Computational Statistics and Data Analysis 68 (2013) 296–310

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Dealing with multiple local modalities in latent class profile analysis

Hsiu-Ching Chang[a], Hwan Chung[b,*]

[a] Department of Clinical Epidemiology and Biostatistics, Blue Cross Blue Shield of Michigan, 500 Rencen, Detroit, MI 48226, USA

[b] Department of Statistics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 136-701, Republic of Korea

ARTICLE INFO

Article history:
Received 18 March 2011
Received in revised form 8 July 2013
Accepted 8 July 2013
Available online 16 July 2013

Keywords:
Deterministic annealing
Latent stage-sequential process
Markov chain Monte Carlo
Maximum likelihood
Recursive formula

ABSTRACT

Parameters for latent class profile analysis (LCPA) are easily estimated by maximum likelihood via the EM algorithm or Bayesian method via Markov chain Monte Carlo. However, the local maximum problem is a long-standing issue in any hill-climbing optimization technique for the LCPA model. To deal with multiple local modalities, two probabilistic optimization techniques using the deterministic annealing framework are proposed. The deterministic annealing approaches are implemented with an efficient recursive formula in the step for the parameter update. The proposed methods are applied to the data from the National Longitudinal Survey of Youth 1997 (NLSY97), a survey that explores the transition from school to work and from adolescence to adulthood in the United States.

1. Introduction

Many longitudinal studies on behavioral research are turning to the idea of stage-sequential process. A common theme of stage-sequential process is that, at any moment, individuals can be sorted into distinct qualitative classes, and they can change their class membership over time. The latent class analysis (LCA) is perhaps the most straightforward mixture model now being used to identify mutually exclusive groups of individuals based on their responses to manifest items. A number of methods for analyzing the stage-sequential processes have been derived from the family of LCA. These new methods include the hidden Markov model (HMM) (Baum et al., 1970; Rabiner, 1989; Welch, 2003; Cappé et al., 2005; Zucchini and MacDonald, 2009). In HMM, the measurement model at each time point is specified with an LCA, and the stage-sequential development is summarized in terms of transition probabilities among latent classes over time. A transition across classes is typically represented with a first-order Markov chain, on the assumption that class membership at time t depends only on class membership at time $t-1$. When there are data at three or more time points, however, it may be more useful to characterize the stage-sequential development by a small number of common patterns of the sequences of class membership than to summarize it by transition probabilities with the assumption of a first-order Markov chain. Recently, Chung et al. (2011) proposed the latent class profile analysis (LCPA), which provides an LCA principle for the systematic identification of sequential patterns of latent class membership over time. We may view the LCPA as a standard LCA, where the identification process can be divided into two steps. In the first step, LCPA identifies *latent classes* in which individuals show similar responses to manifest items at each measurement occasion. In the second step, LCPA examines individuals' latent class membership across all time points to classify the study population into two or more *latent profiles* based on their class sequencing. In other words, the

* Corresponding author. Tel.: +82 2 3290 2246; fax: +1 82 2 924 9895.
E-mail address: hwanch@korea.ac.kr (H. Chung).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.07.016

measurement model at each time point is specified in terms of latent classes, and stage-sequential development of class memberships is explained by a small number of latent profiles. Note that the standard LCA can be applied to this two-step procedure.

Like any other finite mixture models, the first and most crucial step in LCPA is to choose an appropriate number of classes and profiles. Recently, Chung and Chang (2012) proposed two Bayesian approaches, reversible jump Markov chain Monte Carlo (MCMC) and Dirichlet process, in order to select the number of classes and profiles for the LCPA model. Once a model is selected, the parameters of the LCPA model are needed to be estimated. The expectation–maximization (EM) algorithm (Dempster et al., 1977) and the data augmentation using the MCMC (Tanner and Wong, 1987) can be easily implemented to draw statistical inferences of the parameters for the LCPA model. However, as the number of measurement occasions increases in the LCPA model, the computational burden of EM or MCMC will become exponentially intensive. On the contrary, if one adapts a recursive scheme in the update step, calculations for parameter estimates can be simplified for the LCPA model. In this paper, we propose to formulate each update step for model parameters with a recursive formula, which are directly analogous to the forward–backward algorithm (Chib, 1996; MacKay, 1997).

The parameter estimation for the LCPA model benefits from the recursive formula; yet, the recursive algorithm still requires careful examination for the existence of multiple local modes of the objective function. There is a plethora of literature aiming at finding a numerically robust parameter estimation tool to deal with local maximum problems (Biernacki et al., 2003; Reddy and Rajaratnam, 2010). For example, Reddy and Rajaratnam (2010) convolute the objective function by kernel functions to flatten the surface and reduce the number of local modes. They demonstrate that the optimal solutions would be reached effectively by adjusting the smoothing factor at each iteration. Instead, we implement deterministic annealing EM (DAEM) (Ueda and Nakano, 1998) and deterministic annealing variant of variational Bayes (DAVB) (Katahiral et al., 2008) in order to find the parameter estimates on the global mode of the objective function. Both DAEM and DAVB, formulated by the recursive formula, are based on the deterministic annealing approach, in which ω is included as an annealing parameter to control the annealing rate. By adjusting the value of ω, the annealing process tracks multiple local modes and identifies the globalized optimum as a result.

For their analysis of a data set from the National Longitudinal Survey of Youth 1997 (NLSY97), Chung et al. (2011) used 100 different sets of initial parameters in order to avoid local optimum entrapment in the LCPA model. They selected a set of estimates providing the highest log-likelihood value among those from 100 different initializations. We conduct DAEM and DAVB with the predetermined annealing schedule and compare their estimates with those provided by the standard EM algorithm with 100 sets of starting values.

The organization of the rest is as follows. We introduce the LCPA model in Section 2. In Section 3, we provide two deterministic annealing approaches, DAEM and DAVB, both stated in the recursive way for parameter estimation. Section 4 carries out simulation studies in order to evaluate the performance of the proposed estimation methods. In Section 5, we apply the standard EM, DAEM, and DAVB to alcohol drinking items drawn from the NLSY97 to assess each algorithm's performance. In Section 6, we discuss the advantages and limitations of the deterministic annealing approaches and finally, conclude the paper.

2. Latent class profile analysis

2.1. Motivating example: adolescents' drinking behaviors

We begin this section to motivate the LCPA model with an application given in Lanza and Collins (2006), where the standard LCA with repeated measures was used to identify common pathways of heavy drinking in people between ages 18 to 30 years. They used a single binary indicator of heavy drinking at six different time points, allowing a straightforward use of LCA to produce a set of latent classes. Each of the latent classes identified by their standard LCA can be characterized by a particular sequence of drinking behaviors over time. However, in the situation where drinking behavior was measured by more than one item per time, the interpretation of the resulting classes from the standard LCA is much more complex because these classes contain information on two types of latent variables (i.e., latent class and latent profile that were defined in Section 1 simultaneously. Although imposing a set of parameter restrictions may allow for confirmatory tests regarding stage-sequential process using the standard LCA, it is difficult to impose plausible restrictions a priori in the absence of well-established substantive theory.

We believe that the LCPA model can be illuminated with an application given in Section 5, where an LCPA has been applied to identify the subtypes of sequential patterns of early-onset drinking behaviors. To investigate adolescents' drinking behaviors, LCPA identifies discrete subgroups of early-onset drinkers who have similar drinking behaviors (i.e., similar manifest responses to three drinking items) at each time point. We will refer to the subgroups identified in this step as *classes*. In addition, LCPA examines the early-onset drinkers' class memberships over the entire set of time points and classifies them into two or more subgroups based on their class sequencing. We refer to the subgroups identified in this step as *profiles*. By applying an LCPA to the longitudinal study of early-onset adolescent drinking, all drinkers in a class at a certain point in time are expected to be homogeneous in terms of their drinking behaviors, and those individuals in a given class profile will have similar sequential patterns of class membership over time (Chung et al., 2011).