Factorial Tree Models of Brand Switching Data

Thomas P. Novak

Introduction

A wide variety of methods for categorical data analysis have been used to model brand switching data (Charnes, Cooper, Learner & Phillips 1984; Zufryden 1986; Carpenter & Lehmann 1985; Colombo & Morrison 1989; Grover & Srinivasan 1987; Jain, Bass & Chen 1990; Novak 1993; Novak & Stangor 1987). When brand switching data consist of observations from *k* brands over two independent time points, a *k*-by-*k* contingency table, N, is generated, and log-linear models can be used to model the data. Novak (1993) describes a reparameterization of the log-linear model of quasi-symmetry, which allows brand switching data to be represented in terms of additive trees. Here we briefly describe the model used by Novak (1993) and demonstrate how it can be applied to the French car switching data.

The Model

The log-linear model of quasi-symmetry,

$$\ln(n_{ij}) = \mu + \mu_{1(i)} + \mu_{2(j)} + \mu_{12(ij)} \quad \text{where: } \mu_{12(ij)} = \mu_{21(ji)} \tag{1}$$

can, with an appropriately chosen design matrix, be expressed as a Poisson log-linear model (e.g., Agresti, p. 438):

$$E_A[\ln(\underline{\mathbf{n}})] = \mathbf{X}\beta \tag{2}$$

where $\underline{\mathbf{n}}$ is a $k^2\times1$ vector of counts obtained by concatenating the transpose of row from **N**, and E_A denotes asymptotic expectation.

Now, consider a vector, $\underline{\lambda}$, containing the set of $k(k-1)/2$ log-odds ratios defined by:

$$\lambda_{ij} = \ln(n_{ii}n_{jj}/n_{ij}n_{ji}) \quad \text{for all } i > j\ . \tag{3}$$

Let **C** be the $k(k-1)/2$-by-k^2 coefficient matrix which creates this vector, $\underline{\lambda}$, of log-odds ratios:

$$\underline{\lambda} = C[\ln(\underline{\mathbf{n}})]. \tag{4}$$

In addition to the log-linear model in expression (2), consider a second linear model which fits a design matrix, **X***, to the vector $\underline{\lambda}$ rather than to the vector $\ln(\underline{\mathbf{n}})$:

$$E_A[\lambda] = \mathbf{X}^*\beta^*. \tag{5}$$

We can combine expressions (4) and (5) to obtain:

$$CE_A[\ln(\underline{\mathbf{n}})] = \mathbf{X}^*\beta \tag{6}$$

so that

$$E_A[\ln(\underline{\mathbf{n}})] = [X^m \,||\, C^+X^*][\underline{\beta}^{m\prime} \,||\, \underline{\beta}^{*\prime}]', \tag{7}$$

where $\mathbf{C}^+$ is the Moore-Penrose inverse, $\mathbf{X}^m$ are $2(k-1)+1$ columns of the design matrix specifying an intercept and main effects, $\mathbf{C}^+\mathbf{X}^*$ are $k(k-1)/2$ columns of the design matrix specifying interaction terms, $\underline{\beta}^m$ and $\underline{\beta}^*$ are parameters respectively for main effects and interaction, and || specifies horizontal matrix concatenation. Expression (7) is therefore a reparameterized log-linear model, where the model is fit to log-counts, but the interaction parameters $\underline{\boldsymbol{\beta}^*}$ can be interpreted as contributions to log-odds ratios.

Example

Novak (1993) shows that a design matrix **X*** specifying an additive tree (Sattath & Tversky 1977) generates, through the transformation $\mathbf{C}^{+}\mathbf{X}^{*}$, a restricted version of the quasi-symmetry model. Using the stepwise WLS algorithm described in Novak (1993), a "factorial additive tree model" was identified. This model is graphically represented in Figure 1. This model involved developing a design matrix, **X***, in expression (5), where the parameters $\underline{\beta}^*$ corresponded to arc-lengths for unique and common features in an additive tree. Arc lengths were identified which attempted to minimize the WLS residual chi-square statistic:

$$\chi^2 = (\underline{\hat{\lambda}} - X^*\underline{\hat{\beta}}_w^*)'\, Cov(\hat{\lambda})^{-1}\, (\hat{\lambda} - X^*\underline{\hat{\beta}}_w^*) \qquad (8)$$

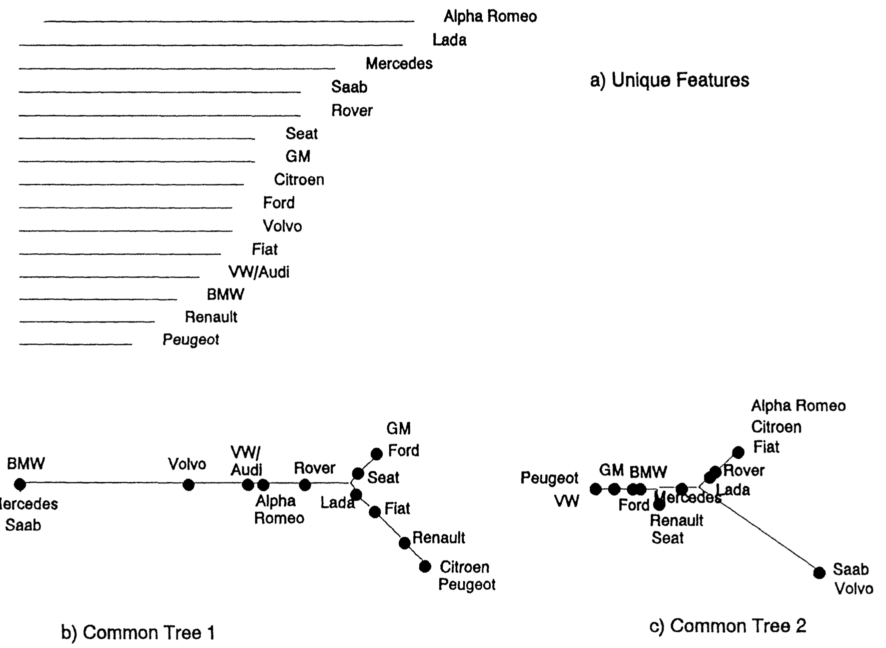

Figure 1: Factorial Tree Representation

The tree representation is interpreted as follows. Sattath and Tversky (1977) and Carroll & Pruzansky (1975) note that dissimilarities which are not representable by a single additive tree may be decomposed as the sum of incompatible (non-nested) additive trees. Figure 1 shows a structure which has been decomposed into three additive trees. Figure a represents unique features corresponding to the *k* diagonal parameters in the quasi-independence model. For clarity, these are drawn as horizontal lines. Technically speaking, these arcs all originate from a single node like spokes in a wheel, and define a singular additive tree. Figure b represents common features organized in a hierarchical tree, and Figure c represents an additional hierarchical tree which is not nested within the first. These two overlapping (non-nested) hierarchical tree structures are referred to as "factorial trees."

The log-odds ratio between a pair of brands, *i* and *j*, is approximated by the sum of arc-lengths between these two brands, added up over each of the three sub-trees. Larger log-odds correspond to greater dissimilarity between the two brands (i.e., less switching) while smaller log-odds correspond to greater similarity between brands (i.e., more switching). Consider Peugeot and Renault. The two are relatively close together in Figures a, b, and c, indicating relative similarity (high degree of switching) between these two cars.

While the factorial tree representation can be used to approximate log-odds ratios, it is perhaps more useful as a guide to unique and common features underlying the log-odds ratios. Figure a shows which cars are, on the whole, more (small arc length) or less (large arc length) likely to switch with other cars. Thus, Alpha Romeo and Lada are less likely, while Renault and Peugeot are more likely. Figures b and c identify common feature structures. Without providing detailed explanation, Figure b shows similarities among expensive luxury cars (BMW, Mercedes, Saab), American cars (GM and Ford), and French cars (Renault, Citroen, Peugeot). Figure c mainly shows the similarity among Saab and Volvo.

A series of log-linear models were fit to the French car switching data, and are summarized in Table 1. Likelihood-ratio chi-square, G^2, residual degrees-of-freedom, and AIC (Kumar & Sashi 1989; Agresti 1990) are reported, where AIC is calculated as:

$$\mathrm{AIC} = G^2 + 2(k^2 - \mathrm{d.f.}) \quad (9)$$

Results for log-linear models of independence, quasi-independence, and quasi-symmetry are reported, as are results for a series of log-linear tree models. These log-linear tree models are reported in order of increasing complexity as arc lengths are added in turn to the first hierarchical tree (see Figure b) and the second hierarchical tree (see Figure c). Note that the models for the second hierarchical tree include all 12 arc lengths from the first hierarchical tree. While the tree-generating algorithm is based upon WLS estimation and fits the model in expression (5), the design matrix **X*** was used in expression (7) with maximum likelihood estimation to obtain the chi-square values in Table 1.

For the models fit, minimum AIC is achieved for model with a) 15 unique brand features, b) a hierarchical tree with 12 common features, and c) an additional hierarchical tree with 8 features. (This model corresponds to collapsing BMW, Renault, Seat, and Mercedes in Figure c, and eliminating the short arc to the immediate right of Mercedes). Table 1 indicates that the factorial tree representation in Figure 1 provides a good approximation of the more complex quasi-symmetry model. Note that a *single* hierarchical tree (corresponding to Figures a and b) has AIC of 406.6, and is not by itself an adequate approximation of the quasi-symmetry model (AIC of 397.8); a second, non-nested tree (i.e., Figure c) is also needed.

Table 1: Goodness of fit for log-linear models

Model	Likelihood Ratio Chi-Square	d.f.	AIC
Independence	22099.9	196	22157.9
Quasi-Independence	783.6	181	783.6
Quasi-Symmetry	129.8	91	397.8
First Hierarchical Tree:			
1 arc	555.3	180	645.3
2 arcs	513.4	179	605.4
3 arcs	477.1	178	571.1
4 arcs	372.3	177	468.3
5 arcs	356.4	176	454.4
6 arcs	344.0	175	444.0
7 arcs	315.0	174	417.0
8 arcs	308.7	173	412.7
9 arcs	304.8	172	410.8
10 arcs	299.4	171	407.4
11 arcs	298.3	170	408.3
12 arcs	294.6	169	406.6
Second Hierarchical Tree:			
1 arc	280.3	168	394.3
2 arcs	270.1	167	386.1
3 arcs	262.8	166	380.8
4 arcs	255.5	165	375.5
5 arcs	250.3	164	372.3
6 arcs	245.7	163	369.7
7 arcs	240.3	162	366.3
8 arcs*	237.9	161	365.9*
9 arcs	237.5	160	367.5
10 arcs	237.2	159	369.2
11 arcs	237.7	158	371.7
12 arcs	235.6	157	371.6

*minimum AIC

Conclusion

The log-linear model of quasi-symmetry can be expressed as a reduced model, where interaction parameters are interpreted as arc lengths in an additive tree fit to log-odds ratios. Log-odds ratios have a clear interpretation as a dissimilarity measure among brands. Further, as described in Novak (1993), log-odds ratios can also be interpreted in terms of heterogeneity of brand choice probabilities in a population of zero-order consumers, where dissimilar brands (large log-odds ratios) implies heterogeneous choice probabilities and thus the opportunity to segment a market based upon brand choice. Log-linear trees thus allow structural features underlying a brand switching matrix to be graphically represented, and provide tests of significance for the graphical representation.