Bayesian Hierarchical Multivariate Formulation with Factor Analysis for Nested Ordinal Data

Terrance D. Savitsky*

September 29, 2012

1 Abstract Body

1.1 Background, Context:

Empirical results from value-added modeling demonstrate high variations among teacher effects derived from student test performance (Aaronson et al. 2007, Goldhaber et al. 2010, Hanushek 1992, Rivkin et al. 2005), prompting policy focus on improving teaching and teacher evaluation. A variety of structured protocols have recently been developed to formalize the assessment of teacher classroom practice and these are rapidly making their way into updated state and school district-mandated teacher evaluation systems, as well as being used by researchers studying teaching. Our analysis employs the “Classroom Assessment and Scoring System” for the secondary education level (CLASS-S) of Hamre, Pianta, Burchinal, Field, LoCasale-Crouch, Downer, Howes, LoParo & Scott-Little (2012). CLASS-S identifies 10 dimensions, grouped into 3 domains, derived from extensive theoretical and empirical work in pre-school and early elementary school classrooms, as the key elements of teaching to support learning.

Evaluations of teachers (by practitioners or researchers) typically use data from multiple observations to draw conclusions about teacher classroom performance over a time span, such as a year instruction. The intent of utilizing multiple observations through an instructional year is to allow the removal of nuisance variation related to the idiosyncratic characteristics of any given measurement event. The unit of an observation score is one of multiple ratings by separate raters of all the segments or disjoint intervals of one or more measurement event class periods (which we label as sessions) from one or more sections of students taught by a given teacher. A scoring creates multivariate ordered score vectors (each with length equal to the number of protocol dimensions, e.g. 10 for CLASS-S).

Research on classroom observation measures is in an emerging phase such that there is little consistent information on the nature of the dependence structure among the CLASS-S dimensions at each level in the nested hierarchy.

*RAND Corporation, 1776 Main Street, Box 2138, Santa Monica, CA 90401-2138 USA
1.2 Purpose, Objective, Research Question, Focus of Study:

Our inferential task seeks to discover any underlying set of less than 10 independent factors that span the enumerated dimensions of CLASS-S. Our focus is at the teacher level, but we must simultaneously capture dependence structures both across the multiple levels of the hierarchy and among the CLASS-S dimensions to extract an unbiased teacher-level correlation matrix along with a sparse underlying factor structure. In other words, we expect each level of the hierarchy to potentially express its own multivariate dependence structure for which we must account.

1.3 Population, Participants, Subjects:

We evaluate our methods on both simulated data and an application acquired from classroom observations on 82 middle and high school algebra teachers based in a large urban school district from the research study data set of Casabianca et al. (2012) that we label in the sequel as “the algebra teacher data”. Our first simulated data set generates multivariate ordered observations from a latent continuous response whose mean is composed with multivariate random effects at each level of the hierarchy. We use these data to compare our formulations with typically-employed alternatives. A second simulated data set assumes that raters cluster in their perspectives or beliefs about teacher performance. We draw 30 raters from a cluster of 3 to generate multivariate rater and session-rater random effects. This clustering of rater perspectives produces ordered observations characterized by mild multi-modality on each of the 10 CLASS-S dimensions.

1.4 Significance, Novelty of Study:

Bayesian ordinal data formulations for an outcome specified over multiple dimensions often assume independence among the dimensions. These models employ either a latent response or set of cumulative probabilities over multiple dimensions specified as a function of a univariate intrinsic item (e.g. a teacher) trait (with associated cutpoints, after removing variation through inclusion of fixed effects) (Qiu et al. 2002, Johnson & Albert 1999). These models collapse across the multiple dimensions for each response to a univariate trait. A fundamental innovation of this paper is the development of multivariate, hierarchical models for ordered outcomes that allows for a separate decomposition and discovery of dependence structures at each level of a hierarchy. Our framework permits an unbiased extraction of the teacher-level multivariate dependence structure.

1.5 Statistical, Measurement, or Econometric Model:

We introduce a parametric model in which the observed ordered outcomes are generated by latent multivariate responses formulated as mixtures of multivariate Gaussian variables from each level of the hierarchy. We develop an approach to scale-identify the resulting model, which is particularly challenging with multivariate random effects.
Next, we expand the parametric approach by generalizing the implied probit link function to a semi-parametric mixture formulation based on Kottas et al. (2005). We expand their construction to fit within a nested hierarchical framework and re-purpose the inferential focus from the estimation of contingency cell probabilities to multivariate correlation matrices at multiple levels. A unique feature of our semi-parametric formulation is that it permits the joint specification of known nested structure with an unstructured term to capture unspecified dependence in the observations to yield unbiased inferences about teacher-level dependence among the dimension scores of an observation protocol.

The models allow for a sparse generating structure for the nested covariance matrices through Bayesian factor analytic models. We employ a parameter expansion approach from Ghosh & Dunson (2009) to improve posterior sampling mixing performance. Factor loadings are sampled without sign restriction under this approach to facilitate mixing, though the resulting sign invariance in the likelihood requires a subsequent transformation step. We correct an error on assignment of signs in Ghosh & Dunson (2009) by employing a post-processing relabeling step using a unimodal loss function from Curtis & Erosheva (2011) to select signs for the loadings based on the work of Stephens (2000); only, Curtis & Erosheva (2011) limit their exposition and application to confirmatory factor analytic approaches. We generalize the approach and application to exploratory factor analytic modeling under which there is no assumed a priori knowledge for the structure of the factor loadings matrix.

1.6 Usefulness, Applicability of Method:

Our first simulation demonstrates the superior performance associated to both modeling the data as hierarchical and ordinal for estimation of teacher level correlations. We employ two comparator methods; the first method, labeled “averaged”, eschews a model-based approach and simply averages the observations over the nested hierarchy to the teacher level and composes the sample correlations at the teacher level. The second method treats the ordered outcomes as continuous in a model that accounts for the multivariate, nested structure. Figure 1 compares the two methods to our Bayesian ordinal parametric formulation to reveal that both comparator methods produce overly smoothed or attenuated correlations such that our method more accurately reproduces the true correlation structure.

Our second simulation compares the performance of our ordinal multivariate parametric and semi-parametric formulations under rater-induced multi-modality. Figure 2 presents histogram distributions of our simulated data for each CLASS dimension and demonstrates how a clustering of rater perspectives on teaching performances produces mild multi-modality. The teacher-level covariance matrix is constructed from a sparse 3 factor design under which the first 3 measures are designed to load onto the first factor, the next 3 onto the second factor and the last 4 onto the third in an orthogonal loadings construction. Figure 3 compares the modeled loading matrices for the parametric (left-hand panel) and semi-parametric (right-hand panel) formulations, where the rows in each of the 2 panels represent each of the CLASS-S measures and the columns represent the factors. The loadings are normalized to express the
percent of variance explained at each factor-measure intersection. We see that the true signal is more strongly expressed for the semi-parametric model under rater clustering, though the parametric model (which is not designed to handle clustering) still performs reasonably well.

Lastly, we apply our model to the algebra teacher data set. Figure 4 highlights that we recover a single teacher-level factor that strongly explains essentially all of the variation among the 10 CLASS-S dimensions. Such a result indicates that among the 82 teachers empaneled for the algebra teacher data, some generally score high across all 10 dimensions and others generally score lower all of the dimensions, but there are no patterns of performance among subsets of the dimensions. A single measure describes the variation in teaching among these teachers. The result is inconsistent with the theoretical model used in developing CLASS-S and with the results from other analyses that do not distinguish the teacher-level effects from other levels of the hierarchy or model the ordinality of the scores (Hamre, Hafen, Pianta, Bell, Gitomer & McCaffrey 2012).

1.7 Conclusion:

Our hierarchical, multivariate formulation with factor analysis provides the first comprehensive approach for the unbiased estimation of the teacher-level dependence structure that intends to uncover the bases for excellent teaching. We have generalized univariate formulations for the analysis of ordinal data to the case of multivariate nested data that facilitates the extraction of a set of multivariate random effects at each level in the hierarchy. We demonstrate how our hierarchal nested data model paired with a factor analysis formulation simultaneously extracts the correlation structure for teacher effects and permits our assessment for an underlying sparse generating process.

Our results on simulated data demonstrate the importance of both modeling the observed data as ordered and accounting for multivariate dependence structures at all levels of the hierarchy expressed in the study design in order to provide an unbiased teacher-level result. The application to the algebra teacher data set reveals the possibility for rich inference possible by accounting for the dependence structures over multivariate dimensions of any protocol employed for the measurement of teacher classroom performance.
1.8 Appendix B: Tables and Figures

Figure 1: Comparisons of 95% credible intervals for teacher level correlations under 3 different formulations - 1. Ordinal Multivariate; 2. Continuous Multivariate; 3. Sample averaging up to teacher level.
Figure 2: Histograms of observed ordinal simulated data responses, faceted by each of 10 dimensions.
Figure 3: Posterior mean teacher level factor loadings, by cell, normalized to amount of total variation explained for each (row) dimension.
Figure 4: Posterior mean teacher level factor loadings, by cell, normalized to amount of total variation explained for each (row) dimension for algebra teacher data.

References


Goldhaber, D., & Hansen, M. (2010), ‘Is it just a bad class? assessing the stability of measured teacher performance.’.


