European Journal of Operational Research 241 (2015) 564–574

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Advanced conjoint analysis using feature selection via support vector machines

Sebastián Maldonado[a,*], Ricardo Montoya[b], Richard Weber[b]

[a] Universidad de los Andes, Mons. Álvaro del Portillo 12455, Las Condes, Santiago, Chile
[b] Department of Industrial Engineering, Universidad de Chile, Av. República 701, Santiago, Chile

ARTICLE INFO

Article history:
Received 22 January 2014
Accepted 25 September 2014
Available online 25 October 2014

Keywords:
Conjoint analysis
Feature selection
Support vector machines
Business analytics

ABSTRACT

One of the main tasks of conjoint analysis is to identify consumer preferences about potential products or services. Accordingly, different estimation methods have been proposed to determine the corresponding relevant attributes. Most of these approaches rely on the post-processing of the estimated preferences to establish the importance of such variables. This paper presents new techniques that simultaneously identify consumer preferences and the most relevant attributes. The proposed approaches have two appealing characteristics. Firstly, they are grounded on a support vector machine formulation that has proved important predictive ability in operations management and marketing contexts and secondly they obtain a more parsimonious representation of consumer preferences than traditional models. We report the results of an extensive simulation study that shows that unlike existing methods, our approach can accurately recover the model parameters as well as the relevant attributes. Additionally, we use two conjoint choice experiments whose results show that the proposed techniques have better fit and predictive accuracy than traditional methods and that they additionally provide an improved understanding of customer preferences.

1. Introduction

Conjoint analysis is one of the research techniques most widely used to identify customers' preferences (see e.g. Green, Krieger, & Wind, 2001). Firms' decisions regarding new product or service design (Kohli & Krishnamurti, 1989) as well as promotional and advertising campaigns increasingly rely on its outputs. Usually, the estimated preferences are the inputs for market simulation techniques that are then used to evaluate different market opportunities. Additionally, conjoint analysis allows estimating consumers' willingness to pay (WTP), defined as the price of indifference between buying and not buying (Gensler, Hinz, Skiera, & Theyson, 2012), and thus it helps to make important pricing decisions. Consequently, appropriate conjoint studies and their derived implications can determine the success or failure of new product introductions or marketing campaigns.

Originally developed in marketing (Green & Rao, 1971), conjoint analysis has had an increasing impact in many other disciplines such as health care (Bridge et al., 2011; Halme & Kallio, 2011), tourism management (Thyne, Lawson, & Todd, 2006), transportation (Hensher, Louviere, & Swait, 1998), and operations management (Dobson & Kalish, 1993), among others. Further applications where this technique has been successfully employed have been presented by Karniouchina, Moore, van der Rhee, and Verma (2009).

In addition, conjoint analysis is relevant for the Operations Research community for at least the following two reasons. First, conjoint analysis can be used in the context of multi-attribute decision making (MADM), since multiple attributes are considered in a preference measurement process. A comparison to an alternative MADM technique (Analytic Hierarchy Process) has been presented in Scholl, Manthey, Helm, and Steiner (2005). Second, optimization techniques are used in the context of conjoint analysis (see e.g. Camm, Cochran, Curry, & Kannan, 2006; Halme & Kallio, 2011, 2014). This fact constitutes an opportunity to develop different types of advanced optimization models to increase the applicability of conjoint analysis.

One of the main outputs of conjoint analysis is to identify the relevant attributes at the consumer level. That is, the (subset of) attributes that the consumer considers when evaluating the proposed alternatives. The usual approach to obtain such subset of attributes (or their ranking) is by post-processing the estimated parameters. For instance, the relative range of part-worths can be used to represent attribute importance when using additive models such as a mixed logit model. Such post-processing task implicitly assumes that consumers use all attributes when facing a conjoint decision. However, as shown later, traditional models can have problems eliminating irrelevant attributes across consumers, especially when there is limited individual-level data. Indeed, despite current developments

* Corresponding author. Tel.: +56 9 61704167.
E-mail address: smaldonado@uandes.cl (S. Maldonado).

http://dx.doi.org/10.1016/j.ejor.2014.09.051
0377-2217/

Table 1
Summary of the relevant literature in Support Vector Machines (SVM), Feature Selection (FS), and Conjoint Analysis (CA)

References	SVM	FS	CA
Schoelkopf and Smola (2002); Vapnik and Chervonenkis (1991)	√		
Blum and Langley (1997); Fan and Li (2001); Song, Smola, Gretton, Bedo, and Borgwardt (2012)		√	
Arora and Huber (2001); Ben-Akiva and Lerman (1985); Dzyabura and Hauser (2011); Gelman and Pardoe (2006); Gilbride and Allenby (2006); Green et al. (2001); Hauser, Toubia, Evgeniou, Befurt, and Dzyabura (2010); Jedidi, Montoya, and Kohli (2013); Kohli and Krishnamurti (1989); Rossi, Allenby, and McCulloch (2005); Toubia, Hauser, and Garcia (2007b)			√
Bradley and Mangasarian (1998); Guyon and Elisseeff (2003); Maldonado and Weber (2009); Maldonado, Weber, and Basak (2011)	√	√	
Chapelle and Harchaoui (2005); Cui and Curry (2005); Evgeniou et al. (2005); Evgeniou, Pontil, and Toubia (2007); Toubia, Evgeniou, and Hauser (2007a)	√		√
Argyriou, Evgeniou, and Pontil (2008)		√	√
This study	√	√	√

in choice modeling that incorporate non-compensatory preferences, typical models based on conjoint analysis do not allow for "attribute non-attendance" (Hensher, Rose, & Greene, 2012). This occurs when customers completely neglect some attributes and focus their attention on a small subset of attributes. As conjoint analysis studies have been incorporating more complex products that are characterized by a larger number of attributes and at the same time more data are available, it is expected that many consumers be more selective regarding the attributes they really consider. Our proposed model contributes in filling this gap in the academic literature and also aims at providing a useful contribution to practitioners.

Several approaches from data mining and machine learning have been presented in the last decade in order to achieve better predictive performance in conjoint analysis (Evgeniou, Boussios, & Zacharia, 2005) and accurate representations of consumer preferences. These approaches have proved to provide important insights and consequently have gained reputation as valid methods to uncover customers' preferences. However, they do not address the problem of effectively and efficiently selecting the relevant attributes used by consumers in their evaluation tasks. Attribute (or feature) selection has proved to be an important characteristic that predictive models need to include (see e.g. Blum & Langley, 1997; Guyon & Elisseeff, 2003). Not only because of a more parsimonious representation but also because it can better identify true underlying preferences that can lead to a higher predictive ability of consumer decisions. Table 1 presents an overview of the relevant literature studied in this work.

We present a novel technique based on Support Vector Machines (SVM) to determine the relevant attributes for estimating customer preferences. The identification of the relevant attributes that customers use to evaluate products, with the corresponding reduction in the dimensionality of customers' utility functions, is achieved by a backward elimination of attributes procedure based on the individual part-worths. Therefore, such attribute selection is performed simultaneously to the estimation of customers' preferences. An extensive simulation exercise shows that the proposed approach outperforms existing methods for attribute selection in the context of choice-based conjoint analysis.

The contribution of the paper is twofold: (i) it presents a framework that simultaneously identifies the most relevant attributes when estimating customer preferences, and (ii) it shows that the understanding of customers' preferences and the predictive performance of the proposed approach can be enhanced considering the most relevant attributes.

The remainder of the paper is organized as follows. Section 2 discusses previous work. In particular, it describes SVM for CBC and provides a general overview of the different attribute selection approaches for SVM. The proposed method for attribute selection based on SVM for conjoint analysis is introduced in Section 3. In Section 4 we present the results of a simulation exercise that underline our method's capabilities. Section 5 describes the application of the proposed approaches in two empirical conjoint studies highlighting the managerial implications that can be derived from the respective analyses. Section 6 summarizes the key results and discusses directions for future research.

2. Previous work

SVM were introduced to conjoint analysis by Evgeniou et al. (2005) and Cui and Curry (2005). Evgeniou et al. (2005) showed that SVM are accurate, robust to noise, and computationally efficient in a conjoint analysis context. Cui and Curry (2005) found that the predictive ability of SVM outperforms competing models such as multinomial logit models in consumer choice experiments. Later, Evgeniou et al. (2007) developed a convex approach for modeling consumer heterogeneity (Natter & Feurstein, 2002) in conjoint analysis, and compared it to Hierarchical Bayes (HB) methods. To the best of our knowledge, the present paper is the first work that adds feature selection to SVM for conjoint analysis.

Section 2.1 describes SVM in the context of choice-based conjoint analysis (CBC) (Chapelle & Harchaoui, 2005; Cui & Curry, 2005; Evgeniou et al., 2005). In Section 2.2 we present the state-of-the-art regarding feature selection using SVM.

2.1. Support vector machines for choice-based conjoint analysis

Consider a product profile with J attributes. Each attribute is defined over n_j levels, $j = 1, \ldots, J$. Suppose a consumer evaluates the profiles of K different products and chooses one profile in each of T choice occasions. Finally, consider a sample of N customers.

Customer i's preferences are modeled by an additive utility function, which is assumed to be a linear combination of the partial utilities (part-worths): $u_i(\mathbf{x}) = \mathbf{w}_i{}^T \cdot \mathbf{x}$, $i = 1, \ldots, N$.

We consider CBC data with the following information $([\mathbf{x}_{it}^1, \ldots, \mathbf{x}_{it}^K], y_{it})$, where $\mathbf{x}_{it}^k \in \Re^J$ and $y_{it} \in \{1, \ldots, K\}$ for $1 \le i \le N$, $1 \le t \le T$, and $1 \le k \le K$. The choice $y_{it} = k$ indicates that at occasion t, consumer i prefers the kth alternative among the K product profiles described by $[\mathbf{x}_{it}^1, \ldots, \mathbf{x}_{it}^K]$. That is, $u_i(\mathbf{x}_{it}^{y_{it}}) \ge u_i(\mathbf{x}_{it}^b)$, $\forall b \in \{1, \ldots, K\} \setminus \{y_{it}\}$ (Chapelle & Harchaoui, 2005). Without loss of generality, and following previous research, we assume that for each choice occasion t all customers choose the first profile, i.e. $y_{it} = 1$, $1 \le i \le N$ and $1 \le t \le T$. Thus, the inequalities can be rewritten as

$$\mathbf{w}_i{}^T \cdot (\mathbf{x}_{it}^1 - \mathbf{x}_{it}^k) \ge 0, \qquad (1)$$

where $1 \le i \le N$, $2 \le k \le K$, and $1 \le t \le T$.

To determine the weights $\mathbf{w}_i$ the *structural risk minimization principle* (Vapnik & Chervonenkis, 1991) has been considered. This approach minimizes the Euclidean norm of $\mathbf{w}_i$, with noise penalization via slack variables ξ_{kt} (l_2-soft margin formulation) that leads to the following quadratic programming problem for each customer $i = 1, \ldots, N$ (Chapelle & Harchaoui, 2005; Evgeniou et al., 2005):

$$\min_{\mathbf{w}_i, \xi} \quad \frac{1}{2}\|\mathbf{w}_i\|^2 + C \sum_{t=1}^{T} \sum_{k=2}^{K} \xi_{kt} \qquad (2)$$

s.t.

$$\mathbf{w}_i{}^T \cdot (\mathbf{x}_{it}^1 - \mathbf{x}_{it}^k) \ge 1 - \xi_{kt} \quad t = 1, \ldots, T; \ k = 2, \ldots, K.$$

$$\xi_{kt} \ge 0 \quad t = 1, \ldots, T; \ k = 2, \ldots, K.$$

Model (2) minimizes ξ_{kt} that represent inconsistencies in the choice data. This formulation simultaneously controls for the complexity of the model by maximizing the margin ($\propto 1/\|\mathbf{w}_i\|^2$). The