The Mixed Integer Linear Product Line Design Problem

(Working Paper)

Ralf Krohn

September 21, 2017

1 Customer Choice Behavior

Discrete choice models are a tool to both analyze and predict individual choice behavior. An individual chooses exactly one alternative from a finite set of available, mutually exclusive, and collectively exhaustive alternatives. In our case the choice set includes all possible own new products, own already existing products as well as competitors' products and the alternative of not choosing any product, i. e., the "no–choice" alternative.
An individual's utility for an alternative is a result of the alternative's attributes as well as the individual's characteristics. Following the utility maximization choice rule an individual chooses the available alternative that dominates all other alternatives by having the highest utility (Koppelman & Bhat 2006). That is, individual $n \in \mathcal{N}$ chooses alternative $j \in \mathcal{J}$, iff

$$U_{n,j} > U_{n,h} \qquad \forall\, h \in \mathcal{J} \wedge h \neq j, \tag{1}$$

where $U_{n,j}$ is individual n's perceived utility value for alternative j. Since the decision making process is mostly not completely understood by the analyst, probabilistic choice models are used to account for unobserved characteristics and incomplete information. For that reason the utility is decomposed into two components:

$$U_{n,j} = v_{n,j} + \epsilon_{n,j} \qquad \forall\, n \in \mathcal{N}; j \in \mathcal{J}, \tag{2}$$

where $v_{n,j}$ denotes the deterministic utility part for an individual n and alternative j, which is observed by the analyst. $\epsilon_{n,j}$ is a stochastic error term that equals the difference between the known deterministic utility and the utility $U_{n,j}$ used by the individual, which is generally unknown.
Since $U_{n,j}$ is a random quantity, the analyst can only make probability statements about (1). The probability $p_{n,j}$ that individual n chooses alternative j is

$$p_{n,j} = P\left(U_{n,j} > U_{n,h} \quad \forall\, h \in \mathcal{J} \wedge h \neq j\right). \tag{3}$$

The multinomial logit model (MNL) (McFadden 1974) is based on the assumption that the error components $\epsilon_{n,j}$ are independent and identically type I extreme–value (also: Gumbel) distributed (iid EV) across alternatives as well as across individuals. Therefore, the logit choice probability

$$p_{n,j} = \frac{\mathrm{e}^{v_{n,j}}}{\sum_{h\in\mathcal{J}} \mathrm{e}^{v_{n,h}}} \qquad \forall\ n \in \mathcal{N}; j \in \mathcal{J} \tag{4}$$

can be derived from (3) (Train 2009). A fundamental property of the MNL is the so–called irrelevance of independent alternatives (IIA). Consider the ratio of any two alternatives' choice probabilities (odds):

$$\frac{p_{n,j}}{p_{n,h}} = \frac{\mathrm{e}^{v_{n,j}} / \sum_{k\in\mathcal{J}} \mathrm{e}^{v_{n,k}}}{\mathrm{e}^{v_{n,h}} / \sum_{k\in\mathcal{J}} \mathrm{e}^{v_{n,k}}} = \frac{\mathrm{e}^{v_{n,j}}}{\mathrm{e}^{v_{n,h}}} = \mathrm{e}^{v_{n,j}-v_{n,h}}. \tag{5}$$

This means that the ratio of choice probabilities of two alternatives does not depend on the probabilities for any other alternative. The most important advantage of this property is the implication that alternatives (here: products) can be added to or removed from the choice set without giving rise to the structure of the MNL (Koppelman & Bhat 2006).
For our purposes, the deterministic utility is mainly describable as a linear function of an individual's gender, age and income as well as the product's features and price.

2 Mixed Integer Linear Product Portfolio Design Problem

Sets and Indices

$\mathcal{I}$	set of customer segments, index i
$\mathcal{M}$	set of all products (existing competitive and own as well as non–existing possible new own), index m
$\mathcal{J}$	set of possible new products $\mathcal{J} \subset \mathcal{M}$, index j

Parameters

k	stipulated total number of new own products
r_j	unit price of product j
c_j	unit cost of product j
f_j	fixed cost of offering product j
w_i	number of customers in customer group i
$p_{i,j}$	choice probability of customer group i for buying product j given that j is the only own product in the market, i. e., the choice set consists of the two alternatives $\{j; \text{no}\}$ which results in $p_{i,j} = \frac{\mathrm{e}^{v_{i,j}}}{\mathrm{e}^{v_{i,\text{no}}}+\mathrm{e}^{v_{i,j}}}$, where $v_{i,j}$ is the deterministic utility of customer group i for buying product j and the none option captures the consumption of competitive products as well as not buying any product

Variables

Z_i	cumulative choice probability of customer group i of not buying products of own portfolio ("no–choice")
$X_{i,j}$	choice probability of customer group i of buying product j
Y_j	$= 1$ if product j is part of the own portfolio; 0 otherwise
F	objective function value (expected profit)

Originally,

$$X_{i,j} = \frac{\mathrm{e}^{v_{i,j}} \cdot Y_j}{\sum\limits_{m \in \mathcal{M}} \mathrm{e}^{v_{i,m}} \cdot Y_m} \qquad \forall\, i \in \mathcal{I}; j \in \mathcal{J}, \tag{6}$$

which yields a mixed integer non–linear program (see the following model). We reformulate $X_{i,j}$ in a linear way by utilizing the IIA property:

Model Formulation

$$\text{Maximize} \quad F = \sum_{j \in \mathcal{J}} \left((r_j - c_j) \cdot \sum_{i \in \mathcal{I}} w_i \cdot X_{i,j} - f_j \right) \tag{7}$$

subject to

$$Z_i + \sum_{j \in \mathcal{J}} X_{i,j} \leq 1 \qquad \forall\, i \in \mathcal{I} \tag{8}$$

$$X_{i,j} \leq p_{i,j} \cdot Y_j \qquad \forall\, i \in \mathcal{I}; j \in \mathcal{J} \tag{9}$$

$$X_{i,j} \leq \frac{p_{i,j}}{1 - p_{i,j}} \cdot Z_i \qquad \forall\, i \in \mathcal{I}; j \in \mathcal{J} \tag{10}$$

$$\sum_{j \in \mathcal{J}} Y_j = k \tag{11}$$

$$X_{i,j} \geq 0 \qquad \forall\, i \in \mathcal{I}; j \in \mathcal{J} \tag{12}$$

$$Z_i \geq 0 \qquad \forall\, i \in \mathcal{I} \tag{13}$$

$$Y_j \in \{0; 1\} \qquad \forall\, j \in \mathcal{J} \tag{14}$$

The objective function (7) maximizes the expected profit in a market with competitive as well as own existing and new products. According to Analogous to Haase & Müller (2015, p. 277), (8)–(10) together with (7) are a linear reformulation of the MNL choice probabilities (6). (8) ensure that a customer group i's final choice probabilities for buying our own products as well as the probability of buying competitors' products or nothing at all sum up to at most 1. The linking constraints (9) allow choice probabilities for a product to be greater than 0 only if the corresponding product is actually offered. Allowing for $p_{i,j}$ yields to a smaller upper bound by the corresponding LP–relaxation than just using $X_{i,j} \leq Y_j$ and to tighter bounds for $X_{i,j}$ (Haase & Müller 2015), because $p_{i,j}$ is distinctly smaller than 1.
(10) ensure that the pre–calculated constant substitution ratios between the choice probabilities for any two alternatives are obeyed. They are derived from $\frac{X_{i,j}}{Z_i} = \frac{p_{i,j}}{1-p_{i,j}}$. But, $X_{i,j} \neq p_{i,j}$ and $Z_i \neq (1 - p_{i,j})$ (unless j is the only established facility).

(11) ensures that k new own products are offered and (12)–(14) define the variables' domains.

References

Haase, K. & Müller, S. (2015). Insights into clients' choice in preventive health care facility location planning. *OR Spectrum*, *37*(1), 273–291.

Koppelman, F. S. & Bhat, C. (2006). A self instructing course in mode choice modeling: multinomial and nested logit models. *US Department of Transportation, Federal Transit Administration*, *31*.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press, New York.

Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.