, p_m$;
- W units (inventory) available for sale.
- Order (demand) in period n: (P_n, Q_n):=(price, size),

$$\mathsf{P}[P_n = p_i, Q_n = j] := \theta_{ij}, \qquad i = 1, ..., m;\ j = 1, ..., W.$$

Assume $\sum_i \sum_j \theta_{ij} \leq 1$, with

$$\theta_0 := 1 - \sum_i \sum_j \theta_{ij} \geq 0$$

being the probability of no order arrival (in a period).

- Decision: after observing the realized (P_n, Q_n), whether or not to supply the order.

- $V(n,d)$: expected revenue, under optimal decisions, starting from period n $(\leq N)$, with d $(\leq W)$ units of inventory left.

- Dynamic programming recursion:

$$V(n,d) = V(n+1,d)[\theta_0 + \Theta(d)]$$
$$+\sum_i \sum_{j \leq d} \theta_{ij} \cdot \max\{jp_i + V(n+1, d-j), V(n+1,d)\},$$

where $\Theta(d) := \sum_i \sum_{j>d} \theta_{ij}$.

- Boundary: $V(N,d) = \sum_i \sum_{j \leq d} \theta_{ij} \cdot (jp_i)$.

Background: Stochastic Knapsack

- First developed in the late 1980's as a model to study admission control in telecommunication networks; Ross and Tsang (1989), Ross and Yao (1990).

- Knapsack capacity: bandwidth, or number of circuits.

- Jobs: random size and random occupancy time.

- Jobs come and go: infinite horizon; long-run average or discounted objective.

- Recent applications in revenue management and dynamic pricing: once committed, the capacity is gone; finite horizon; Kleywegt and Papastavrou (1998, 2001), Papastavrou *et al.* (1996), Van Slyke and Young (2000).

Structure of the Optimal Policy

- Suppose all orders are of unit size, $Q_n \equiv 1$. Denote $\theta_i := \mathsf{P}[P_n = p_i]$, $i = 1, ..., m$; $\theta_0 := 1 - \sum_{i=1}^{m} \theta_i$.

- DP recursion simplifies to:

$$V(n,d) = V(n+1,d)\theta_0 + \sum_i \theta_i \max\{p_i + V(n+1,d-1), V(n+1,d)\},$$

which can be rewritten as

$$V(n,d) = V(n+1,d) + \mathsf{E}\{P_n - [V(n+1,d) - V(n+1,d-1)]\}^+.$$

- Structural properties: $V(n,d)$ is
  - increasing in d, decreasing and concave in n,

- concave in d,
- submodular:

$$V(n,d) - V(n,d-1) \geq V(n+1,d) - V(n+1,d-1).$$

- "submodular-plus":

$$V(n,d+1) - V(n,d) \leq V(n+1,d) - V(n+1,d-1).$$

- Note: submodular and submodular-plus implies concave in d.
- Submodularity implies: if an order, with price p_i, is rejected in $(n+1,d)$, i.e.,

$$V(n+1,d) - V(n+1,d-1) > p_i.$$

then it should be rejected in (n,d) as well.

- Similarly, concavity in d implies: if an order, with price p_i, is rejected in (n, d), then it should be rejected in $(n, d-1)$ as well.

- Hence,

  - Lower orthant property (reject): If an order is rejected in state (n^*, d^*), then it is rejected in all lower states: $(n, d) \leq (n^*, d^*)$.
  - Upper orthant property (accept): If an order is accepted in state (n^*, d^*), then it is accepted in all upper states: $(n, d) \geq (n^*, d^*)$.

- Special case ("secretary problem"-like property): If $p_i \geq V(n^*, 1)$, for some n^*, then type i orders are always accepted in periods $n \geq n^*$.

Proof of the Properties

- Increasing in d and decreasing in n follows directly from the DP recursion:

$$V(n,d) = V(n+1,d)\theta_0 + \sum_i \theta_i \max\{p_i + V(n+1,d-1), V(n+1,d)\},$$

or

$$V(n,d) - V(n+1,d) = \mathsf{E}\{P_n - [V(n+1,d) - V(n+1,d-1)]\}^+.$$

- Concavity in d implies $V(n+1,d) - V(n+1,d-1)$ is decreasing in d; hence the RHS above is increasing in d, and so is the LHS, i.e.,

$$V(n,d) - V(n+1,d) \geq V(n,d-1) - V(n+1,d-1),$$

or

$$V(n,d) - V(n,d-1) \geq V(n+1,d) - V(n+1,d-1),$$

which is submodularity.

- That is, concavity in d (at $n+1$) implies submodularity.
- It can be shown that concavity in d (at $n+1$) implies submodular-plus too.
- Recall submodular and submodular-plus implies concavity in d (at n).
- This closes the loop in the induction proof.
- Apply induction on the DP recursion:

$V(N,d)$ concave in d

$\Rightarrow$ submodular and submodular-plus at $N-1$

$\Rightarrow$ concave (in d) at $N-1$

$\Rightarrow$ submodular and submodular-plus at $N-2$

$\Rightarrow$ concave (in d) at $N-2$; and so forth.

- Finally, concavity in n also follows from submodularity:

$$V(n,d)-V(n+1,d) = \mathsf{E}\{P_n-[V(n+1,d)-V(n+1,d-1)]\}^+,$$

the RHS is increasing in n (since $V(n+1,d)-V(n+1,d-1)$ is decreasing in n due to submodularity); hence so is the LHS.

Switch-Over Policy

- The structural results indicate that for orders of each price,

(a) there is a time *before* which the orders will *not* be accepted,
(b) there is a time *after* which the orders will *always* be accepted.

- This suggests a "switch-over" policy: suppose $p_1 \geq p_2 \geq \cdots \geq p_m$; then,

  - accept only orders with the highest price, p_1, in $[0, t_1]$;
  - accept orders with prices, p_1 and p_2, in $(t_1, t_2]$;
  - accept orders with prices, p_1, p_2 and p_3, in $(t_2, t_3]$;

  and so forth, with $0 < t_1 < t_2 < \cdots < t_{m-1} < T$ as decision variables.

Optimal Switch-Over Times

- Homogeneous batch size: $q_j = \mathsf{P}\{Q = j\}$, for $j = 1, ..., W$.

- Remaining inventory, embedded at Poisson order arrival epochs, forms a Markov chain with the probability transition matrix:

$$M = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ q_1 & 1-q_1 & 0 & \dots & 0 \\ q_2 & q_1 & 1-q_1-q_2 & \dots & 0 \\ & & & \dots & \\ q_W & q_{W-1} & q_{W-2} & \dots & 1-q_1-\dots-q_W \end{pmatrix}.$$

- Applying the switch-over policy in $[0, t_1]$, the expected revenue collected at time t_1 is,

$$p_1 \mathsf{E}[W - \mathbf{z}^T M^{N_1} \mathbf{w}]; \quad \mathbf{z}^T := (0, 0, ..., 0, 1), \quad \mathbf{w}^T := (0, 1, 2, ..., W);$$

where $N_1 := N_1(t_1)$ follows a Poisson distribution with mean $\lambda_1 t_1$.

- Recall, the generating function of a Poisson variable N with mean λ is: $\mathsf{E}[z^N] = e^{-\lambda} e^{\lambda z}$.

- Hence, with $\mu_1 := \mathsf{E}(N_1) = \lambda_1 t_1$,

$$p_1 \mathsf{E}[W - \mathbf{z}^T M^{N_1} \mathbf{w}] = p_1 [W - e^{-\mu_1} \mathbf{z}^T e^{\mu_1 M} \mathbf{w}],$$

where the matrix exponent $e^A := \sum_{n=0}^{\infty} A^n / n!$.

- Similarly, for the next interval $(t_1, t_2]$, with $\mu_2 := \mathsf{E}(N_1 + N_{12})$, and p_{12} denoting the average price, we can derive the expected revenue collected as follows:

$$\begin{aligned} & p_{12} \mathsf{E}[\mathbf{z}^T M^{N_1} \mathbf{w} - \mathbf{z}^T M^{N_1 + N_{12}} \mathbf{w}] \\ = \quad & p_{12} [\mathbf{z}^T (e^{-\mu_1} e^{\mu_1 M} - e^{-\mu_2} e^{\mu_2 M}) \mathbf{w}]. \end{aligned}$$

- Putting together the above and rearranging terms, with $\pi_k := p_{1k} - p_{1,k+1}$, and μ_ℓ as decision variables, we have

$$\min \sum_{\ell=1}^{m} \pi_\ell G(\mu_\ell); \qquad \text{s.t.} \sum_{\ell=1}^{m} \left(\frac{1}{\Lambda_\ell} - \frac{1}{\Lambda_{\ell+1}}\right)\mu_\ell = T, \qquad \mu_\ell \geq 0,$$

where $G(\mu) := \mathbf{z}^T e^{-\mu} e^{\mu M} \mathbf{w}$, $\Lambda_\ell := \lambda_1 + \cdots + \lambda_\ell$, and

$$\Lambda_\ell (t_\ell - t_{\ell-1}) = \mu_\ell - \mu_{\ell-1}, \qquad \ell = 1, ..., m.$$

- This is a separable convex programming problem, since

$$G'(\mu) = \mathbf{z}^T e^{-\mu}(M - I) e^{\mu M} \mathbf{w} \leq 0,$$

$$G''(\mu) = \mathbf{z}^T e^{-\mu}(M - I)^2 e^{\mu M} \mathbf{w} \geq 0.$$

Asymptotic Optimality

- The objective value under the optimal policy (from dynamic programming) satisfies:

$$V^* \leq (\lambda_1 p_1 + \cdots + \lambda_k p_k)T + \lambda_{k+1} p_{k+1} t,$$

where $k = 1, ..., m-1$ and $t \in [0, T)$ are such that

$$\Lambda_k T \leq W, \qquad \Lambda_{k+1} T > W, \qquad \Lambda_k T + \lambda_{k+1} t = W.$$

(The upper bound on V^* results from a policy with *clairvoyance* or *anticipative*.)

If $\Lambda_1 T > W$, then $V^* \leq p_1 W$.

If $\Lambda_m T \leq W$, then $V^* \leq p_{1m} \Lambda_m T$.

- Let $V^{\rm sw}$ denote the objective value under the best switch-over policy. Then, with k and t as above,

$$\begin{aligned} V^{\rm sw} &\geq p_{1k}W - (p_{1k} - p_{1,k+1})\mathsf{E}[W - N(\Lambda_k(T-t))]^+ \\ &\quad -p_{1,k+1}\mathsf{E}[W - N(\Lambda_k T + \lambda_{k+1}t)]^+. \end{aligned}$$

(The lower bound corresponds to a feasible policy that accepts the top k prices throughout the horizon and the $k+1$-st over the last t time units.)

If $\Lambda_1 T > W$, then $V^{\rm sw} \geq p_1 \mathsf{E}[N(W) \wedge W]$.

If $\Lambda_m T \leq W$, then $V^{\rm sw} \geq p_{1m} \mathsf{E}[N(\Lambda_m T) \wedge W]$.

- The switch-over policy is asymptotically optimal, in the sense that $V^{\rm sw}/V^* \to 1$ when $W \to \infty$.

- To show the asymptotic optimality of the switch-over policy and compare it with another policy with equally spaced switching times, consider batch size following negative binomial distributions, with a range of W values.

W	(r,p)	Optimal	Switch	% off	Equal	% off
20	(4,0.33)	17.59	17.47	0.67%	17.34	1.37%
	(8,0.5)	17.68	17.50	0.99%	17.28	2.27%
40	(4,0.33)	34.39	34.16	0.67%	33.99	1.18%
	(8,0.5)	34.60	34.26	0.97%	34.14	1.34%
60	(4,0.33)	50.08	49.62	0.92%	48.74	2.69%
	(8,0.5)	50.39	49.84	1.09%	49.14	2.48%
160	(4,0.33)	101.43	100.87	0.55%	72.85	28.18%
	(8,0.5)	102.38	101.84	0.52%	72.72	28.97%
180	(4,0.33)	104.54	104.37	0.16%	72.69	30.47%
	(8,0.5)	105.48	105.31	0.16%	73.14	30.66%
200	(4,0.33)	105.74	105.69	0.05%	72.73	31.20%
	(8,0.5)	105.78	105.76	0.01%	72.56	30.82%

Optimal Pricing

- $T = m$: m equal time segments, each of unit length.
- $p_1 \geq p_2 \geq \cdots \geq p_m$: the original price p_1 is given, others are decision variables.
- Want to maximize the following objective:

$$\begin{aligned} & p_1[W - G(\mu_1)] + \cdots + p_m[G(\mu_{m-1}) - G(\mu_m)] \\ = \;& p_1 W_1 - (p_1 - p_2)G(\mu_1) - \cdots - (p_{m-1} - p_m)G(\mu_{m-1}) \\ & \quad - p_m G(\mu_m). \end{aligned}$$

- Let the new decision variables be

$$r_m = p_m; \qquad r_i = p_i - p_{i+1}, \quad i = 1, ..., m-1;$$

Then, the optimization problem becomes:

$$\min \sum_{i=1}^{m} r_i G(\mu_i), \qquad \text{s.t.} \quad r_1 + \cdots + r_m = p_1.$$

- Arrival rate is a function of the price: $\lambda(p)$; for instance,

$$\lambda(p) = a - bp, \qquad \lambda(p) = ae^{-bp}, \qquad \lambda(p) = \frac{a}{p^b};$$

where a and b are positive parameters.

- Optimality equations are (recall $p_i = r_i + \cdots + r_m$):

$$G(\mu_j) + \sum_{i=1}^{m} r_i G'(\mu_i)\Lambda'_{i\wedge j} = \eta, \qquad j = 1, ..., m;$$

where

$$\mu_i = \lambda(p_1) + \cdots + \lambda(p_i) := \Lambda_i;$$

and η is the Lagrangian multiplier.

- Hence, $r_1 = p_1 - r_2 - \cdots - r_m$, and

$$r_j = \frac{G(\mu_{j-1}) - G(\mu_j)}{G'(\mu_j)\lambda'_j} - \sum_{k=j+1}^{m} r_k \frac{G'(\mu_k)}{G'(\mu_j)} \qquad j = 2, ..., m.$$

Optimal Pricing: Examples

- Pricing under three types of demand functions: linear, exponential, and power; for $m = 8$ and $m = 3$; $p_1 = 1$:

$(W; a, b)$	$i =$	1	2	3	4	5	6	7	8	obj. val.
(40; 14, 15)	p_i	1	0.69	0.69	0.69	0.69	0.69	0.68	0.62	25.64
(40; 37.33,, 40)	p_i	1	0.63	0.61						22.94
(40; 2, 15)	p_i	1	0.60	0.60	0.60	0.60	0.60	0.59	0.56	20.77
(40; 2, 40)	p_i	1	0.55	0.54						19.95
(40; 1.5, 5.33)	p_i	1	0.57	0.57	0.57	0.57	0.57	0.54	0.37	20.46
(40; 2, 40)	p_i	1	0.52	0.44						20.10

- Price reduction with respect to the available inventory W: exponential demand as above, fix $m = 8$ and $(a, b) = (15, 2)$, while changing the W values:

$(W; a, b)$	$i =$	1	2	3	4	5	6	7	8	obj. val.
(50; 15, 2)	p_i	1	0.52	0.52	0.52	0.52	0.52	0.52	0.52	21.28
(30; 15, 2)	p_i	1	0.73	0.73	0.73	0.73	0.73	0.71	0.61	19.35
(25; 15, 2)	p_i	1	0.82	0.82	0.82	0.82	0.81	0.78	0.65	18.12
(20; 15, 2)	p_i	1	0.93	0.93	0.93	0.93	0.92	0.86	0.69	16.43
(15; 15, 2)	p_i	1	1	1	1	1	1	1	0.78	14.00
(10; 15, 2)	p_i	1	1	1	1	1	1	1	0.88	9.90

- Non-optimal pricing schemes under the above cases ($m = 3$):

$(W; a, b)$	$i =$	1	2	3	obj. val.	% off opt
(40; 37.33, 40)	p_i	1	1	1	8.00	65%
(40; 2, 40)	p_i	1	0.67	0.33	19.00	5%
(40; 2, 40)	p_i	1	1	0.50	18.21	10%

Concluding Remarks

- Exploiting problem structure leads to switch-over policies that are near optimal, and asymptotically optimal.

- Optimal switch-over times and optimal prices are readily derived through convex programming.

- Extensions:

  - p_1 can be a decision variable as well;
  - demand depends on time (as well as price);
  - incorporating replenishment decisions and service measures;
  - two-dimensional switch: remaining quantity as well as remaining time.