Tractable consideration set structures for assortment optimization and network revenue management

Arne K. Strauss\(^1\) and Kalyan Talluri\(^2\)

\(^1\)Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom.

\texttt{arne.strauss@wbs.ac.uk}

\(^2\)Imperial College Business School, London SW7 2AZ, United Kingdom.

\texttt{kalyan.talluri@imperial.ac.uk}

7 October 2016

Abstract

Discrete-choice models are widely used to model consumer purchase behavior in assortment optimization and revenue management. In many applications, each customer segment is associated with a consideration set that represents the set of products that customers in this segment consider for purchase. The firm has to make a decision on what assortment to offer at each point in time without the ability to identify the customer’s segment. A linear program called the Choice-based Deterministic Linear Program (\textit{CDLP}) has been proposed to determine these offer sets. Unfortunately, its size grows exponentially in the number of products and it is NP-hard to solve when the consideration sets of the segments overlap. The Segment-based Deterministic Concave Program with some additional consistency equalities (\textit{SDCP}+) is an approximation of \textit{CDLP} that provides an upper bound on \textit{CDLP}'s optimal objective value. \textit{SDCP}+ can be solved in a fraction of the time required to solve \textit{CDLP} and often achieves the same optimal objective value. This raises the question under what conditions can one guarantee equivalence of \textit{CDLP} and \textit{SDCP}+. In this paper, we obtain a structural result to this end, namely that if the segment consideration sets overlap with a certain tree structure or if they are fully nested, \textit{CDLP} can be equivalently replaced with \textit{SDCP}+. We give a number of examples from the literature where
this tree structure arises naturally in modeling customer behavior.

**Keywords:** discrete choice models, assortment optimization, network revenue management, consideration sets

Received: May 2015; accepted: December 2016 by Bill Cooper after two revisions.

1 Introduction

This paper is concerned with three closely related problems: The **Assortment Optimization** problem is to determine an assortment of products to maximize the expected revenue from a heterogeneous population of customers. This population is modeled as being made up of multiple latent customer segments. **Revenue management (RM)** is dynamic assortment optimization for products that share a common resource. In **Network Revenue Management (NRM)**, products may use inventory from multiple resources (for example, a hotel stay for three nights uses three days of inventory; airline itineraries involving multiple flights use seats on the connecting flight legs). Network interactions arise as the decision to offer a product depends on the future revenues attainable from the sale of the other products that share resources on the network.

Underlying all three problems is a model of how consumers choose a product to purchase. Talluri and van Ryzin (2004) introduced RM based on a discrete-choice model of customer purchases. Discrete-choice models represent purchase probability as a function of available products and customer characteristics. Such modeling was later applied to NRM by Gallego et al. (2004) and Liu and van Ryzin (2008) who formulated a Choice-based Deterministic Linear Program (CDLP), the main object\(^1\) of study in this paper. The linear program **CDLP** has an exponential number of variables and is difficult to solve except for a few restricted choice models.

In our model, the customer population consists of multiple segments. Each segment is associated with a subset of products called the consideration set, along with segment-specific parameters of the choice model. Assuming that all products are available, the consideration set represents the set of products that customers in that segment would consider for purchase. Consideration sets play a crucial role in the solvability of the assortment optimization problem: If the segments’ consideration sets overlap, the assortment optimization

\(^1\)We present our results in the context of NRM as this problem generalizes both RM as well as assortment optimization. Subsequent work on this problem can be found in Bodea et al. (2009), Kunnumkal and Topaloglu (2010), Zhang and Adelman (2009), Meissner and Strauss (2012), Méndez-Díaz et al. (2014) and Kunnumkal (2014).
problem is NP-hard when there are just two segments, even for simple choice models such as the Multinomial-Logit (MNL) (Bront et al., 2009).

Motivated by this intractability of CDLP, Talluri (2014) develops a weaker formulation called Segment-based Deterministic Concave Program (SDCP), weaker in the sense that its optimal objective value gives an upper bound on the optimal objective value of CDLP. The idea is to solve a collection of small subproblems, each corresponding to a segment, with some constraints that loosely link them together. SDCP is generally poor in approximating CDLP when segment consideration sets overlap, i.e. its optimal objective function value is significantly higher than that of CDLP, and it also performs poorly in revenue simulations. To improve this situation, Meissner et al. (2013) propose an extension of the SDCP formulation called SDCP+ (defined in §2.4) that obtains a significantly tighter relaxation of CDLP for the case of overlapping consideration sets. In their numerical experiments, SDCP+ achieves the CDLP optimal objective function value in many instances despite being faster to solve by an order of magnitude.

The main contribution of this paper is that we identify two consideration set structures for which CDLP is equivalent to SDCP+, i.e. a solution to CDLP can be obtained by instead solving the simpler problem SDCP+ with significantly fewer decision variables in a fraction of computation time. We give a number of applications from the literature where such consideration set structures naturally occur in the modeling of customer behavior. Our results depend only on the consideration set structure and not on the structure of the network of resources and, importantly, apply for any general discrete-choice model.

The remainder of the paper is organized as follows: In §2 we introduce the notation, the demand model, the basic dynamic program and the two approximations of the dynamic program, namely the CDLP and the SDCP+. In §3 we present the main structural results. In §4 we illustrate some applications from the literature where the desired structure of the consideration set is naturally present. In §5 we summarize our conclusions.

2 Models

We introduce the notation in §2.1. In §2.2, we state a dynamic programming formulation of the choice-based NRM problem. Its intractability motivates the formulation of approximations: in §2.3 we define the Choice-
based Deterministic Linear Program (CDLP), and in §2.4 the Enhanced Segment-Based Deterministic Concave Program (SDCP+).

2.1 Notation

A product is a specification of a price (usually with restrictions such as advance purchase requirements) and a set of resources that the product uses. For instance, a product for a network airline would be the combination of fare class (price and restrictions) and the itinerary (the flight legs of the itinerary); in a hotel network, a product is a multi-night stay for a particular room type at a certain price point.

We define a discrete-time booking horizon that consists of $T$ intervals, indexed by $t$. The sale process begins at time 0 and all resources perish instantaneously at time $T$. We make the standard assumption that the time intervals are small enough so that the probability of more than one customer arriving in a time period is negligible.

The underlying network has $m$ resources (indexed by $i$) and $n$ products (indexed by $j$), and we refer to the set of all resources as $I$ and the set of all products as $J$. The resources used by $j$ are represented by a resource-product incidence matrix $A$, with $a_{ij} = 1$ if product $j$ uses resource $i$, and $a_{ij} = 0$ otherwise. Columns of $A$ are the 0-1 incidence vectors $A_j$. We denote the vector of capacities at time $t$ as $c_t$, so the initial set of capacities at time 0 is $c_0$.

We assume that there are $L := \{1, \ldots, L\}$ customer segments, each with distinct purchase behavior. In each period, a customer arrives with probability $\lambda$ and belongs to segment $l$ with probability $p_l$. We denote $\lambda_l = p_l \lambda$ and assume $\sum_{l \in L} p_l = 1$, so $\lambda = \sum_{l \in L} \lambda_l$. We assume time-homogenous arrivals (homogenous in rates and segment mix), but the model and all solution methods in this paper can be clearly extended to the case where rates and mix change by period. Customers in segment $l$ have a consideration set $C_l \subseteq J$ of products that they consider to purchase (see Shocker et al. Shocker et al. (1991) for a survey on the consideration-set modeling literature).

In each period the firm offers a subset $S$ of its products for sale, called the offer set. Given an offer set $S$, an arriving customer purchases a product $j$ (at the price $r_j$) in the set $S$ or decides not to purchase any (no-purchase). The no-purchase option is indexed by 0 and is always present for the customer.
A segment-\(l\) customer’s choice probabilities are not affected by the availability of products \(j \in J \setminus C_l\). A segment-\(l\) customer purchases \(j \in S\) if \(j \in S \cap C_l\) with probability \(P_j^l(S)\), \(S \subseteq J\). These functions are either given by an oracle or by a functional form such as in the MNL model where \(P_j^l(S) = v_j / (\sum_{k \in C_l \cap S} v_k)\) for a set of “weights” \(v_l \in \mathbb{R}^{\mid C_l\mid}\) that capture the attractiveness of the products for each segment \(l\).

Whenever we specify probabilities for a segment \(l\) for a given offer set \(S\), we just write it with respect to \(S_l = C_l \cap S\) (note that \(P_j^l(S) = P_j^l(S_l)\)). So when the firm offers set \(S\), it sells \(j \in S\) with probability \(P_j(S) = \sum_{l \in L} p_l P_j^l(S_l)\) and makes no sale with probability \(P_0(S) = 1 - \sum_{j \in S} P_j(S)\).

We define the vector \(P^l(S) = [P_1^l(S_l), \ldots, P_n^l(S_l)]\) (recall the no-purchase option is indexed by 0, so it is not included in this vector). We define the vector \(P(S) = [P_1(S), \ldots, P_n(S)]\). Notice that \(P(S) = \sum_{l \in L} p_l P^l(S)\). We define the vectors \(Q^l(S) = AP^l(S)\) and \(Q(S) = AP(S)\) to denote the expected resource consumption for an offer set \(S\) by segment \(l\). Likewise, the expected revenue function for segment \(l\) is \(R^l(S) = \sum_{j \in S_l} r_j P_j^l(S_l)\) and the expected revenue from a given arrival, \(R(S) = \sum_{j \in S} r_j P_j(S)\).

In our notation and demand model we broadly follow Bront et al. (2009) and Liu and van Ryzin (2008).

### 2.2 Dynamic Program

We describe the stochastic dynamic program to determine the optimal offer set at each point in time. While computationally intractable, it gives a conceptual reference point of the value we are trying to approximate with tractable methods.

Let \(V_t(c_t)\) denote the maximum expected revenue that can be earned over the remaining time horizon \([t, T]\), given remaining capacity \(c_t\) in period \(t\). Let \(J(c_t)\) denote the set of products that can be offered given remaining available capacity, i.e., \(J(c_t) := \{j \in J : A_j \leq c_t\}\). Then \(V_t(c_t)\) satisfies the Bellman equation

\[
V_t(c_t) = \max_{S \subseteq J(c_t)} \left\{ \sum_{j \in S} \lambda P_j(S) \left( r_j + V_{t+1}(c_t - A_j) \right) + \left( \lambda P_0(S) + 1 - \lambda \right) V_{t+1}(c_t) \right\}, \quad \forall t, \forall c_t, \quad (1)
\]

with the boundary condition \(V_T(c_T) = 0\) for all \(c_T\). Let \(V^{DP}(c_0)\) denote the optimal value of this dynamic program from 0 to \(T\), for the given initial capacity vector \(c_0\). Solving the dynamic program (1) is intractable because the state space explodes even for small problems. Therefore, we are forced to look at
approximations to the dynamic program (1).

2.3 Choice Deterministic Linear Program (CDLP)

The Choice-based Deterministic Linear Program (CDLP) approximation defined in Gallego et al. (2004) and Liu and van Ryzin (2008) has $2^n$ decision variables $w_S$ (recall that $n$ denotes the number of products). The decision variables can be interpreted as the amount of time set $S$ is offered:

$$
\max \sum_{S \subseteq J} \lambda R(S) w_S \\
\text{s.t.} \sum_{S \subseteq J} \lambda w_S Q(S) \leq c_0 \\
(CDLP) \quad \sum_{S \subseteq J} w_S = T \\
0 \leq w_S, \ \forall S \subseteq J.
$$

That is, we maximize the total expected revenue, subject to the constraint that the total expected capacity consumption on each resource $i$ must be less than or equal to the initially available capacity $c_{0i}$. The second constraint (2) says that we offer product sets over $T$ time units.

Liu and van Ryzin (2008) show that the optimal objective value of CDLP is an upper bound on $V^{DP}$. They also show that the problem can be solved efficiently by column generation for the MNL model with non-overlapping segment consideration sets. Bront et al. (2009) and Rusmevichientong et al. (2014) investigate this further and show that column generation is NP-hard if the consideration sets for the segments overlap for the MNL choice model with two segments.

2.4 Enhanced Segment-Based Deterministic Concave Program (SDCP+)

Talluri (2014) proposed an upper bound on CDLP called the Segment-based Deterministic Concave Program (SDCP). SDCP optimizes the offer set for each segment separately. SDCP and CDLP have the same objective values when the consideration sets for the different segments are disjoint. We do not elaborate on the SDCP formulation as it is not considered further in this paper, but it corresponds to the formulation (SDCP+) given below without the constraints (4).
In applications, the segments’ consideration sets can overlap in a variety of ways and, as the choice probabilities depend on the offer set, they do not have any structure that we can exploit. We call a set of constraints valid for a linear programming approximation of the dynamic program (1) if adding the constraints preserves the property that its optimal objective value still forms an upper bound on $V^{DP}$. Meissner et al. (2013) develop a set of valid inequalities for SDCP called product cuts—the constraints (4) below—that tighten the SDCP bound. We call the formulation SDCP+ in light of the additional constraints (4). Let $S_{lm}$ represent subsets of $C_l \cap C_m$, i.e., subsets in the intersection of the consideration sets of segments $l$ and $m$. SDCP+ is:

$$\max \sum_{l=1}^{L} \sum_{S_l \subseteq C_l} \lambda_l R^l(S_l)w^l_{S_l} \quad \text{s.t.} \quad (SDCP+)$$

$$\sum_{l=1}^{L} y_l \leq c_0$$

$$\sum_{S_i \subseteq C_i} \lambda_i Q^i(S_l)w^i_{S_i} \leq y_l, \quad \forall l \in L$$

$$\sum_{S_i \subseteq C_i} w^i_{S_i} = T, \quad \forall l \in L$$

$$\sum_{S_i \subseteq C_i, |S_i| \geq |S_{lm}|} w^i_{S_i} - \sum_{S_m \subseteq C_m, |S_m| \geq |S_{lm}|} w^m_{S_m} = 0, \quad \forall S_{lm} \subseteq C_l \cap C_m, \forall \{l, m\} \subset L : C_l \cap C_m \neq \emptyset \quad (4)$$

$$w^l_{S_l} \geq 0, \quad \forall S_l \subseteq C_l, \forall l \in L,$$

$$y_l \geq 0, \quad \forall l \in L.$$

The vector $y_l$ represents capacity allocation to segment $l$ subject to total available capacity $c_0$. We maximize total expected revenue from each segment $l$ subject to several constraints. The first represents that the capacity allocations are limited by the overall available network capacity. The second set enforces that each segment can only consume at most as many resources as have been allocated to it. The third ensures that we offer product sets (possibly the empty set) over the full time horizon. The intuition behind the product cuts (4) is the following: SDCP+ can be seen as a collection of segment-level implementations of CDLP linked via the constraint (3) and tightened via the product cuts (4). For any set $S_{lm} \subseteq C_l \cap C_m$, the length of time that set $S_{lm}$ is offered to segment $l$ (possibly alongside other products) must be equal to the length of time that it is being offered to segment $m$ (again possibly alongside other products). The numerical
experiments of Meissner et al. (2013) show that generating just a few of these constraints can be sufficient to obtain close approximations to the optimal CDLP objective function value.

3 Analysis of CDLP and SDCP+

We wish to understand when the optimal objective value of SDCP+ is the same as CDLP. To this end, we first develop a simple example to illustrate that there can be a strict gap between the optimal objective values of CDLP and SDCP+, even if all product cuts are satisfied. The underlying reason for that gap is that there is no solution to CDLP that can be projected onto the segment consideration sets so as to coincide with the segment-level optimal solution. When are these two formulations equivalent then? We explore this issue in the remainder of this section following the example.

Example 1 Our example has five products and three segments and their respective consideration sets are shown in Figure 1. For this example we show that there is a gap between the optimal objective values of SDCP+ and CDLP even if we generate all constraints of type (4). Assume $T = 1$, capacity $c = 1$, revenue $r_j = 1$ for all products $j \in J = \{1, 2, 3, 4, 5\}$, and purchase probabilities defined as follows: $P_A^j(\{1, 2\}) := 0.5$ for $j = 1, 2$, $P_B^j(\{2, 5\}) := 0.5$ for $j = 2, 5$, $P_B^j(\{1, 2, 3\}) := 1/3$ for $j = 1, 2, 3$, $P_C^4(\{4\}) := 0.5$ for $j = 3, 5$, and 0 for all other sets. We show that there is a feasible solution to SDCP+ for this example with objective value 1. A feasible solution to SDCP+ is given by $y_l = 1/3$ for all segments $l \in \{A, B, C\}$, and $w_{A}^l(\{1, 2\}) = w_{B}^l(\{2\}) = w_{B}^l(\{1, 2, 3\}) = 0.5$, $w_A^l(\{2, 5\}) = 0.5$, and $w_{C}^l = 0$ otherwise for all $l \in \{A, B, C\}$, $S_l \subseteq C_l$. This solution is feasible to SDCP+ since $\lambda_l \sum_{S_i \subseteq C_l} Q^l(S_i)w_{S_i}^l = 1/3 = y_l$ for all segments $l$, and the product cuts (4) for all pairs of segments $\{l, m\}$ and sets $S_{lm} \subseteq C_l \cap C_m$, $S_{lm} \neq \emptyset$ are satisfied as reported in Table 1. They also hold for $S_{lm} = \emptyset$ since the solution satisfies $\sum_{S_i \subseteq C_l} w_{S_i}^l = 1$ for all segments $l$. The other constraints are likewise satisfied as can easily be checked. The objective value is 1 since $\lambda_l \sum_{S_i \subseteq C_l} R^l(S_i)w_{S_i}^l = 1/3$ for each
Next, we show that there is no corresponding solution to CDLP with the same objective value. Note that CDLP has $2^5 = 32$ variables corresponding to subsets $S \subset J = \{1, 2, 3, 4, 5\}$. Under the single-leg example described above, we can enumerate all 32 subsets $S$ and calculate the corresponding objective coefficient $\lambda R(S)$. We find that $\lambda R(S) \leq 2/3$ for all $S \subset J$, with equality reached for the sets $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{2, 3, 5\}$, $\{1, 2, 3, 4\}$ and $\{1, 2, 3, 5\}$. It follows that there can be no feasible solution to CDLP that has objective value greater than $2/3$ since the objective is a convex combination of these coefficients (note that $T = 1$).

Moving on from the example, we seek to obtain a structural result on when CDLP and SDCP+ are equivalent. Since the overlap of the consideration sets plays a critical role in Example 1, let us represent the overlap structure in a graph. Specifically, we define a bipartite intersection graph as follows: There are two types of nodes, one type called segment node, the other is called intersection node. Each node of the former type corresponds to a segment, each of the latter represents a set of the form $C_k \cap C_l$ for some segment pair $(k, l)$. If there are two pairs of segments $(m, n)$ and $(k, l)$ with $C_m \cap C_n = C_k \cap C_l$, then $S = C_m \cap C_n$ is represented by a single intersection node. Edges from segment node $k$ connect to all the sets of the form $C_k \cap C_l \neq \emptyset$ for any $l \neq k$. In graph theory, a connected graph without cycles is called a tree, and a disjoint union of trees is called a forest (Bondy and Murty, 1976).

The intersection graph of the example of Figure 1 has a cycle, as can be seen from Figure 2. This turns
out to be the critical feature: If the segment consideration sets do not have a cycle and are arranged say in the form of a tree (or, in general, a forest), then the product cuts are sufficient to ensure equivalence between CDLP and SDCP+, as stated in Proposition 1 below. Before establishing this result, we provide some intuition for it: The intersection tree tells us which segments are directly or indirectly connected to each other, in the sense that a solution $w^l$ for some segment $l$ is dependent on the solution $w^k$ for any segment node $k$ that is reachable from segment node $l$. Example 1 illustrates this point: It is not possible to arrange the segment-level solutions in a way such that they are consistent. By consistent, we mean that there is a feasible solution to CDLP that can be projected onto the segments to obtain the SDCP+ solution. For instance, if we would arrange the SDCP+ solution so that the sets $S_1^A = \{1, 2\}$, $S_1^B = \{1, 2, 3\}$, and $S_1^C = \{3, 5\}$ are offered in parallel (recall that they are all offered for the same duration), then we cannot find a single set $S \subset J$ whose projection onto the segment consideration sets results in $S_1^A$, $S_1^B$ and $S_1^C$, respectively. To see this, note that $S_1^A$ requires us not to offer product 5, whereas $S_1^C$ does require us to offer product 5. In fact, there is no possibility to arrange the segment-level solutions so that all sets that would be offered in parallel are consistent, and this shows that the segment-level solutions of Example 1 do not have a corresponding CDLP solution.

The product cuts ensure that any offer set in the intersection of any two segments’ consideration sets is being offered to both segments for the same time (possibly alongside offering other products). This is the case in Example 1: All sets are offered for the same duration. Suppose we have a tree-structured intersection graph with a segment node $l$ connected to a segment node $k$ via a single intersection node. For a given solution $w^l$, the product cuts allow us to arrange the segment-level solutions $w^k$ in a way such that they represent the projections of a feasible global solution onto the respective segments. We can repeat this argument to construct a global CDLP solution by moving along the tree. However, if there is a cycle, then this pairwise approach to construct a global solution does not work any more because there is no guarantee that the last segment-level solution ($w^C$ in Example 1) in the sequence along the cycle is consistent with the first segment-level solution ($w^A$ in Example 1).

**Proposition 1.** If the intersection graph is a forest, then CDLP and SDCP+ have the same optimal objective value (henceforth referred to by the shorthand notation CDLP = SDCP+).

**Proof**
Any solution to CDLP is a solution to SDCP+ as shown in Meissner et al. (2013), hence CDLP ≤ SDCP+. It remains to show CDLP ≥ SDCP+. We use a merging procedure in this proof. For clarity, we first explain it on a simple example with two segments, A and B, displayed in Figure 3 along with their corresponding consideration sets. Figure 3 shows a feasible solution to SDCP+. We wish to construct a feasible solution to CDLP from this solution with the same objective value. For simplicity assume that the only sets with positive weights that contain U are S2 in the consideration set of A and S3 in the consideration set of B. Note that the product cuts imply the restriction wA(S2) = wB(S3) since S2 and S3 are the only offer sets with positive weight in the segment-level solution for segments A and B respectively that contain the set U in the intersection of the consideration sets. Moreover, wA(S1) + wA(S2) = T and wB(S3) + wB(S4) = T. This implies that wA(S1) = wB(S4). So we construct weights for the CDLP formulation as wCDLP(S1∪S4) = wA(S1) = wB(S4) and wCDLP(S2∪S3) = wA(S2) = wA(S1). This CDLP solution satisfies wCDLP(S1∪S4) + wCDLP(S2∪S3) = T, as well as the capacity constraints, and has the same objective value as the SDCP+ solution.

In the proof of Proposition 1 (see appendix), we essentially repeat this argument for the more complicated case with L segments and arbitrary consideration sets using an induction argument (made possible by the tree structure).

In addition to the tree-structured intersection graphs, we identify another structure that guarantees equivalence of CDLP and SDCP+. We show that nested consideration sets also guarantee CDLP = SDCP+ even though such consideration sets do not have the tree structure.

**Proposition 2.** For a nested consideration set structure C1 ⊆ C2 ⊆ ... ⊆ CL, CDLP = SDCP+.

**Proof**
See appendix.
Let us consider why we do not necessarily need a tree structure in the intersection graph: the main induction step in the proof of Proposition 1 is a merging procedure between a leaf node and the rest of the intersection graph. What we require is that at every step we should be able to find a leaf node. Once we identify the leaf node (segment), and remove it, the intersection graph of the remaining segments can be quite different from the original. Indeed, that is the reason for the tractability of the nested consideration structure in Proposition 2 even though the original intersection graph is not a tree. Hence, we can write a more general version of Proposition 1 as follows:

**Proposition 3.** CDLP is equivalent to SDCP+ when the intersection graph has a sequence of segment-nodes, such that the first node in the sequence is a leaf node, and after removal of each leaf node and a re-drawing of the intersection graph with the remaining segments the next segment-node in the sequence is also a leaf node in the new intersection graph.

## 4 Applications

In this section we present some applications from the literature where the consideration set structures that we have described appear naturally in the modeling. These applications are: RM of advance tickets and ticket options for sport events (§4.1), RM for primary care clinics (§4.2), dynamic pricing of home delivery time slots (§4.3), low cost airlines (§4.4), and retail (§4.5).

### 4.1 Revenue management of advanced tickets and options for sports tickets

Sport event ticket options have become so popular that there is a software company called TTR that specializes in selling Internet platforms to teams and events that wish to offer options. Balseiro et al. (2011) consider a scenario where advance tickets for the tournament are sold before it starts—hence the identities of the two teams playing in a tournament final are unknown at this time. However, fans of a specific team are only interested in attending the event if their team makes it to the final. To address this uncertainty, the authors propose team-specific call options under which a customer can pay a small non-refundable amount in advance for the right to attend the event if and only if the specified team makes it to the final and if he pays an additional amount once the finalists are known. Such options allow event organizers in principle
to oversell capacity many times because only fans with advance tickets or with options for the two finalist teams will be able to attend the event. Fans are segmented by the teams they support; assuming there are $L$ teams in the tournament in total, we therefore have $L$ customer segments. A customer from segment $l$ has the choice between buying an advance ticket $A$ that would give access to the final regardless of who will be playing, and an option $O^l$ for team $l$. In other words, a customer in segment $l$ has a consideration set $\{A, O^l\}$. Thus we obtain a star consideration set structure (Figure 4).

Balseiro et al. (2011) use the CDLP to solve the problem, and show for their specific model that an equivalent, more compact formulation exists that actually is special case of SDCP+. The main result of our paper provides a more general explanation for the equivalence of CDLP and SDCP+, namely that the consideration set structure is a tree.

4.2 Revenue management for primary care clinics

Gupta and Wang (2008) present an application of revenue management under patient choice of primary care providers and appointment time-slots. Specifically, the problem is to manage physicians’ consultation time slot availabilities over a finite booking horizon so as to maximize revenues. Each physician has a panel of patients for whom he is the designated primary-care provider; these patient groups correspond to customer segments with preferences for particular physicians. Patients of any segment can choose between all available combinations of all appointment time-slots and all physicians, hence the consideration sets are all identical. Same-day patients form another segment and are assumed to be willing to accept any available slot with any physician on the workday. In this application, the product is a combination of physician-time combination. All patient segments consider all products, and therefore the intersection tree has a star structure as in the previous example. Gupta and Wang (2008) proposed various heuristics to tackle the problem; our main
result tells us that we can use the tractable SDCP+ formulation in lieu of CDLP as an alternative solution approach.

4.3 Dynamic pricing of home delivery time-slots

Another application related to appointment scheduling is the work of Asdemir et al. (2009) who look at the question of how to dynamically price delivery time-slots for attended home delivery over a finite booking horizon using dynamic programming. Different prices can be quoted for different delivery time-slots at any given point in time. There are different customer segments in a given area based on a choice model that reflects their preferences for specific time-slots as well as price sensitivity. Their model considers all geographic areas as independent of each other. All segments in a given area are assumed to consider all available delivery time-slots. If the time slot prices have to be the same across all segments within a given geographical area (which would be reasonable so as to avoid customer dissatisfaction due to perceived unfairness of this group-based discrimination), we again have a star-shaped tree as the intersection graph. Similar to the previous application, the geographic area and time-slot combination is the product.

As solution method, Asdemir et al. (2009) propose a dynamic program that has an exponentially growing state space in the number of delivery time-slots. Alternatively, one could again use SDCP+ (equivalent to CDLP owing to the tree structure) to obtain an approximate policy that can be calculated even for large problem instances where dynamic programming becomes intractable.

4.4 Nested consideration set structure (low-cost airline model)

Consider a fully nested consideration set structure where the $L$ consideration sets are nested as $C_1 \subseteq C_2 \subseteq \ldots \subseteq C_L$. This models buy-up/buy-down amongst unrestricted products where complete dilution is possible. This type of structure is encountered on single-leg flights or in a retail context where customers remove products from consideration based on certain cutoff values for the products’ attributes or qualities, and these can be ranked linearly. The latter example was proposed by Feldman and Topaloglu (2015) to motivate work on assortment optimization under MNL with nested consideration sets. Proposition 2 shows that this is tractable, following the structure defined in Proposition 3.
4.5 Retail

When defining segments and their consideration sets, there is often a certain degree of subjectivity and modeling flexibility. For instance, consider the segmentation study based on retail scanner data by Kamakura and Russell (1989). Initially, they discover nine customer segments with intersection graph as illustrated on the left of Figure 5. There are four loyal segments for the brands A, B, C and P, respectively, and five switching segments. They found that segment 5 and segment 3 could be merged without major impact on the preference structure. Furthermore, one could refine the sets of the switching segments by including only the products with purchase probabilities greater than 10% (the threshold mentioned in Kamakura and Russell (1989)). The resulting simplified consideration sets of the modified segments 1’, 2’, 3’, 4’ were \{A\}, \{A,B\}, \{A,B,C\} and \{B,C,P\}, respectively. The corresponding intersection graph is depicted in the middle of Figure 5 along with the simple tree structure on the right that results from removing leaf nodes as in Proposition 3. This serves as an example of how a modeler could sensibly change the segmentation so as to obtain a tractable structure; effectively it comes down to balancing the loss in modeling accuracy with
tractability of the subsequent optimization.

5 Conclusions

Discrete-choice models are widely used to model consumer purchase behavior in assortment optimization and revenue management. The firm has to make a decision on what assortment to offer at each point in time without the ability to identify the customer’s segment. In many applications, each customer segment is associated with a consideration set that represents the set of products that customers in this segment consider for purchase. The formulation $CDLP$ has been proposed to determine these offer sets but its size grows exponentially in the number of products and it is computationally intractable for even modest-sized applications when segment consideration sets overlap. The formulation $SDCP+$ runs much faster than $CDLP$ and often obtains the same optimal objective function value. In this paper we show that $CDLP$ and $SDCP+$ are equivalent if the intersection graph of the segment consideration sets is a tree or if the consideration sets are nested. We give a number of examples from the literature that naturally exhibit these structures.

References


Appendix

Proof of Proposition 1

Proof

Any solution to $CDLP$ is a solution to $SDCP+$ as shown in Meissner et al. (2013), hence $CDLP \leq SDCP+$. It remains to show $CDLP \geq SDCP+$. 

Consider the case of a single segment $L = 1$, and a given feasible solution $(w_t^L, y_t)_S$ to $SDCP+$, then $w'^{CDLP}_S := w_t^L$ for all $S \subseteq J$ is a feasible solution to $CDLP$ with the same objective value.

Next, we consider $L > 1$. Without loss of generality, the discussion will focus on an intersection graph that is a finite tree rather than a forest since the same arguments can be made for each tree that makes up the forest. Assuming that it is a tree, there must be at least two leaves, i.e. nodes with degree 1. By definition, intersection nodes have at least degree 2, so there exists a segment node that is a leaf. Without loss of generality, let this node correspond to the consideration set of segment $L$, and let $SDCP+$ represent the problem $SDCP+$ with the segment $L$ removed. Consider a feasible solution $(w, y)$ to $SDCP+$, where $(w, y)$ is shorthand notation for $w_t^l$ for $S_t \subseteq C_l$, for all segments $l \in \mathcal{L} := \{1, 2, \ldots, L\}$, and $y_{il}$ for all resources $i$ and $l \in \mathcal{L}$.

This solution induces a feasible solution $(\bar{w}, \bar{y})$ to $SDCP+$ by defining $\bar{w}_t^l := w_t^l$ for all $S_t \subseteq C_l$, for all $l \in \mathcal{L} := \mathcal{L} \setminus \{L\}$, and $\bar{y}_t := y_t$ for all $l \in \bar{L}$. The solution $(\bar{w}, \bar{y})$ produces an objective value equal to that of $SDCP+$ less $\sum_{S_t \subseteq C_l} \lambda_l R^l(S_l)w_{SL}^L$. By the induction assumption, there exists a feasible solution $\bar{w}_S^{CDLP}$ for all $S \subseteq \mathcal{J} := \cup_{l=1}^{L-1} C_l$ to $CDLP$ with the same objective value, and $\bar{w}_S^{CDLP}$ induces $(\bar{w}, \bar{y})$ meaning that $\bar{w}_t^l = \sum_{S_l \subseteq J_l \subseteq C_l} \lambda_t Q^l(S_l)\bar{w}_S^{CDLP}$ for all $l \in \mathcal{L}$, $S_l \subseteq C_l$ for $l \in \mathcal{L}$, and $\bar{y}_t = \sum_{S_t \subseteq C_l} \lambda_t Q^l(S_t)\bar{w}_t^l$ for all $l \in \bar{L}$.

Now we construct a feasible solution $w_S^{CDLP}$ for all $S \subset J$ to $CDLP$ that induces $(w, y)$ for $SDCP+$ with the same objective value. Since $L$ is a leaf of the intersection tree, all interactions with other segments...
are via a set $S^{\text{int}}$ that is associated with the intersection node to which $L$ is connected. Let us denote all segments that are connected to this intersection node by $L^{\text{int}}$.

Consider a set $U \subseteq S^{\text{int}}$ that is \textit{maximal} for segment $L$ with respect to $S^{\text{int}}$, that is there is no set $S_L \subseteq C_L$ such that $U \subset S_L \cap S^{\text{int}}$ and positive support $w^{L}_S > 0$. Note that for a feasible solution to $\text{SDCP}^+$, the product cuts ensure that if a set is maximal for $L$ with respect to $S^{\text{int}}$, it is maximal for all segments $l \in L^{\text{int}}$ with respect to $S^{\text{int}}$. Moreover, from the definition of maximal

$$\sum_{S_L \subseteq C_L | S_L \cap S^{\text{int}} \supseteq U} w^L_{S_L} = \sum_{S_L \subseteq C_L | S_L \cap S^{\text{int}} = U} w^L_{S_L}, \forall l \in L^{\text{int}}.$$  

We select an arbitrary \textit{maximal} set $U \subseteq S^{\text{int}}$ and segment $l \in L^{\text{int}}$. The following argument shows that the total weight $\tau(U)$ that we offer sets that intersect with $S^{\text{int}}$ exactly in $U$ is the same in solutions $w^L$ and $\bar{w}^{\text{CDLP}}$:

$$\tau(U) = \sum_{S_L \subseteq C_L | S_L \cap S^{\text{int}} = U} w^L_{S_L}$$  

$$= \sum_{S_L \subseteq C_L | S_L \cap S^{\text{int}} = U} w^L_{S_L}$$  

$$= \sum_{S_L \subseteq C_L | S_L \cap S^{\text{int}} = U} \bar{w}^{\text{CDLP}}_{S_L}$$  

The first equality holds by definition, the second due to maximality and the product cuts being satisfied by the solution $w$ to $\text{SDCP}^+$, the third since $w^L_{S_L} = \bar{w}^L_{S_L}$, the fourth because $w^{\text{CDLP}}$ induces $\bar{w}$, and the final one as a result of a reformulation.

As a consequence, we can merge the solution $w^L_{S_L}$ with $\bar{w}^{\text{CDLP}}$ over total weight $\tau(U)$ to obtain $w^{\text{CDLP}}$ for all sets that intersect with $S^{\text{int}}$ only in the fixed set $U$. We illustrate the process in Figure 6 by drawing two parallel bars of equal length representing the weight $\tau(U)$, each bar with intervals corresponding to the support of the solutions $w^L$ and $\bar{w}^{\text{CDLP}}$ (the order of the sets does not matter). Merging the sets as depicted ensures that the constructed solution $w^{\text{CDLP}}$ induces $w^L$ as well as $w^l$ for $l \in \hat{L}$ (the latter due to the induction assumption on $\bar{w}^{\text{CDLP}}$).
Now remove all the solution components \( w_{S_1}^L \) and \( w_{S_1}^L \) with positive support for all \( l \in \mathcal{L}^{\text{int}} \) with \( S_i \cap \mathcal{S}^{\text{int}} = U \) and \( S_L \cap \mathcal{S}^{\text{int}} = U \). After the removal, the product cut equations for the remaining solution remain valid because of the equalities (5–6). We repeat this merging process by taking a maximal \( U \subseteq S^{\text{int}} \) at each stage till we conclude with \( U = \emptyset \). At every stage, as \( U \) is a maximal set, all the sets that contained \( U \), namely sets of the form \( U \subseteq S_i \cap \mathcal{S}^{\text{int}}, l \in \mathcal{L}^{\text{int}} \) were maximal sets in previous stages and therefore accounted for by equalities (5–6) for the set \( S_i \); now combining it with the product cuts for the set \( U \), we again obtain equalities (5–6).

The solution \( w^{\text{CDLP}} \) that emerges from this process is feasible to \( \text{CDLP} \): it holds that \( \sum_{U \subseteq S^{\text{int}}} \tau(U) = T \) (note that \( U = \emptyset \subseteq S^{\text{int}} \)), and therefore, by construction, \( \sum_{S} w_S^{\text{CDLP}} = T \). That the capacity constraint of \( \text{CDLP} \) is satisfied follows from the induction assumption that \( w^{\text{CDLP}} \) induces \((\bar{w}, \bar{y})\), with \( \bar{w} := w \) and \( \bar{y} := y \), combined with the fact that \((w, y)\) is feasible to \( \text{SDCP}^+ \) and that we constructed \( w^{\text{CDLP}} \) in a way such that we only added capacity consumption equal to that of segment \( L \) under solution \( w^L \). So the combined solution also satisfies the induction for \( L \) segments.

The objective value of \( \text{CDLP} \) equals that of \( \text{SDCP}^+ \) because in the merging process we only add products of \( C_L \setminus \mathcal{S}^{\text{int}} \) to the solution \( w^{\text{CDLP}} \), and since these products do not influence other segments as they are only in the consideration set of segment \( L \), we only add the contribution of segment \( L \) to the objective without a change of the contribution of other segments.

\[ \square \]

**Proof of Proposition 2**

**Proof**

Any solution to \( \text{CDLP} \) induces a feasible solution in \( \text{SDCP}^+ \) with the same value. To show equivalence,
we have only to construct a feasible solution to CDLP from a feasible solution to SDCP+ with the same objective value. We refer to a subset of $S \subseteq C_l$ with a positive weight $w^l_S$ in SDCP+ as positive in $l$.

Assume without loss of generality that a solution to SDCP+ has at least one set $S_1 \subseteq C_1$ with $w^1_{S_1} > 0$. Note that segment node 1 is a leaf in the intersection graph due to the fully nested structure. Then there are sets $S_l \subseteq C_l$ such that $w^1_{S_1} \geq w^l_{S_l}$ for all $l > 1$. Moreover, we can describe a sequence of maximal nested sets $S_l$ for $l = 2, \ldots, L$ containing $S_1$ that, due to the product cuts (4), have positive $w^1_{S_1}$ and these variables are non-increasing: namely, $S_1 \subseteq S_2 \subseteq \ldots \subseteq S_L$ and $w^1_{S_1} \geq \ldots \geq w^L_{S_L} > 0$. Moreover, as the consideration sets are nested, the product cuts (4) imply that $S_{l-1} = S_l \cap C_{l-1}$, as no positive maximal set in $l-1$ contains $S_{l-1}$.

As each is maximal within its consideration set, these sets have the property that $S_l$ is not contained in any set $S \subseteq C_l$ of segment $l$ with positive weight $w^l_S > 0$. Therefore, we can create a solution to CDLP by giving the maximal set $S_L$ a weight $w^L_{S_L}$ and subtracting $w^L_{S_L}$ from $w^1_{S_1}, \ldots, w^L_{S_L}$. Now repeat this operation, at each step peeling off the maximal sequence of weights, to obtain a solution to CDLP. Note that this procedure terminates in a finite number of steps.

The two solutions have the same objective value: $\sum_S \lambda R(S)w_S = \sum_l \lambda_l \sum_{S_l} R^l(S_l)[\sum_{S : S \cap C_l = S_l} w_S] = \sum_l \lambda_l \sum_{S_l} R^l(S_l)w^l_{S_l}$ where $\sum_{S : S \cap C_l = S_l} w_S = w^l_{S_l}$ holds by construction for all $S_l \subseteq C_l$ for all $l$. \hfill $\Box$

21