Generalized Robust Conjoint Estimation

Giorgos Zacharia, Theodoros Evgeniou & Costas Boussios

Artificial Intelligence Laboratory and
The Center for Biological and Computational Learning
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: We develop a framework within which robust models of preferences are computationally efficiently estimated using quadratic optimization methods. Within this framework general highly non-linear models can be computationally efficiently estimated while at the same time avoiding overfitting problems that such models typically have. We compare these models with standard logistic regression and recently proposed polyhedral conjoint methods.

Motivation: Traditional preference modeling methods such as conjoint analysis [4] have been used for many preference data analysis applications [12] but have been more naturally suited for controlled data gathering situations for example through questionnaires. However a lot of web-based information about choices is typically not gathered in such a controlled way and therefore is more noisy and often sparse. It is therefore important to develop a new generation of preference modeling methods that are robust to handle such data, especially for products with a very large number of attributes, and can be computed efficiently without sacrificing the complexity of the models.

Previous Work: The market research community has traditionally approached utility estimation problems through function estimation. Conjoint analysis is one of the main methods for modeling preferences from data [4, 9]. For simplicity, we deal only with full-profile preference data - full product comparisons - for the case of choice based conjoint analysis (CBC) [5] instead of other metric based ones.

Another approach to utility estimation is within the Discrete Choice Analysis field, where typically users' preferences are modeled as random variables of logit models [2, 6].

Finally, a different approach was implemented by [8], reformulated the problem as an ordinal regression estimation, and used Support Vector Machines (SVM) [11] to predict transitive ranking boundaries.

Recent work by Toubia, Simester, and Hauser [7] adresses the problem of choosing questions and estimating preference models using polyhedral methods for preference modeling.

Approach: We reformulate the problem of preference modeling as classifying vectors in the space of differences of the attribute values of the options.

If we label vectors $(\mathbf{x}_i^{c_i} - \mathbf{x}_i^j)$ for $j \neq c_i$ with $+1$, while vectors $(\mathbf{x}_i^j - \mathbf{x}_i^{c_i})$ for $j \neq c_i$ with -1, then searching for a utility function that satisfies the comparison constraints can be seen as equivalent to searching for a hyperplane that separates the vectors with the $+1$ labels from the ones with the -1. Our formulation is then equivalent to the well known SVM classification method [11], applied on the vectors of differences with the labels defined as above.

Difficulty: Conjoint analysis and discrete choice empirisists have always faced the trade off between model complexity and ease of model estimation ("curse of dimensionality" [10]: as the number of dimensions increases, an exponential increase in sample size is needed to maintain reliable model estimation).

We also explore ways that incorporate prior knowledge about the modeled domain to guide the training of our models. We use virtual examples to incorporate positivity constraints (if a product is prefered from another one, then a slightly better product is also prefered) in the estimation method [1].

Impact: The robustness of the methodology enables the estimation of preference models when the data is noisy and the number of attributes describing the choices is very large while the amount of example choices - past information - is small. The approach can therefore be useful also for analyzing data from very small questionnaires regarding multi-attribute choices with large numbers of attributes, or data that are noisy like for example that describing users' choices or clicks on the internet.

Future Work: The machinery developed for SVM as well as statistical learning theory can be used for solving in new ways problems in the field of conjoint analysis. For example, one can extend the use of virtual examples we used here for adding positivity constraints on the utility function. Empirical evidence for learning from examples settings shows that if the original data used to estimate a model are extended to include virtual examples then the performance of the estimated models increases [1]. Another direction of research is to develop active learning [3] type methods for the problem of adaptively designing questionnaires, like for example in [7].

References:

[1] C. Burges B. Scholkopf and V. Vapnik. Incorporating invariances in support vector learning machines. *Artificial Neural Networks, ICANN'96, Springer Lecture Notes in Computer Science*, 1112, 1996.

[2] M. Ben-Akiva et al. Modeling methods for discrete choice analysis. *Marketing Letters*, 8:273–286, 1997.

[3] M. Warmuth et al. *Active learning in the drug discovery process*, volume 14. 2002.

[4] P. Green and V. Srinivasan. Conjoint analysis in consumer research: Issues and outlook. *Consumer Research*, 5(2):103–123, 1978.

[5] J. Swait J. Louviere, D. Hensher. *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, New York, NY, 2000.

[6] D. McFadden. *Conditional Logit Analysis of Qualitative Choice Behavior*. New York Academic Press, 1974.

[7] D. Simester O. Toubia and J. Hauser. Fast polyhedral adaptive conjoint estimation. *Working paper, Sloan School of Management*, 2001.

[8] K. Obermayer R. Herbrich and T. Graepel. *Large Margin Rank Boundaries for Ordinal Regression*. The MIT Press, 1999.

[9] Sawtooth Software. Hb-reg: Hierarchical bayes regression. *sawtoothsoftware.com/hbreg.shtml*.

[10] C. J. Stone. Additive regression and other nonparametric models. *Annals of Statistics*, 13:689–705, 1985.

[11] V. N. Vapnik. *Statistical Learning Theory*. Wiley, New York, 1998.

[12] D. Wittink and P. Cattin. Commercial use of conjoint analysis: An update. *Journal of Marketing*, 53(3):91–96, 1989.