Latent Markov Model

Jeroen K. Vermunt

The latent Markov model (LMM) can either be seen as an extension of the LATENT CLASS MODEL for the analysis of longitudinal data or as an extension of the discrete-time MARKOV CHAIN MODEL for dealing with measurement error in the observed variable of interest. It was introduced in 1955 by Wiggins and also referred to as latent transition or hidden Markov model. The LMM used to separate true systematic change from spurious change resulting from measurement error and other types of randomness in the behavior of individuals.

Suppose a single categorical variable of interest is measured at T occasions, and that Y_t denote the response at occasion t, $1 \leq t \leq T$. This could, for example, be a respondent's party preference measured at 6-month intervals. Let D denote the number of levels of each Y_t, and y_t a particular level, $1 \leq y_t \leq D$. Let X_t denote an occasion-specific latent variable, C number of categories of X_t, and x_t a particular LC class at occasion t, $1 \leq x_t \leq C$. The corresponding LMM has the form

$$
\begin{aligned}
P(\mathbf{Y} = \mathbf{y}) &= \sum_{\mathbf{x}} P(X_1 = x_1) \prod_{t=2}^{T} P(X_t = x_t | X_{t-1} = x_{t-1}) \\
&\quad \prod_{t=1}^{T} P(Y_t = y_t | X_t = x_t).
\end{aligned}
$$

The two basic assumptions of this model are that the transition structure for the latent variables has the form of a first-order Markov chain and that each occasion-specific observed variable depends only on the corresponding latent variable. For identification and simplicity of the results, it is typically assumed that the error component is time-homogeneous:

$$
P(Y_t = y_t | X_t = x_t) = P(Y_{t-1} = y_t | X_{t-1} = x_t),
$$

for $2 \leq t \leq T$. If no further constraints are imposed, one needs at least 3 time points to identify the LMM. Typical constraints on the latent transition probabilities are time-homogeneity and zero restrictions.

The enormous effect of measurement error in the study of change can be illustrated with a hypothetical example with $T = 3$ and $C = D = 2$. To illustrate this, suppose $P(X_1 = 1) = .80$, $P(X_t = 2|X_{t-1} = 1) = P(X_t =$

$2|X_{t-1} = 1) = .10$, and $P(Y_t = 1|X_t = 1) = P(Y_t = 2|X_t = 2) = .20$. If we estimate a stationary manifest first-order Markov model for the hypothetical table, we find .68 in state one at time point one, and transition probabilities of .29 and .48 out of the two states. These are typical biases encountered when not taking measurement error into account: the size of the smaller group is overestimated, the amount of change is overestimated, and there seems to more change in the small than in the large group.

It is straightforward to extend the above single-indicator LMM to multiple indicators. Another natural extension is the introduction of covariates or grouping variables explaining individual differences in the initial state and the transition probabilities. The independent classification error (ICE) assumption can be relaxed by including direct effects between indicators at different occasions. Furthermore, mixed variants of the LMM have been proposed, such as models with MOVER-STAYER structures. In the social sciences, LMM are conceived as tools for categorical data analysis. However, as in standard LATENT CLASS ANALYSIS, these models can be extended easily to other scale types.

The PANMARK program can used to estimated the more standard LMMs. The LEM program can also deal with more extended models, such as models containing covariates and direct effects between indicators.

References

Langeheine, R., and Van de Pol, F. (1994). Discrete-time mixed Markov latent class models. A. Dale and R.B. Davies (eds.), *Analyzing Social and Political Change: a Casebook of Methods*, 171-197. London: Sage Publications.

MacDonald, I.L., and Zucchini, W. (1997). *Hidden Markov Models and Other Types of Models for Discrete-valued Time Series*. London: Chapman & Hall.

Vermunt, J.K. (1997). *Log-linear Models for Event Histories*. Thousand Oakes: Sage Publications.

Wiggins, L.M. (1973). *Panel Analysis*. Amsterdam: Elsevier.