An Approximate Dynamic Programming Approach to Network Revenue Management

Vivek F. Farias *   Benjamin Van Roy †

November 26, 2006

PRELIMINARY DRAFT

Abstract

We use the linear programming approach to approximate dynamic programming as a means of approximating the value function for the dynamic capacity allocation problem. We study the use of a separable concave approximation architecture and show how this architecture may be used in the context of models with very general stochastic arrival processes. We establish via computational experiments that our algorithm achieves significant performance gains (of up to about 8%) relative to a popular benchmark algorithm (the deterministic linear program).

1 Introduction

Network Revenue Management is a generic term for a menagerie of revenue management problems wherein a vendor who is endowed with limited quantities of several resources must sell products, each composed of one or more resource type, over some finite sales season so as to maximize revenues. Revenue management for the airline industry is perhaps the most notable source for such problems: An airline typically operates flights on each leg of a network of cities and offers for sale 'fare products' composed of seats on one or more of these legs. Each fare product is associated, among other things, with some fixed price which the airline receives upon its sale. Since demand for fare products is stochastic and capacity on each leg limited, the airline's problem becomes one of deciding which of its fare products to offer for sale at each point in time of some finite sales period so as to maximize expected revenues. This paper is concerned with a new algorithmic approach to this widely studied problem.

For most models of interest, the dynamic capacity allocation problem we have alluded to can be cast as a dynamic program, albeit one with a computationally intractable state-space for even moderately large networks. As such, revenue management techniques have typically resorted to heuristic control strategies. Early heuristics for the problem were based primarily on the solutions to a set of single resource problems

*Electrical Engineering, Stanford University, email: vivekf@stanford.edu.
†Management Science & Engineering and Electrical Engineering, Stanford University, email: bvr@stanford.edu.

solved for each leg. The state of the art technique involves 'bid-price' controls. A generic bid-price control scheme might work as follows: At each point in time the scheme generates a bid-price for a seat or unit of capacity on each leg of the network. A request for a particular fare at that point in time is then accepted if and only if the revenue garnered from the sale is no smaller than the sum of the bid prices of the resources or seats that constitute that fare. There is a vast array of available algorithms that may be used in the generation of bid-prices. There are two important dimensions along which such an algorithm must be evaluated. One, of course, is revenues generated from the strategy. Since bid-prices must be generated in real time, a second important dimension is the efficiency of the procedure used to generate them. A simple elegant solution to this problem which has found wide-spread acceptance involves the solution of a single linear program referred to as the 'deterministic' LP (DLP). This approach and associated bid-price techniques have found widespread use in modern RM systems and are believed to have generated incremental revenues on the order of 1-2 % greater than previously used single-resource type heuristics.

We take an approximate dynamic programming (DP) approach to this problem. In particular, we propose an algorithm that attempts to generate a good approximation to the optimal value function for the dynamic capacity allocation problem within the span of concave functions of capacity on individual network legs, using the linear programming approach to approximate DP ([2, 3]). The revenue manager then acts greedily with respect to to this approximate value function over the sales season. For the approximation architecture we propose to use, this control policy can be interpreted as a bid-price policy for which generating bid prices at each point in time is equivalent to a table look-up, making it significantly quicker to execute than the DLP technique.

Our algorithm applies to stochastic customer arrival processes where arrival rates can themselves be stochastic. This represents a substantial generalization of the deterministic arrival rate arrival process models considered in the literature and allows for the integration of a broad class of demand forecasting models. We demonstrate via a sequence of computational examples that our algorithm consistently produces higher revenues than a strategy using bid-prices computed via re-solution of the DLP at each time step. While the performance gain relative to the DLP is modest ($\sim 1\%$) for a model with time homogenous arrivals, this gain increases significantly when arrival rates vary stochastically. Even for a very simple stochastic arrival rate process (where arrival rates can be in one of three modes), we report relative performance gains of between up to about 8 %.

Computing our approximate value function entails the solution of a linear program, the approximate linear program (ALP), at the start of the sales season. We study structural properties possessed by the optimal solution to the ALP. In particular, we show that the ALP generates a provably stronger approximation to the optimal value function than does the DLP (which produces an upper bound to the optimal value function) and also demonstrate that the optimal solution to the ALP assigns marginal seat values that are non-neagtive and non-decreasing in time. Practical solution of the ALP entails the use of a constraint sampling procedure; enforcing these constraints explicitly, strengthens the constraint sampling procedure.

The literature on both general dynamic capacity allocation heuristics, as well as bid-price controls is vast and predominantly computational; the book [9] provides an excellent review. Closest to this work is the paper by Adelman, [1], which also proposes an approximate DP approach to computing bid prices via an *affine* approximation to the value function. We make a simple observation that allows an exponential reduction in the number of constraints for the corresponding ALP. In spite of affine approximation being a computationally attractive approximation architecture, our computational experiments suggest that affine approximations are not competitive with an approach that uses bid-prices computed via re-solution of the DLP at each time step. There are broadly two classes of schemes used in the generation of bid prices. One class of schemes is based on mathematical programming formulations of essentially static versions of the problem. The DLP approach is representative of this class and apparently the method of choice in practical applications [9]. This class of approaches generates incremental revenues of approximately 1% over single resource based techniques [8, 7] in realistic simulations. A second class of approaches is based on 'decompositions' of the value function. These decompositions are largely ad-hoc. Examples include the pro-rated EMSR scheme and the DAVN scheme [9]. These schemes generate gains in the vicinity of 0.5 % over single resource based techniques [7, 8] and continue to have a following among practitioners [4]. A concise description for both classes of schemes is provided in [9]. There is also an emerging literature on optimization techniques for models that incorporate some manner of demand forecasting. A recent example is the paper by Mishra, [6], that evaluates various stochastic programming techniques for a linear (with additive noise) model of demand evolution.

The remainder of this paper is organized as follows: In section 2, we formally specify a model for the dynamic capacity allocation problem. In section 3 we review the benchmark DLP heuristic. Section 4 presents an ADP approach to the dynamic capacity allocation problem and specifies our approximation architecture. That section also discusses some simple structural properties possessed by our approximation to the value function. Section 5 presents a series of computational examples comparing the performance of our algorithm with the DLP approach as also an approach based on an affine approximation to the value function. Section 6 concludes.

2 Model

We consider an airline operating L flight legs. The airline may offer up to F fares for sale at each point in time. Each fare f is associated with a price p_f and requires seats on one or more legs. A matrix $A \in \mathbb{Z}_+^{L\times F}$ encodes the capacity on each leg consumed by each fare: $A_{l,f} = k$ if and only if fare f requires k seats on leg l. For concreteness we will restrict attention to the situation wherein a given fare can consume at most 1 seat on any given leg although our discussion and algorithms carry over without any change to the more general case. Initial capacity on each leg is given by a vector $x_0 \in \mathbb{Z}_+^L$. Time is discrete. We assume a T period horizon with at most one customer arrival in a single period. A customer for fare product f arrives in the tth period with probability $\lambda_f(m_t)$. Here $m_t \in \mathcal{M}$ (a finite set) and represents the current demand 'mode'. m_t evolves according to a discrete time Markov process on $\mathcal{M}$ with transition kernel P_t. We note that the discrete time arrival process model we have described may be viewed as a uniformization of an appropriately defined continuous time arrival process. At the start of the tth period the airline must decide which subset of fares from the set $\{f : A_f \preceq x_t\}$ it will offer for sale; an arriving customer for fare f is assigned that fare should it be available, the airline receives p_f, and $x_{t+1} = x_t - A_f$.

We define the *state-space* $\mathcal{S} = \{x : x \in \mathbb{Z}_+^L, x \preceq x_0\} \times \{1, 2, \ldots, T\} \times \mathcal{M}$. Encoding the fares offered for sale at time t by a vector in $\{0,1\}^F \equiv \mathcal{A}$, a control policy is a mapping $\pi : \mathcal{S} \to \mathcal{A}$ satisfying $A\pi(s) \leq x(s)$ for all $s \in \mathcal{S}$. Let Π represent the set of all such policies. Let $R(s,a)$ be a random variable representing revenue generated by the airline in state $s \in \mathcal{S}$ when fares $a \in \mathcal{A}$ are offered for sale, and define

$$J^\pi(s) = E_\pi \left[\sum_{t'=t(s)}^{T} R(s_{t'}, \pi(s_{t'})) \big| s_{t(s)} = s \right].$$

We let $J^*(s) = \max_{\pi \in \Pi} J^\pi(s)$, denote the expected revenue under the optimal policy π^* upon starting in state s.

J^* and π^* can, in principle, be computed via Dynamic Programming. In particular, define the dynamic programming operator T for $s \in \{s' : t(s') < T\}$ according to

$$\begin{aligned}(TJ)(s) = &\sum_{f: A_f \leq x(s)} \lambda_f(m(s)) \max \left[p_f + E\left[J\left(S'_f\right)\right], E\left[J\left(S'\right)\right] \right] \\ &+ \Big(1 - \sum_{f: A_f \leq x(s)} \lambda_f(m(s))\Big) E\left[J(S')\right]. \end{aligned} \tag{1}$$

where $S'_f = (x(s) - A_f, t(s)+1, m_{t(s)+1})$ and $S' = (x(s), t(s)+1, m_{t(s)+1})$. We defined $(TJ)(s) = 0$ for all $s \in \{s' : t(s') = T\}$. J^* may then be identified as the unique solution to the fixed point equation $TJ = J$ that satisfies $J(s) = 0$ for all $s \in \{s' : t(s') = T\}$. π^* is then the policy that achieves the maximum in (1).

We will focus on three special cases of the above model:

- (M1) Time homogenous arrivals: Here we have $|\mathcal{M}| = 1$. That is the arrival rate of customers for the various fare products is constant over time and the arrival process is un-correlated in time.

- (M2) Multiple demand modes, deterministic transition time: Here we consider a model with $\mathcal{M} = \{\text{med,hi,lo}\}$. We have $m_t = \text{med}$ for $t \leq T/2$. With probability $\tilde{p}$, $m_t = \text{lo}$ for all $t > T/2$ and with probability $1 - \tilde{p}$ and $m_t = \text{hi}$ for all $t > T/2$. This is representative of a situation where there is likely to be a change in arrival rates at some known point during the sales season. The revenue manager has a probabilistic model of what the new arrival rates are likely to be.

- (M3) Multiple demand modes, random transition time: Here we consider a model with $\mathcal{M} = \{\text{med, hi,lo}\}$, with the transition kernel P_t defined according to

$$P_t(m_{t+1} = y | m_t = x) = \begin{bmatrix} 1-q & q\tilde{p} & q(1-\tilde{p}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{xy}.$$

where $q, \tilde{p} \in (0, 1)$. This arrival model is similar to the second with the exception that instead of a change in demand modes occuring at precisely $t = T/2$, there is now uncertainty in when this transition will occur. In particular, the transition time is now a geometric random variable with expectation $1/q$.

3 Benchmark Heuristic: The Deterministic LP (DLP)

The Dynamic Programming problem we have formulated is computationally intractable and so one must resort to various sub-optimal control strategies. We review the DLP-heuristic for generating bid prices. This heuristic makes the simplifying assumption that demand is deterministic and equal to its expectation. In doing so, the resulting control problem reduces to the solution of a simple LP (the DLP) and the optimal control policy is static. In particular, if demand for fare class f over a $T - t$ period sales season, $D_{t,f}$, were deterministic and equal to expected demand, $\mathbb{E}[D_{t,f}|m_t]$, the maximal revenue that one may generate with an initial capacity $x(s)$ is given by the optimal solution to the DLP:

$$\begin{array}{lll} DLP(s): & \max & p'z \\ & \text{s.t.} & Az \leq x(s) \\ & & 0 \leq z \leq \mathbb{E}[D_{t(s)}|m_{t(s)} = m(s)] \end{array}$$

Denote by $r^*(s)$ a vector of optimal shadow prices corresponding to the constraint $Az \leq x(s)$ in $DLP(s)$. The bid price control policy based on the DLP solution is then given by:

$$\pi^{\text{DLP}}(s)_f = \begin{cases} 1 & \text{if } A_f' r^*(s) \leq p_f \text{ and } A_f \leq x(s) \\ 0 & \text{otherwise} \end{cases}$$

The above description of the DLP heuristic assumes that the shadow prices r^* are recomputed at each time step. While this many not always be the case, a general computational observation cf. [9] is that frequent re-computation of r^* improves performance. This is consistent with our computational experience.

In the case of model M2, one might correctly point out that a simple modification of the DLP is likely to have superior performance. In particular, one may consider retaining the probabilistic structure of the demand mode transition model and solving a single stage stochastic program with recourse variables for capacity allocation in the event of a transition to the hi and lo demand modes respectively. We do not consider such a stochastic programming approach as it is intractable except for very simple models (such as M2); for more general models, the number of recourse variables grows exponentially with horizon length.

4 Bid Price Heuristics via Approximate DP

Given a component-wise positive vector c, the optimal value function J^* may be identified as the optimal solution to the following LP:

$$
\begin{array}{ll}
\max & c'J \\
\text{s.t.} & (TJ)(s) \leq J(s) \quad \forall s \in \mathcal{S} \\
& J(s) = 0 \quad \forall s \in \{s' : t(s') = T\}
\end{array}
$$

The linear programming approach to approximate DP entails adding to the above LP, the further constraint that the value function J lie in the linear span of some set of basis functions $\phi_i : \mathcal{S} \to \mathbb{R},\ i = 1, 2, \ldots, k$. Encoding these functions as a matrix $\Phi \in \mathbb{R}^{|\mathcal{S}| \times k}$, the approximate LP (ALP) computes a vector of weights $r \in \mathbb{R}^k$ that optimally solve:

$$
\begin{array}{ll}
\max & c'\Phi r \\
\text{s.t.} & (T\Phi r)(s) \leq (\Phi r)(s) \quad \forall s \in \mathcal{S} \\
& (\Phi r)(s) = 0 \quad \forall s \in \{s' : t(s') = T\}
\end{array}
$$

Given a solution r^* to the ALP (assuming it is feasible), one then uses a policy that is greedy with respect to Φr^*. Of course, the success of this approach depends crucially upon the choice of the set of basis functions Φ. In the next two subsections we examine affine and concave approximation architectures. The affine approximation architecture for the network RM problem was proposed by Adelman [1] in the context of the M1 model. The concave architecture is the focus of this paper. In the sequel we assume that $c_{s_0} = 1$ and that all other components of c are 0.

4.1 Separable Affine Approximation

Adelman [1] studies considers the use of affine basis functions in the M1 model. In particular, [1] explores the use of the following set of $(L+1)T$ basis functions defined according to

$$
\phi_{l,t}(x, t') = \begin{cases} x_l & \text{if } l \leq L \text{ and } t = t' \\ 1 & \text{if } l = L+1 \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}
$$

The ALP here consequently has $(L+1)T$ variables but $|\mathcal{S}|F$ constraints. [1] proposes the use of a column generation procedure to solve the ALP. We can, in fact, show that the ALP can be reduced to an LP with $(L+1)T$ variables and $2^L TF$ constraints making practical solution of the ALP to optimality possible for relatively large networks [1]. This is the content of the following lemma whose proof may be found in the appendix.

Lemma 1. *A vector r is a feasible solution to the ALP for the M1 model with affine approximation if and only if the following set of constraints is satisfied*

$$
(T\Phi r)(x, t) \leq (\Phi r)(x, t)\ \forall t, \forall x \in \{0, 1\}^L
$$

[1]The largest computational examples in [1] for example would be reduced to LPs with 1 million constraints, 10 thousand variables and 20 non-zeros per constraint, which can very effectively be solved on a personal computer

In spite of being a computationally attractive approximation architecture, affine approximations have an obvious weakness: the greedy policy with respect to an affine approximation to the value function is insensitive to intermediate capacity levels so that the set of fares offered for sale at any intermediate point in time depends only upon the time left until the sales season ends. In particular the greedy policy with respect to an affine approximation, π^{aff}, will satisfy $\pi^{\mathrm{aff}}(x,t) = \pi^{\mathrm{aff}}(\tilde{x},t)$ provided x and $\tilde{x}$ are positive in identical components. We observe in computational experiments that a policy that is greedy with respect to an affine approximation to the value function is in fact not competitive with a policy based on re-computation of bid-prices at each time step via the DLP. While one possible approach to consider is frequent re-solution of the ALP with affine approximation, this is not a feasible option given that bid-prices must often be generated in real time. The affine approximation is incapable of capturing the concavity of the J^* in the remaining inventory x. This motivates us to consider a separable concave approximation architecture which is the focus of this paper.

4.2 Separable Concave Approximation

Consider the following set of basis functions, $\phi_{i,l,t,m}$, defined for integers $i \in [0,(x_0)_l]$, $l \in [1,L]; t \in [1,T]$, and $m \in \mathcal{M}$ according to:

$$\phi_{i,l,t,m}(x',t',m') = \begin{cases} 1 & \text{if } x'_l = i \text{ , } t = t' \text{ and } m = m' \\ 0 & \text{otherwise} \end{cases}$$

The ALP in this case will have $(1'x_0 + L)T|\mathcal{M}|$ variables and $|\mathcal{S}|F$ constraints. Note that optimal solution is intractable since $|\mathcal{S}|$ is exponentially large. One remedy is the constraint sampling procedure in [3] which requires sampling constraints from $\mathcal{S}$ according to the state-distribution induced by an optimal policy. Assuming a sales season of T periods and an initial inventory of x_0, we propose using the following procedure with parameter N:

1. Simulate a bid price control policy starting at state $s_0 = (x_0, 0, m_0)$, using bid prices generated by re-solving the DLP at each time step. Let $\mathcal{X}$ be the set of states visited over the course of several simulations. We generate a set with $|\mathcal{X}| = N$

2. Solve the following Relaxed LP (RLP):

$$\begin{array}{lll} \max & (\Phi r)(s_0) & \\ \text{s.t.} & (T\Phi r)(s) \le (\Phi r)(s) & \text{for } s \in \mathcal{X} \\ & r_{i,l,t,m} \ge r_{i+1,l,t,m} & \forall i,l,t,m \\ & r_{i,l,t,m} = 0 & \forall i,l,m; t = T \end{array}$$

3. Given a solution r^* to the RLP, use the following control policy over the actual sales season:

$$\pi^{\text{con}}(s)_f = \begin{cases} 1 & \text{if } \sum_{l:A_{l,f}=1} r^*_{x(s)_l,l,t(s),m(s)} \leq p_f \text{ and } A_f \leq x(s) \\ 0 & \text{otherwise} \end{cases}$$

Several comments on the above procedure are in order. Step 1 in the procedure entails choosing a suitable number of samples N; [3] provides some guidance on this choice. Our choice of N was heuristic and is described in the next section. Step 2 of the procedure entails solving the RLP whose constraints are samples of the original ALP. We will shortly prove several simple structural properties that an optimal solution to the ALP must posses. Adding these constraints to the RLP strengthens the quality of our solution. Also, note that the inequality constraints on the weights enforce concavity of the approximation. Finally note that the greedy policy with respect to the our approximation to J^* takes the form of a bid price policy as in the case of affine approximation. However, unlike affine approximation the resulting policy decisions depend on available capacity as well as time.

4.3 Structural Properties of the ALP Solution

The optimal solution to the DLP provides an upper bound to the true value function J^*, i.e. $DLP(s) \geq J^*(s)$. There are several proofs of this fact for the time homogenous model M1. For example, see Gallego and Van Ryzin [5] or Adelman [1]. The DLP continues to be an upper bound to the true value function for the more general model we study here (via a simple concavity argument and the use of Jensen's inequality). We can show that the ALP with separable concave approximation provides a tighter upper bound than does the DLP for model M2, and generalizations to M2 which allow for more than a single branching time. The same result for time homogenous arrival rates (i.e. for model M1) follows as a corollary. We are at present unable to establish such a result for the general model.

Lemma 2. *For model M2 with initial state s, $J^*(s) \leq ALP(s) \leq DLP(s)$*

The proof of the lemma can be found in the appendix. The above result is not entirely conclusive. In particular, while it is indeed desirable to have a good approximation to the true value function, a tighter approximation does not guarantee an improved policy. Nonetheless, stronger approximations to the true value function imply stronger bounds on policy performance.

The optimal solution to the ALP also shares two simple structural properties with the optimal value function for time homogenous arrivals (i.e. model M1):

Lemma 3. *The marginal value of a seat under an optimal ALP solution is non-negative and non-decreasing in the sales horizon when arrival rates are constant. Let r be any feasible solution to the ALP, and r^* be an optimal solution. Then, for model M1, $r^*_{i,l,t} \geq 0$ and $\sum_{i=0}^{x_l} r_{i,l,t} \leq \sum_{i=0}^{x_l} r_{i,l,t-1}$ for all i, l, t and $0 \leq x \leq x_0$.*

The proof of this lemma can be found in the appendix. The second property need not hold in general;

counterexamples are simple to construct. We explicitly enforce the constraints implied by Lemma 3 in our computational experiments.

5 Computational Results

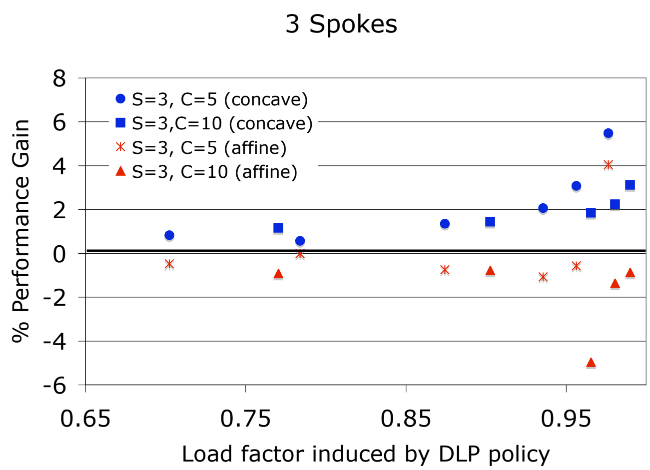

Figure 1: Performance relative to the DLP for model M1

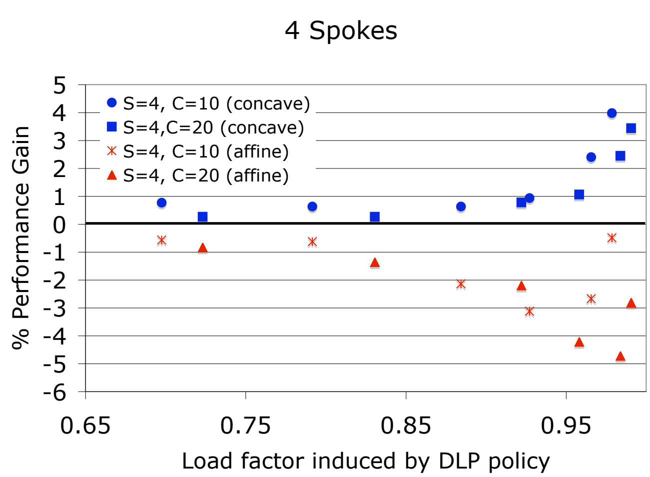

Figure 2: Performance relative to the DLP for model M1

It is difficult to establish theoretical performance guarantees for our algorithm. Indeed, we are unaware of any algorithm for the dynamic capacity allocation problem for which non-asymptotic theoretical performance guarantees are available. As such, we will establish performance merits for our algorithm via a computational study. We will consider two simple test networks each with a single 'hub' and either three or four spoke cities. This topology is representative of actual airline network topologies. Each leg in our network represents two separate aircraft (one in each direction) making for a total of $f = 15$ itineraries on

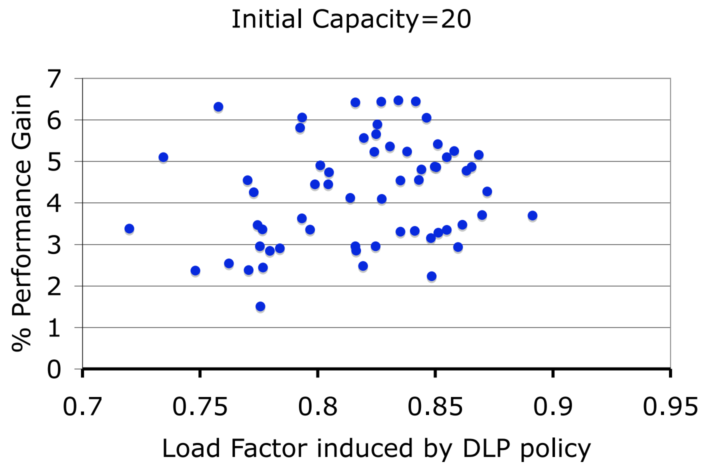

Figure 3: Performance relative to the DLP for model M2

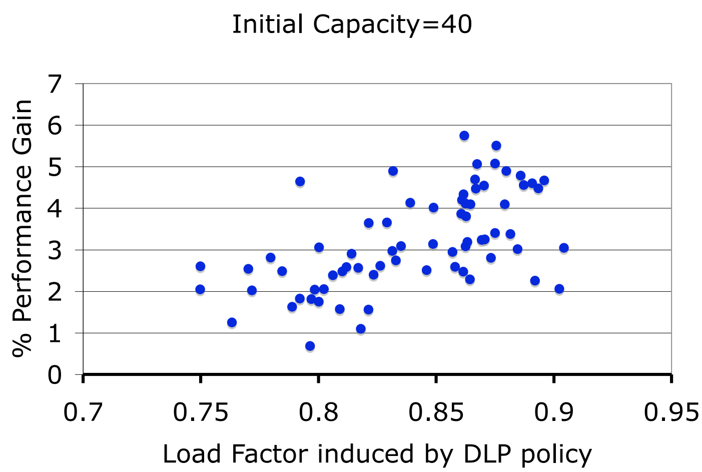

Figure 4: Performance relative to the DLP for model M2

the 3 spoke network and $f = 24$ itineraries on the 4 spoke network. Arrival rates for each itinerary, demand mode i.e. (f, m) pair were picked randomly from the unit f-dimensional simplex and suitably normalized. Route prices were generated uniformly in the interval $[50, 150]$ for single leg routes and $[50, 250]$ for two leg routes. We consider a random instantiation of arrival rates and probabilities for each network topology and for each instantiation measure policy performance upon varying initial capacity levels and sales horizon. We compare performance against the DLP with re-solution at each time step. In the case of model M1, we also include policies generated via the separable affine approximation architecture in our experiments. We solve RLPs with $50,000$ sampled states, this number being determined by memory constraints. We now describe in detail our experiments and results for each of the three models.

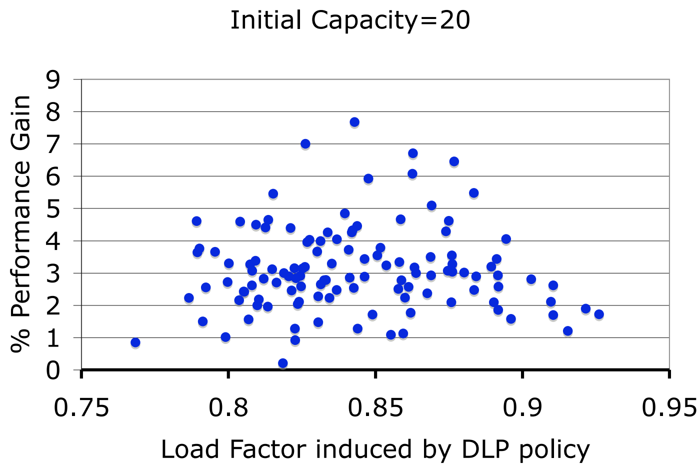

Figure 5: Performance relative to the DLP for model M3

5.1 Time homogenous arrivals (M1)

We consider three and four spoke models. The arrival probabilities for each fare class were drawn uniformly at random on the unit simplex and normalized so that the probability of no customer arrival in each period was 0.7. For both models, we consider fixed capacities (of 5,10 and 20) on each network leg and vary the sales horizon T. For each value of T we record the average load-factor (i.e. the average fraction of seats sold) under the DLP policy; we select values of T so that this induced load factor is > 0.7. We plot in Figures 1 and 2 the performance of the the ADP based approaches with affine and separable approximations relative to the DLP heuristic for two different initial capacity levels. The x-intercept for a data point in both plots is the average load factor induced by the DLP heuristic for the problem data in question at that point.

The plots suggest a few broad trends. The affine approximation architecture is almost uniformly dominated by the DLP heuristic when the DLP is re-solved at every time step, while the separable architecture uniformly dominates both heuristics in every problem instance. We note that since a bid price computation in the ADP approach is simply a lookup it is far quicker than solving the DLP, so that together these facts support the plausability of using an ADP approach with separable approximation. Another trend is performance gain. This is actually quite low at low induced load factors ($< 0.5\%$) but can be as high as 5% at high load factors. At moderate load factors (that are at least nominally representative) the performance gain is on the order of 1%. We anticipate the gain to be larger for more complex networks.

It is difficult to expect higher performance gains than we have observed for the M1 demand model. In particular, at low load factors, the problem is trivialized (since it is optimal to accept all requests). Moreover, it is well known (see [5]) that in a certain fluid scaling (which involves scaling both initial capacity x_0 and sales horizon T by some scaling factor N), the DLP heuristic is optimal as N gets large. The purpose of our experiments with this model is to illustrate the fact that the separable affine approximations we employ are robust in this simple demand setting.

5.2 Multiple demand modes (M2, M3)

Model M1 is potentially a poor representation of reality. This leads us to consider incorporating a demand forecasting model such as that in models M2 and M3. In our experiments, the arrival probabilities for each demand mode were drawn uniformly at random on the unit 24-dimensional simplex and normalized so that the probability of no customer arrival in each period was 0.55 for the 'med' demand mode, 0.7 for the 'lo' mode, and 0.1 in the 'hi' mode. The probability of transitioning from the med to lo demand mode, p, was set to 0.5 in both models, and we set $m_0 =$ med.The probability of transitioning out of the med demand state, q, was set to $2/T$ in model M3. The sales horizon T was varied so that the load-factor induced by the DLP policy was approximately between 0.8 and 0.9. We generate a random ensemble of 40 such problems for a network with 4 spokes and consider initial capacity levels of 20 seats and 40 seats. We measure the performance gain of our ADP with separable concave approximation derived bid price control over the DLP. The DLP is resolved at every time step so that it may recompute expected total remaining demand for each fare class conditioned on the current demand mode.

For model M2, we plot in Figures 3 and 4 the performance of the the ADP based approach with separable concave approximation relative to the DLP heuristic with initial capacity levels of 20 and 40 respectively. We note that the relative performance gain here is significant (up to about 8%) in a realistic operating regime. In the case of model M3, Figure 5 illustrates similar performance trends.

We see that the approximate DP approach with concave approximation offers substantiative gains over the use of the DLP even with very simple stochastic variation in arrival rates. We anticipate that these gains will be further amplified for more complex models of arrival rate variability (for example in models with a larger number of demand modes etc.).

6 Conclusion

We have explored the use of separable concave functions for the approximation of the optimal value function for the dynamic capacity allocation problem. The approximation architecture is quite flexible and we have illustrated how it might be employed in the context of a general arrival process model wherein arrival rates vary stochastically according to a Markov process. Our computational experiments indicate that the use of the LP approach to Approximate DP along with this approximation architecture can yield significant performance gains over the DLP (of up to about 8%) , even when re-computation of DLP bid prices is allowed at every time step. Moreover, our control policy is a bid price policy where policy execution requires a table look-up at every epoch making the methodology ideally suited to real time implementation. State of the art heuristics for the dynamic capacity allocation problem typically resort to using point estimates of demand in conjunction with a model that assumes simple time homogenous arrival processes in order to make capacity allocation decisions dynamically. As such, our algorithm may be viewed as a viable approach to moving beyond the use of point estimates and instead integrating forecasting and optimization.

Several issues remain to be resolved. For example, in the interest of very large-scale implementations, it would be useful to explore the use of simpler basis functions that are nonetheless capable of capturing the concavity of the true value function. A second source of complexity is the number of constraints in the ALP. The concave architecture we use imparts a great deal of structure to the ALP. In the case of affine approximation this structure allows for a dramatic reduction in constraints; it would be useful to explore similar constraint reductions for the ALP with concave approximation.

Acknowledgment

The authors would like to thank Mike Harrison for useful discussions that led to our consideration of some of the models here, as well as his suggestions and pointers to the literature.

References

[1] D. Adelman. Dynamic bid price in revenue management. to appear in Operations Research, 2005.

[2] D.P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51:6:850–865, 2003.

[3] D.P. de Farias and B. Van Roy. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 29:3:462–478, 2004.

[4] I. C. Bilegan E. A. Boyd. Revenue management and e-commerce. Pros-RM white paper. URL: `http://www.prospricing.com`, 2005.

[5] G. Gallego and G. van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.

[6] Victor De Miguel and Nishant Mishra. What multistage stochastic programming can do for network revenue management. London Business School working paper, 2006.

[7] P.P.Belobaba. Revenue and competitive impacts of o-d control: Summary of pods results. In *First Annual INFORMS Revenue Management Section Meeting*, New York, NY, 2001.

[8] P.P.Belobaba and S. Lee. Pods update: Large network o-d control results. In *Proc. 2000 AGIFORS Reservations and Yield Management Study Group Symposium*, New York, NY, 2000.

[9] Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. Springer Science+Business Media, 2004.

A Proofs for Section 3

Lemma 1. *A vector r is a feasible solution to the ALP for the M1 model with affine approximation if and only if the following set of constraints is satisfied*

$$(T\Phi r)(x,t) \leq (\Phi r)(x,t)\ \forall t, \forall x \in \{0,1\}^L$$

Proof: Let $x' \in \{0,1\}^L$. We will show that if $(T\Phi r)(x',t) \leq (\Phi r)(x',t)$, then it must be that $(T\Phi r)(\tilde{x},t) \leq (\Phi r)(\tilde{x},t)$ for all $\tilde{x}$ satisfying $\tilde{x}_l = 0$ iff $x'_l = 0$. Let us denote by r_t the weights corresponding to the basis functions $\phi_{l,t}$ for $l \leq L$ and let θ_t denote the weight for $\phi_{L+1,t}$. From Lemma 3 it follows that any feasible r satisfies $r_t \geq r_{t+1}, \theta_t \geq \theta_{t+1}$. Now,

$$\begin{aligned}
&(T\Phi r)(\tilde{x},t) - (\Phi r)(\tilde{x},t)\\
&= \sum_{f:A_f \leq \tilde{x}} \lambda_f \left(\sum_l (r_{t+1} - r_t)(\tilde{x}_l - A_{f,l}) + \max\left(\sum_{l:A_{f,l}=1} r_{t+1,l}, p_f \right)\right)\\
&\quad + \left(1 - \sum_{f:A_f \leq \tilde{x}} \lambda_f \right)\left(\sum_l (r_{t+1} - r_t)\tilde{x}_l \right) + \theta_{t+1} - \theta_t\\
&\leq \sum_{f:A_f \leq \tilde{x}} \lambda_f \left(\sum_l (r_{t+1} - r_t)(x'_l - A_{f,l}) + \max\left(\sum_{l:A_{f,l}=1} r_{t+1,l}, p_f \right)\right)\\
&\quad + \left(1 - \sum_{f:A_f \leq \tilde{x}} \lambda_f \right)\left(\sum_l (r_{t+1} - r_t)x'_l \right) + \theta_{t+1} - \theta_t\\
&\leq 0
\end{aligned}$$

where the first inequality follows from the fact that $\tilde{x} \geq x'$, $\{f : A_f \leq x'\} = \{f : A_f \leq \tilde{x}\}$, $r_t \geq r_{t+1}$ and $\theta_{t+1} - \theta_t$ and the second inequality is true since $(T\Phi r)(x',t) - (\Phi r)(x',t) \leq 0$. This completes the proof.

Lemma 2. *For model M2 with initial state s, $J^*(s) \leq ALP(s) \leq DLP(s)$*

Proof: Consider the following linear program:

$$\begin{array}{lll}
sDLP(s): & \max & p'z_0 + \Pr(s_{t(s)/2} = \text{lo})p'z_1 + \Pr(s_{t(s)/2} = \text{hi})p'z_2\\
& \text{s.t.} & A(z_0 + z_1) \leq x(s)\\
& & A(z_0 + z_2) \leq x(s)\\
& & 0 \leq z_0 \leq \mathbb{E}[D_{t(s)}] - E[D_{t(s)/2}]\\
& & 0 \leq z_1 \leq \mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{lo}]\\
& & 0 \leq z_2 \leq \mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{hi}]
\end{array}$$

It is clear that $sDLP(s) \leq DLP(s)$. This is because $z_0 + \Pr(s_{T/2} = \text{lo})z_1 + \Pr(s_{T/2} = \text{hi})z_2$ is a feasible solution to $DLP(s)$ of the same value as $sDLP(s)$. We will first show that $ALP(s) \leq sDLP(s)$. The dual to $sDLP(s)$ is given by:

$$\begin{aligned}
\min \quad & x(s)'y_{1,1} + x(s)'y_{1,2} + \tilde{D}_0' y_{2,0} + \tilde{D}_1' y_{2,1} + \tilde{D}_2' y_{2,2} \\
\text{s.t.} \quad & A'(y_{1,1} + y_{1,2}) + y_{2,0} \geq p \\
& A' y_{1,1} + y_{2,1} \geq p \Pr(s_{T/2} = \text{lo}) \\
& A' y_{1,2} + y_{2,2} \geq p \Pr(s_{T/2} = \text{hi}) \\
& y_{1,1}, y_{1,2}, y_{2,0}, y_{2,1}, y_{2,2} \geq 0
\end{aligned}$$

where $\tilde{D}_0 = \mathbb{E}[D_{t(s)}] - E[D_{t(s)/2}]$, $\tilde{D}_1 = \mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{lo}]$ and $\tilde{D}_2 = \mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{hi}]$. Consider the following solution to the ALP for M2: Set

$$r^*_{i,l,t,med} \begin{cases} = (y^*_{1,1})_l + (y^*_{1,2})_l \\ \qquad \text{for } i > 0, t < t(s)/2 \\ = (\mathbb{E}[D_t] - E[D_{t(s)/2}])' y^*_{2,0} + \Pr(s_{t(s)/2} = \text{lo}) r^*_{0,1,t(s)/2,lo} \\ \quad + \Pr(s_{t(s)/2} = \text{hi}) r^*_{0,1,t(s)/2,hi} \\ \qquad \text{for } i = 0, l = 1, t < t(s)/2 \end{cases}$$

$$r^*_{i,l,t,lo} \begin{cases} = (y^*_{1,1})_l / \Pr(s_{t(s)/2} = \text{lo}) & \text{for } i > 0, t \geq t(s)/2 \\ = (\mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{lo}])' y^*_{2,1} / \Pr(s_{t(s)/2} = \text{lo}) & \text{for } i = 0, l = 1, t \geq t(s)/2 \end{cases}$$

$$r^*_{i,l,t,hi} \begin{cases} = (y^*_{1,2})_l / \Pr(s_{t(s)/2} = \text{hi}) & \text{for } i > 0, t \geq t(s)/2 \\ = (\mathbb{E}[D_{t(s)/2} | s_{t(s)/2} = \text{hi}])' y^*_{2,2} / \Pr(s_{t(s)/2} = \text{hi}) & \text{for } i = 0, l = 1, t \geq t(s)/2 \end{cases}$$

It is routinely verified that this solution is in fact feasible for the ALP and has value equal to $sDLP(s)$. The fact that $ALP(s) \geq J^*(s)$ follows from the monotonicity of the T operator and the fact that J^* is the unique fixed point of T. This completes the proof.

Lemma 3. *The marginal value of a seat under an optimal ALP solution is non-negative and non-decreasing in the sales horizon when arrival rates are constant. Let r be any feasible solution to the ALP, and r^* be an optimal solution. Then, for model M1, $r^*_{i,l,t} \geq 0$ and $\sum_{i=0}^{x_l} r_{i,l,t} \leq \sum_{i=0}^{x_l} r_{i,l,t-1}$ for all i, l, t and $0 \leq x \leq x_0$.*

Proof: That $\sum_{i=0}^{x_l} r_{i,l,t} \leq \sum_{i=0}^{x_l} r_{i,l,t-1}$ is in fact an explicit constraint in the ALP. In particular this is precisely the constraint corresponding to allowing no sales in state $(x_l e_l, t-1)$ where e_l is the lth unit vector. One may show via a simple induction on t that $\sum_{i=0}^{(x_l-1)^+} r^*_{i,l,t} \leq \sum_{i=0}^{x_l} r^*_{i,l,t-1}$. The first claim can then also be established via induction on t. In particular, it is clear that $r^*_{i,l,T-1} \geq 0$ for all i, l. Assume the statement true for some $0 < t \leq T-1$ and consider states of the form $(0, t-1)$. By the induction hypothesis $(T\Phi r^*)(0, t-1) \geq 0$. Since $(\Phi r^*)(0, t-1) \geq (T\Phi r^*)(0, t-1)$, we must have $r^*_{0,l,t} \geq 0$ for all l. We now use the fact that $\sum_{i=0}^{(x_l-1)^+} r^*_{i,l,t} \leq \sum_{i=0}^{x_l} r^*_{i,l,t-1}$ to conclude that $r^*_{i,l,t} \geq 0$ for all i, l.