Journal of Multivariate Analysis 126 (2014) 75–85

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

A class of smooth models satisfying marginal and context specific conditional independencies

R. Colombi [a], A. Forcina [b,*]

[a] Statistical Laboratory, University of Bergamo, Italy

[b] Dipartimento di Economia, Finanza e Statistica, University of Perugia, Italy

ARTICLE INFO

Article history:
Received 6 November 2012
Available online 20 January 2014

AMS subject classifications:
37C05
46N30
65H12

Keywords:
Categorical data
Marginal log-linear parameterizations
Smooth parameterizations

ABSTRACT

We study a class of conditional independence models for discrete data with the property that one or more log-linear interactions are defined within two different marginal distributions and then constrained to 0; all the conditional independence models which are known to be non-smooth belong to this class. We introduce a new marginal log-linear parameterization and show that smoothness may be restored by restricting one or more independence statements to hold conditionally to a restricted subset of the configurations of the conditioning variables. Our results are based on a specific reconstruction algorithm from log-linear parameters to probabilities and fixed point theory. Several examples are examined and a general rule for determining the implied conditional independence restrictions is outlined.

1. Introduction

Conditional independence models for discrete data are determined by a set of constraints on log-linear interactions defined within different marginal distributions of a contingency table. The family of hierarchical and complete marginal log-linear parameterizations (HCMP for short) introduced by Bergsma and Rudas [4] provides a general framework for combining log-linear constraints defined on a collection of marginal distributions into an overall joint distribution. Methods for determining whether and how a conditional independence model may be translated into a HCMP have been studied by Rudas et al. [14] and Forcina et al. [10] among others; the fact that a HCMP exists, is a sufficient condition for the model to be smooth.

It has been shown [4, Theorem 3] that, when the same interaction is defined in two different marginals, the Jacobian of the mapping from log-linear parameters to probabilities is singular for the uniform distribution. Though, formally, this does not imply that the model itself has singularities, all known models with singularities correspond to cases where one or more interactions are constrained more than once. In this paper we study the class of conditional independence models where the same interaction is constrained in two or more marginal distributions and we show, essentially, that any such model is non-smooth but can be turned into a smooth model by assuming that the conditional independencies hold only for specific combinations of the categories of the conditioning variables.

Following Bergsma and Rudas [4], we may assume, without loss of generality, that the marginal distributions of interest have been arranged in a non decreasing order and that they will be reconstructed one at a time starting from the smallest. Because the full joint distribution is simply the last marginal in this list, we need only to consider how to determine a given marginal distribution when one or more log-linear interactions to be constrained have already been defined

* Corresponding author.
E-mail address: forcina@stat.unipg.it (A. Forcina).

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2014.01.001

and/or constrained in a previous marginal. A useful tool for reconstructing an ordered set of marginal distributions is the mixed parameterization (e.g., [2]) by which we may combine the probabilities from previous marginals with the log-linear interactions defined in the marginal distribution under consideration. Because the mapping produced by the mixed parameterization is one to one and smooth, if there is a reconstruction algorithm based on the mixed parameterization which converges everywhere, any model defined by linear constraints on the corresponding log-linear parameters must be smooth. By using results from the theory of fixed point algorithms, we study a new reconstruction algorithm that allows certain log-linear interactions to be redefined and show that this may either converge, and thus the model is smooth, remain at the starting point irrespective of the starting value, implying that the resulting distribution is not uniquely determined by the log-linear parameters or, simply not converge. A formal proof of these properties is derived under complete independence; our results provide some evidence to support the conjecture that there exist independence models defined by constraining the same interactions in two or more marginals which are smooth everywhere in the parameter space.

The results derived in this paper determine the set of interaction parameters which may be redefined and the corresponding set of interactions which should be omitted as a replacement: this second set of interactions has the property that, when one or more of its elements are missing, and thus unconstrained, the conditional independence of interest holds only on a subset of the configuration of the conditioning variables. Log-linear models which allow context specific conditional independencies have been studied in detail by Hojsgaard [11] who also derives a Markov property for undirected graphs involving context specific conditional independencies. A special case of the results derived here was considered by Roverato et al. [13].

In Section 2, we introduce the basic notations, define marginal log-linear interactions and review the properties of the mixed parameterization. In Section 3, after presenting a set of motivating examples, we introduce a new algorithm for reconstructing a marginal distribution when interactions defined in previous marginals have to be constrained again and we analyze its convergence properties. In Section 4 we study the consequences on the original conditional independence statements of omitting constraints on a specific subset of higher order interactions and show that this results in context specific restrictions.

2. Notations and preliminary results

We study the joint distribution of d discrete random variables where $X_j, j = 1, \ldots, d$, takes values in $(0, \ldots, r_j)$. For conciseness, we denote variables by their indices and use capitals to denote non-empty subsets of $V = \{1, \ldots, d\}$; such subsets will determine the variables involved either in a marginal distribution or in an interaction term. The collection of all non-empty subsets of a set $M \subseteq V$ will be denoted by $\mathcal{P}(M)$. In the following we write $i_1 i_2 \ldots i_k$ as a shorthand notation for $\{i_1, i_2, \ldots, i_k\}$. For a given $M \subseteq V$, the marginal distribution in M is determined by the cell probabilities $p_M(\boldsymbol{x}_M) = P(X_j = x_j, \forall j \in M)$. We introduce a shorthand notation that allows to specify the values of selected subsets of the arguments in a marginal probability and on the log-linear interactions to be defined below. Let $J \subset I \subset M$, then $p_M(\boldsymbol{x}_J, \boldsymbol{x}_{I\setminus J}, \boldsymbol{x}_{M\setminus I})$ denotes the marginal probability where $\boldsymbol{x}_J$ is the value of X_h, $h \in J$, $\boldsymbol{x}_{I\setminus J}$ the value of X_h, $h \in I\setminus J$ and $\boldsymbol{x}_{M\setminus I}$ the values of X_h, $h \in M\setminus I$. We will also write $\boldsymbol{0}_{I\setminus J}$ to state that $X_h = 0$, $\forall h \in I\setminus J$.

2.1. Marginal and conditional log-linear interactions

Though there are many different ways of coding marginal log-linear parameters, parameters defined by different codings are linearly related; thus there is no loss of generality in using the *reference category* coding, where comparisons are with respect to the category taken as reference; here we use the category denoted with "0" as reference.

Definition 1. A reference category log-linear interaction I within M is defined by the following expression

$$\eta_{I;M}(\boldsymbol{x}_I \mid \boldsymbol{x}_{M\setminus I}) = \sum_{J\subseteq I} (-1)^{|I\setminus J|} \log p_M(\boldsymbol{x}_J, \boldsymbol{0}_{I\setminus J}, \boldsymbol{x}_{M\setminus I}), \tag{1}$$

where, $\forall i \in I, x_i > 0$.

Example 1. The logit of X_i at x_i computed within M is

$$\eta_{i;M}(x_i \mid \boldsymbol{x}_{M\setminus i}) = \log p_M(x_i, \boldsymbol{x}_{M\setminus i}) - \log p_M(0_i, \boldsymbol{x}_{M\setminus i}), \quad x_i > 0$$

and the log-odds ratio for $X_i = x_i, X_j = x_j$ is

$$\eta_{H;M}(x_i, x_j \mid \boldsymbol{x}_{M\setminus H}) = \log p_M(x_i, x_j, \boldsymbol{x}_{M\setminus H}) - \log p_M(0_i, x_j, \boldsymbol{x}_{M\setminus H}) - \log p_M(x_i, 0_j, \boldsymbol{x}_{M\setminus H}) + \log p_M(0_i, 0_j, \boldsymbol{x}_{M\setminus H})$$

where $H = i \cup j$.

It may be easily verified that, given $h \in M\setminus I$ and $H = I \cup h$, (1) implies the following recursive relation

$$\eta_{H;M}(\boldsymbol{x}_H \mid \boldsymbol{x}_{M\setminus H}) = \eta_{I;M}(\boldsymbol{x}_I \mid x_h, \boldsymbol{x}_{M\setminus H}) - \eta_{I;M}(\boldsymbol{x}_I \mid 0_h, \boldsymbol{x}_{M\setminus H}), \tag{2}$$