Ranking Support Matrix Machine for Vehicle Face Prediction

Alex Burnap, Hareen Kancharla, Zheng Shen, Yi Ying Zhu

April 24, 2012

Abstract

We study the relationship between the demographics of an individual, and the preferred geometric shape of the vehicle face (shape of headlights) of a compact vehicle using a novel machine learning algorithm, the ranking support matrix machine (RSMM). We provide derivation of the RSMM for ranking bilinear transformations between two spaces, and theoretical justification behind its advantages over traditional methods used in marketing science. Empirical justification for its use is studied by benchmarking on the commonly used MovieLens data set. The RSMM is then trained on a data set of historical vehicle purchases to predict the preference function between the individuals and vehicle faces. Each vehicle face is then decomposed using Fourier decomposition of edges within vehicle face images. The combination of the RSMM and the face decomposition results in a predictive mapping between an individual and the most appealing vehicle face.

1 Introduction

What features of a car are important to a prospective buyer? That question has determined the success and failure of hundreds of vehicle manufacturers since the dawn of the automobile, and its answer lies in understanding the relationship between characteristics of the buyer and the features of the car. One of the most heavily used methods in ascertaining this relationship is conjoint analysis [1]. This relatively well developed method aims to model the expected utility of the vehicle as a linear combination of vehicle features, with the idea that a prospective buyer will purchase the vehicle with the highest expected utility [2].

Though conjoint analysis has shown empirical success since its inception cite, we show the predictive power of our novel algorithm, the ranking support matrix machine (RSMM), is able to overcome several shortcomings of conjoint analysis. We apply a RSMM to a data set of 30,000 vehicle purchases from 2006. This data set includes the information of 26 distinct demographics for each vehicle buyer including age, gender, education, etc. Additionally, this data set only contains purchase history for compact cars in an effort minimize latent dimensions. Latent dimensions for us are defined as those that exist in our data, but are not in the set of features we assumed to span the space of vehicles. In our case, the set of vehicle features we consider only comprises the "face" of the car, namely the headlights. The headlights were picked to give a visual representation of the predictive performance of our algorithm.

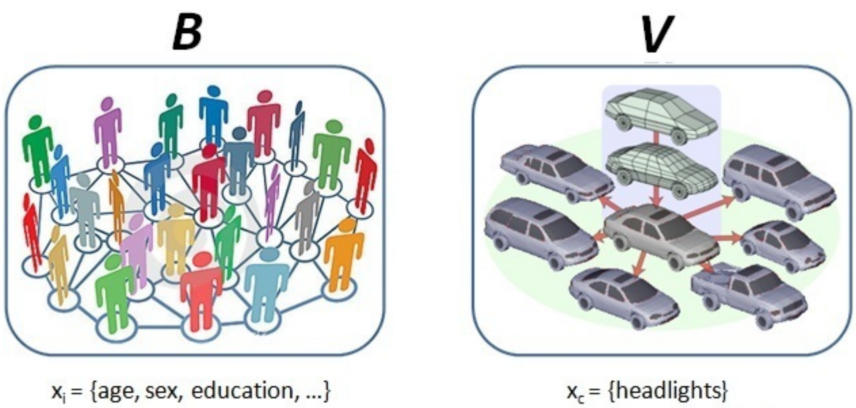
1.1 Advantages of RSMM

Though conjoint analysis and its variants have a proven track record, there exist two major ways in which ideas from machine learning and statistical learning theory can prove to be beneficial. In particular, training the utility function using a support vector machine can deduce a nonlinear kernel space [3], thus circumventing the assumption of independent and irrelevant alternatives necessary for the linear kernel space of conjoint analysis [4]. Secondly, by deducing the utility function from a variational calculus point of view, we may bound the complexity capacity of the utility function's form as well as the error in choosing the form [5]. This has advantages over the traditional method of maximum likelihood parameter deduction for conjoint analysis; the latter has the tendency to over-fit data by erroneously accounting for variance that actually resides in latent dimensions [6].

Our work differs from conjoint analysis in that it deduces the utility function by ranking sets of vehicle features rather than picking off the single highest ranked set of vehicle features. This allows us to understand the discrete distribution over several point estimates of utility from vehicle features rather than just the highest utility point estimate. Moreover, the most significant difference and potentially useful aspect of our work is though taking into account the heterogeneity of prospective vehicle buyers [7]. Traditional conjoint analysis assumes a homogenous population of vehicle buyers, thus clumping diversity of buyer demographics into a single utility function. The RSMM deduces a ranking of utility functions that act as bilinear transformations from the space of prospective vehicle buyers to the space of vehicle features.

2 Definitions

We define the space of car buyers B and the space of vehicles V. We define a single car buyer $\vec{x_i} \in B$ and a single vehicle $\vec{x_c} \in V$, where $\vec{x_i} = \{age, sex, education, \ldots\}$ and $\vec{x_c} = \{c_1 h_1 \ldots c_K h_K\}$. In this case h_k are the basis functions for the headlights and c_k are their respective weights.



Let $g(\vec{x_i}, \vec{x_c}) : B, V \to U$ be a mapping defining a bilinear transformation returning the utility function of an single car buyer. This utility function is interpreted as the appeal that a particular vehicle face has on an particular individual. We are interested in finding a ranked set of utilities corresponding to most appealing vehicle face to least appealing vehicle face. Although it is not possible to assign ordinal rankings, we can infer rankings using pairwise comparisons.

We now define a new space corresponding to all pairwise comparisons of $g_a - g_b$ named R. Since we are using finite data of size n, we define the attainable space R_n and the correct space of utility rankings R_*. The goal is now to find the most optimal approximation of R_n to R_*over the space of all possibilities of g. We define this mapping of functionals as $\mathcal{F}$. Under this framework we assume that given infinite informative data, i.e. independent and identically distributed, we may converge onto the true topology of R and in doing so the true mapping g.

3 Methodology

3.1 Overview

The chief goal of our system is to predict the most appealing vehicle "face", comprised of the headlights and the grill, given the demographic information of an individual. To achieve this task, we chose to use the philosophy of statistical learning theory underlying support vector machines [8, 9]. We assume the variance within a data set is shared by the complexity capacity of our model's functional form and the functions themselves. We can then put a variational bound on the corresponding Vapnik-Chervonenkis dimension of our mapping $\mathcal{F}$. This can be thought of as getting as close to the correct ranked utility space as possible, thus the best predictive power for $\vec{g}(\vec{x_i}, \vec{x_c})$. We can write the ranked utility space as a linear combination of orthogonal support vectors, corresponding to the set of unique factors sufficient for calculating the ranked utility space. The support vectors may be then viewed as active constraints with linear weights, thus the optimally correct ranked utility space becomes a quadratic programming problem with guaranteed global optimality of a KKT point [10]. It may be noted that this problem could be converted to its corresponding dual form, but we have chosen to leave it in its primal due to its advantages with regards to computational speed [11]. Additionally, although the use of nonlinear kernels may be trained in the primal, we only used the linear case. As noted in 1.1, training a nonlinear kernel has the advantage of circumventing the assumption of independent and irrelevant alternatives.

A practical issue then becomes extracting the basis functions making up the face of cars within our data set. This was done by decomposing the shape of the headlights into its most principally important Fourier modes from a designated center point. The shape of the headlights themselves were found using edge detection algorithms on normalized car images.

3.2 RSMM

Derivation of the RSMM followed the formulation of the ranking support vector machine [5] as well as the formulation of bilinear regression [12]. The ranking support vector machine extends the work of the original support vector machine cite by considering the input space as made up of pairwise comparisons [13].

$$\begin{array}{ll} min & \|\vec{\omega}\|^2 + \widetilde{C}\sum \xi_{a,b} \\ s.to & \forall \vec{g}_{a,b} \in X : \vec{\omega}^T \vec{g_a} - \vec{\omega}^T \vec{g_b} \geq 1 - \xi_{a,b} \end{array}$$

As is standard in support vector machine methodology, we rewrite the constrained problem as an unconstrained problem using the hinge loss function.

$$min \qquad \|\vec{\omega}\|^2 + \widetilde{C}\sum \max[0, y_{a,b}(\vec{\omega}^T \vec{g_a} - \vec{\omega}^T \vec{g_b}) - \vec{1} + \xi_{a,b}]$$

By considering the vector of weights $\vec{\omega}$ instead as a matrix $\widetilde{\Omega}$ as well as expanding g, we are able to formulate a bilinear ranking transformation between $\vec{x_i}$ and $\vec{x_c}$. Note that the norm of $\widetilde{\Omega}$ is assumed to be the Frobenius norm, thus preserving convexity.

$$min \qquad \|\widetilde{\Omega}\|^2 + \widetilde{C}\sum \max[0, \vec{y_{a,b}}(\vec{x_{i,a}}^T \widetilde{\Omega} \vec{x_{c,a}} - \vec{x_{i,b}}^T \widetilde{\Omega} \vec{x_{c,b}}) - 1 + \xi_{a,b}]$$

This can be equivalently written in the form of a ranking support vector machine by considering instead the Kronecker product of $\vec{x_i}$ and $\vec{x_c}$ and breaking the matrix $\widetilde{\Omega}$ into a vector $\hat{w}$ the length of the number of elements of $\widetilde{\Omega}$,

$$min \qquad \|\vec{\hat{\omega}}\|^2 + \widetilde{C}\sum \max[0, \vec{y_{a,b}}(\omega^T (\vec{x_{c,a}} \otimes \vec{x_{i,a}}) - \vec{\omega}^T (\vec{x_{c,b}} \otimes \vec{x_{i,b}})) - 1 + \xi_{a,b}]$$

3.3 Feature Extraction

The feature extraction portion of our work involved applying image processing techniques on vehicle images from a head-on view, identifying the edges of the headlights, and decomposing them into a Fourier expansion of modes. This representation of features using Fourier basis functions falls under the methodology of shape descriptors [14]. Shape descriptors were chosen when compared to two other encoding methods [15, 16] as shown in Table 1.

	Size of Encoded Data	Fidelity of Curvature	Robustness to Noise
Image Encoding	High (All Pixels)	Low Fidelity	High
Shape Descriptors	Low (Boundary Size x 2)	High Fidelity	High
Relief Graphs	Low (Edge Size x 2)	Medium Fidelity	Low

Table 1: Comparison of three encoding methods.

A flow diagram of feature extraction process is shown in Figure 1. The performance of our feature extraction then depends on three factors- quality of headlight edge detection, accuracy in choosing the centroid of the headlight, and number of Fourier modes considered in the expansion.

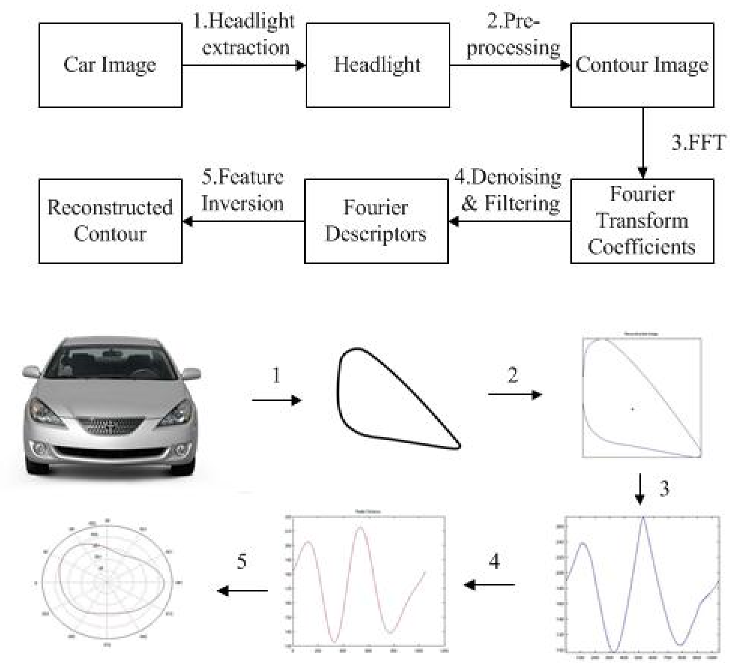


Figure 1: Flow diagram of the feature extraction process. Edge detection of headlights is done on the raw vehicle image. The centroid is then calculated and the distance from edge to centroid is expanded into Fourier modes.

The centroid distance function represents the difference between the centroid of the shape to the boundary coordinates of the image. The boundary coordinates (x_b, y_b) are obtained by using the eight connection contour tracking technique using the Palividis Algorithm [17]. The boundary trace is then normalized to a fixed number of samples N. The centroid (x_c, y_c) for this arbitrary shape is given by,

$$x_c = \frac{1}{N}\sum x_b \qquad y_c = \frac{1}{N}\sum y_b$$

We then write the distance r_b from each point along the edge of the headlight (x_b, y_b) to the centroid (x_c, y_c) as simple Euclidean distance,

$$r_b = \sqrt{(x_b - x_c)^2 + (y_b - y_c)^2}$$

We expand this set of distances into Fourier descriptor coefficients using,

$$F_n = \frac{1}{N} \sum r(b) \exp\left(\frac{-i2\pi nb}{N}\right), \quad n = 0, 1, \ldots N-1$$

The lower frequency descriptors contain information about the general features of the shape, while the higher frequency descriptors contain information about finer details of the shape and noise produced during edge detection. These high frequency descriptors are set to zero, and inverse discrete Fourier transform is applied to reconstruct a shape using only the low frequency descriptors. As the number of features are increased the reconstructed image tends towards the original image as shown in Figure 2.

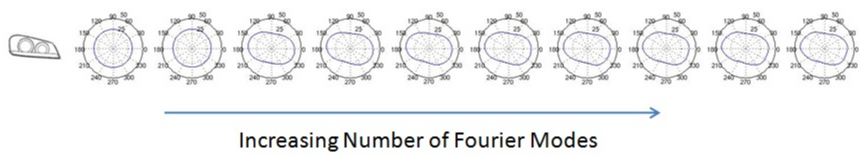


Figure 2: Reconstructed headlight shape with increasing number of Fourier modes.

4 Results

4.1 Benchmarking

To determine fidelity of the RSMM, we trained our model using the commonly benchmarked MovieLens data set. Our metric of performance was the Normalized Discounted Cumulative Gain (NDCG). As shown in Table 2, the RSMM performed roughly around the mean of other currently state of the art algorithms. This result was obtained without slack variable optimization.

Algorithm	NDCG
Random	0.4158
Affinity Model	0.4489
Pairwise Regression	0.4972
RSMM	0.4614

Table 2: Table of ranking performance for current state of the art algorithms.

4.2 Vehicle Purchase Data Set

The RSMM was first run on the vehicle purchase data set to train $\vec{\omega}$. Next, two arbitrarily picked car buyers$\vec{x}$ were input along with all 82 models of vehicle $\vec{x_c}$. The output ranking of the top 5 most appealing vehicles is shown in Table 3.

Rank	$\vec{x_1}$	$\vec{x_2}$
1	Lexus ES	Volvo S40
2	Lincoln LS	Lexus LS
3	Acura TL	Chevy Impala
4	VW Phaeton	Suzuki Forenza
5	Pontiac GTO	Jaguar S-Type

Table 3: Ranked output of two arbitrarily chosen car buyers from testing data set.

By taking a linear combination of the top 5 vehicles in the ranking, we are able to to assume non-point estimate behavior of predictions. Though the proper weighting between these vehicles is an optimization problem in its own right, we take an arbitrarily chosen weighting of,

$$Overall\ Preference = 0.5\,(Rank\ 1) + 0.3\,(Rank\ 2) + 0.1\,(Rank\ 3) + 0.05\,(Rank\ 4) + 0.05\,(Rank\ 5)$$

Using this weighting on preferred vehicles, we then expand each vehicle into its associated Fourier modes describing headlight shape. The resulting predicted headlight shape is shown next to the true headlight of the actual purchased vehicle.

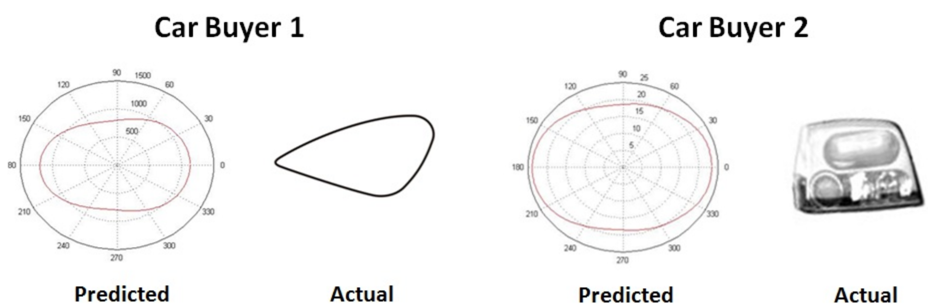


Figure 3: Comparison of predicted headlight shape to actual headlight shape for two separate car buyers.

5 Extensions

We have shown the fidelity of the RSMM algorithm on the MovieLens data set, and its application to marketing science by predicting preferred geometric shape of vehicle headlights. A natural extension along the automotive marketing science route would be the predict entire vehicle preferences. This would not be limited to purely geometric considerations, as other arguably more influential features such as gas economy and vehicle acceleration may be taken into account. The applications of the RSMM are of course not limited to marketing science. For example, recommendation systems would be prime candidates for application as one could rank relevancy among a great number of products.

In terms of algorithmic extensions, there still exist obvious means to increase accuracy and computational speed. Optimization of the slack variables $\widetilde{C}$ provides regularization on the constraint set, and thus controls a tradeoff that may lead to higher accuracy. This optimization can be done by sweeping over terms for all $\widetilde{C}$ using cross validation. If expert knowledge within a domain is accessible, the ratios within $\widetilde{C}$ corresponding to differences in constraint importance may be changed to better reflect the true physics behind a problem. Computational speed may be increased for certain problems if sparsity is present within the data sets by using a conjugate linear gradient methodology to solve the quadratic programming problem instead of the Newton method in our study. Finally, training a nonlinear kernel may have the ability to achieve higher accuracy using either the primal or dual forms of the RSMM.

References

[1] P. Green, “Conjoint analysis in consumer research: issues and outlook,” Journal of consumer research, vol. 5, no. 2, pp. 103-123, 1978.

[2] J. R. Hauser, “Conjoint Analysis , Related Modeling , and Applications Conjoint Measurement or Conjoint Analysis,” 2002.

[3] B. Laboratories, L. Technologies, and C. J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” Pattern Recognition, vol. 43, pp. 1-43, 1997.

[4] O. Netzer et al., “Beyond conjoint analysis: Advances in preference measurement,” Marketing Letters, vol. 19, no. 3–4, pp. 337-354, Jul. 2008.

[5] C. Cortes and V. Vapnik, “Support-vector networks,” Machine learning, vol. 297, pp. 273-297, 1995.

[6] Bishop, Christopher M. Pattern Recognition and Machine Learning. New York: Springer, 2006. Print.

[7] T. Evgeniou, M. Pontil, and O. Toubia, “A Convex Optimization Approach to Modeling Consumer Heterogeneity in Conjoint Estimation,” Technology Management, pp. 1-62, 2007.

[8] Vapnik, V.: Statistical Learning Theory. John Wiley, New York (1998)

[9] O. Bousquet and S. Boucheron, “Introduction to statistical learning theory,” Lectures on Machine Learning, 2004.

[10] Papalambros, Panos Y., and Douglass J. Wilde. Principles of Optimal Design. Cambridge UP, 1988. Print.

[11] O. Chapelle and S. S. Keerthi, “Efficient algorithms for ranking with SVMs,” Information retrieval, vol. 13, no. 3, pp. 201–215, 2010.

[12] S. Park, “Pairwise preference regression for cold-start recommendation,” Proceedings of the third ACM conference on, pp. 21-28, 2009.

[13] T. Joachims, “Optimizing search engines using clickthrough data,” in Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, 2002, pp. 133–142.

[14] Dengsheng Zhang, Guojun Lu, A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval, Journal of Visual Communication and Image Representation, Volume 14, Issue 1, March 2003, Pages 39-57, ISSN 1047-3203

[15] M. Yang, K. Kpalma, and J. Ronsin, “A survey of shape feature extraction techniques,” Methods, vol. 2008, 2008.

[16] Selin Baskan, M.Mete Bulut, Volkan Atalay, Projection based method for segmentation of human face and its evaluation, Pattern Recognition Letters, Volume 23, Issue 14, December 2002, Pages 1623-1629

[17] T. Pavlidis. Algorithms for Graphics and Image Processing. Computer Science Press, 1982, pp.143.