Models and heuristics for product line selection

Name: Lu Zhang, Shuxuan Zeng

- **Contributions**

This paper opened a new start on extending research on single product design optimization toward product line decisions and offered several heuristics approach to solve the problem.

Product line decisions mainly involve determinations of:

- Size and composition of product line
- Composition of alternative collection of products
- Alternative sets of promotional benefits

The paper mainly talks about the methodology of how to optimize the product line. When the product line size K (number of items in the array) is fixed, how to obtain the composition which could maximize the objective function of buyer/seller's welfare. When changing the value, we are able to determine product lines with various sizes.

The research paper introduced heuristics for subset selection (subset size K) from a relative large set of candidate product (set size N). The problem could be constructed based on perspectives from both buyer's and seller's side.

- **The buyer's welfare problem**

Suppose now we are considering a case, in which an employer would like to provide benefit packages for his/her employees. How to select a subset of size K out from N candidate packages so that the employees' benefits could be maximized, given that each package cost the company about the same.

The individual set was constructed as $I = \{1,2,\ldots M\}$, where each element represents an employee. The set of products was interpreted as $J=\{1,2,\ldots N\}$, $j \in J$ is candidate package. u_{ij} denotes utility of individual i choosing benefit package j, for $i \in I$ and $j \in J$.

The optimization question could be interpreted in the following way:

We are aiming for selecing a subset $S \subset J$, $|S|=K$, for which

$$\max Z=\sum_{i \in I} \sum_{j \in J} x_{ij} u_{ij} \qquad \text{for } i \in I \text{ and } j \in J.$$

$$y_j = \begin{cases} 0, & for\ package\ j\ not\ in\ S \\ 1, & for\ package\ j\ in\ S \end{cases}$$

$$x_{ij} = \begin{cases} 0, & employee\ i\ does\ not\ choose\ j \\ 1, & if\ employee\ i\ chooses\ j \end{cases}$$

$$\sum_{j \in J} x_{ij} = 1, \text{ for } i \in I$$

$$\sum_{j \in J} y_j = K, \text{ for } j \in J$$

$$0 \le x_{ij} \le y_j \le 1, \text{ for } i \in I \text{ and } j \in J.$$

- **The seller's welfare problem**
  Sometimes the two parties, buyer and seller, are likely to be in conflict for subset chosen. In this sense, it is essential to construct the problem from seller's perspective. Say v_{ij} denotes the value to seller when customer i chooses product j. To consider the problem more thoroughly, we include utility u_{i0} and v_{i0} which represent the utilities (to buyer and seller) of customer choosing none product.

When taking the importance of customer into consideration, we assign weight w_i to different individuals. In this regard, the objective function is updated as:

$$\max Z=\sum_{i \in I} \sum_{j \in J} w_i x_{ij} u_{ij} \qquad \text{for } i \in I \text{ and } j \in J.$$

what's worth noticing is that the units of u_{ij} and v_{ij} should be normalized among individuals. To simplify the problem, u_{ij} here is 0-1 variable.

To solve buyer's problem, heuristics methods could often help to achieve close to optimum solution and dramatically reduce computation time.

- The greedy and interchange heuristics:
  We start with empty subset S, and add element to S per time (S is built up sequentially)
  Add j to S for max $(Z + \sum_{i \in I} x_{ij} u_{ij})$, until S can no more be increased.
  At this point, interchange each element in S to element not in S (if S could be further augmented).
  Z_G is the maximum value Z achieved by greedy and interchange heuristics.
- Lagrangian relaxation
  When lessening up the side constrains, the integer programming problem is much easier to solve. The value Z_0 achieved by lagragian relaxation could be taken as an upper bound to the optimum problem.
- Worst case scenario $Z_G / Z_0 = 1 - |\frac{K-1}{K}|^K$, as $K \to \infty$, the ratio goes to $(1 - e^{-1})$, which is about 63% of optimal.

- **Solving the seller's problem**

  Assumptions:
  1. Each consumer *i* chooses at most 1 product *j* such that product *j* products the highest utility to given that it exceeds the status quo utility u_{i0}.
  2. The utility to the seller is v_{ij} given consumer *i* chooses product *j*
  3. Total return Z = sum of returns across buyers and sellers

     Hence, the formulation of the problem is:

$$y_j = \begin{cases} 1, \ for\ product\ j\ in\ S \\ 0, for\ package\ j\ not\ in\ S \end{cases}$$

$$y_0 = 1$$

$$x_{ij} = \begin{cases} 1, if\ u_{ij}y_j \geq u_{il}y_l, l = 1, \ldots, N \\ 0,\ otherwise \end{cases}$$

If $x_{ij} = 1$ for only the smallest j for which $x_{ij} = 1$ above. Then the problem becomes:

max $Z=\sum_{i \in I} \sum_{j \in J} x_{ij} v_{ij}$ for i = 1,…,M and j = 1,…,N. subject to

1. $x_{ij} \leq y_j$
2. $\sum_{j=0}^{N} y_j \leq K + 1$
3. x_{ij}, y_j are 0-1 variables

From the formulation, only greedy heuristics is applicable among methods described in the buyer's problem. But, it could lead to arbitrarily bad results. $Z_G / Z_0 = K^{-1}$. As $K \to \infty$, the ratio goes to 0.

However, we can consider the cases when some constraints apply on u_{ij} and/or v_{ij},

1. Heuristic H1: Almost equal v_{ij}:
   i. Modify buyer utility as:

$$u_{ij} = \begin{cases} v_{ij},\ if\ \ u_{ij} \geq u_{i0} \\ 0, ow \end{cases}$$

   ii. Apply buyer's greedy to u_{ij}: then the $Z_{H1} / Z_0 \geq [v_{min}/v_{max}][1 - |\frac{K-1}{K}|^K]$

- **Empirical applications**

An insurance company wants to consider if it is a good idea to introduce home owner's insurance. An hybrid conjoint analysis is applied. The following data is collected:

1. Self-explicated desirabilities for levels of each attribute
2. Attribute importance data
3. Full-profile evaluations

Each profile differs from the base profile in the sense that one or more degraded features, but with a reduction in annual premium. Each respondent receives 2 balanced sets of four profiles each. Each set of 4 is ranked in terms of preference and then each item is rated in terms of likelihood of consideration at policy review time. A utility function is computed for each respondent. Following this step a computer choice simulator is used to examine the impact on probability of considering each of 32 different policies constructed by the marketing group.

Comparative analysis: each heuristics described earlier is applied to empirical data. Maximum value of K is set at 5.

1. The seller greedy heuristic: G
2. Modified buyer heuristic H0:

$$u_{ij} = \begin{cases} 1, & if \quad u_{ij} > u_{i0} \\ & 0, ow \end{cases}$$

3. Modified buyer heuristic H1:

$$u_{ij} = \begin{cases} v_{ij}, & if \quad u_{ij} \geq u_{i0} \\ & 0, ow \end{cases}$$

Empirical results are as follows:

TABLE 2

A Comparison of Three Heuristics Used to Find Solutions to the Seller's Welfare Problem (as a Function of the Probability Cutoff Value u_{i0})

Heuristic (M = 117; N = 32; K = 5)	Cutoff Value					
	0.25		0.5		0.75	
	Products	Z Value	Products	Z Value	Products	Z Value
G: The seller's greedy	1	10,900	1; 2	7,451	1; 15; 7; 18	5,795
H_0: A modified buyer's greedy (1, 0 u_{ij}'s)	1; 2; 3; 4; 5	10,766	1; 2; 3; 4; 5	7,451	1; 15; 7; 2; 3	5,778
H_1: A modified buyer's greedy (v_{ij}'s)	1; 2; 3; 4; 5	10,766	1; 15; 2; 3; 4	7,403	1; 15; 7; 2; 3	5,778
Complete enumeration over all 32 products (optimal)	1	10,900	1; 2	7,451	1; 15; 7; 18	5,795
Upper Bounds (B_1, B_2, and B_3)		10,900		7,476		5,827

Notably, product 1 is highly popular among buyers and also displays highest seller utlity of the set. In this case, u_{ij} tends to be positively correlated with v_{ij}.

G does very well and the other 2 heuristics also do reasonably well. And when the cutoff values = 0.5, 0.75, the optimal solution is not unique.

3 different bounds are obtained from the Lagrangian relaxation. This is the bound on the number of buyers who can be satisfied. All 3 procedures produce the same bound. More likely to be a coincidence.

These are:

1. Method B_1: find the maximum v_{ij} (over all $N = 32$ products) for each buyer, v_i^*, (assuming $u_{ij} > u_{i0}$) and sum these v_i^*'s across all M buyers.

2. Method B_2: replace all u_{ij}'s $> u_{i0}$ with 1's and replace all u_{ij}'s $\leq u_{i0}$ with 0's. Treat this problem as a buyer's problem and solve by Lagrangian relaxation; record the bound obtained from Lagrangian relaxation (restricted to five products in this example). This is a bound on the number of buyers who can be satisfied. (We then add up the b largest v_i^*'s in B_1, rather than all M terms.)

3. Method B_3: replace all u_{ij}'s $> u_{10}$ with their counterpart v_{ij}'s and replace all u_{ij}'s $\leq u_{i0}$ with 0's. Also treat this problem as a buyer's problem and solve it by Lagrangian relaxation; record the bound obtained from Lagrangian relaxation (restricted to five products in this example).

Since the empirical results only illustrates the case when u_{ij} tends to be positively correlated with v_{ij}. A simulation is conducted where u_{ij} tends to be negatively correlated with v_{ij}.

Results are the following:

Z_H/Z_O	The Seller's Greedy G	Other Heuristics H_0	H_1
0.80–0.849	0	5	27
0.85–0.899	0	363	528
0.90–0.949	0	615	434
0.95–0.999	225	17	11
1.0	775	0	0
	1,000	1,000	1,000

Again, heuristic G demonstrates superiority in recovering the optimal solution.

One major aspect needs to be considered.

A major aspect (that is difficult to express precisely) entails a special kind of dependence among the u_{ij}—the presence of a subset of products with high utilities for all individuals and other products which are moderately acceptable to *different segments* of the individuals. Thus, the behavior of the seller's greedy appears to be closely related to the market's potential for segmentation. (However, if the u_{ij}'s are positively correlated, across buyers, as was the case in the insurance example, the seller's greedy tends to perform well.)

As would be surmised, the preceding conditions are data dependent and not well captured by Monte Carlo studies. Hence, in future research we plan to study the impact of negative correlations between the u_{ij}'s and the v_{ij}'s based on actual data sets, in which the dependence structure among the u_{ij}'s is known at the outset; we could then vary the v_{ij}'s, subject to the structure.

In conclusion, this paper opens up many future research opportunities as outlined at the end of the paper.