Journal of the Korean Statistical Society 40 (2011) 469–485

Contents lists available at SciVerse ScienceDirect

Journal of the Korean Statistical Society

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

Modeling rating transitions

Rafael Weißbach [a,*], Thomas Mollenhauer [b]

[a] University of Rostock, Germany

[b] NORD/LB, Credit Risk Control, Hannover, Germany

ARTICLE INFO

Article history:
Received 4 October 2010
Accepted 3 May 2011
Available online 24 May 2011

AMS 2000 subject classifications:
60J28
62M02
62P05

Keywords:
Rating
Metric
Multiple Markov process
Counting process
Likelihood ratio

ABSTRACT

The time-continuous discrete-state Markov process is a model for rating transitions. One parameter, namely the intensity to migrate to an adjacent rating state, implies an ordinal rating to have an intuitive metric. State-specific intensities generalize such state-stationarity. Observing Markov processes from a multiplicative intensity model, the maximum likelihood parameter estimators for both models can be studied with the score statistic, written as a martingale transform of the processes that count transitions between the rating states. A Taylor expansion reveals consistency and asymptotic normality of the parameter estimates, resulting in a χ^2-distributed likelihood ratio of state-stationarity against the state-specific model. This extends to time-stationarity. Simulations contrast the asymptotic results with finite samples. An application to a sufficiently large set of credit rating histories shows that the one-parameter model can be a good starting point.

1. Introduction

The homogeneous Markov process, with stationary transition intensities, remains the starting point for rating transition modeling (Bluhm, Overbeck, & Wagner, 2002, pg. 197ff). Several defects of the model, such as instationarity, non-Markovian behavior or intra-temporal dependence have been investigated (Altman & Kao, 1992; Bangia, Diebold, Kronimus, Schlagen, & Schuermann, 2002; Frydman & Schuerman, 2008; Koopman & Lucas, 2008; Lando & Skødeberg, 2002; Kiefer & Larson, 2007; Weißbach & Walter, 2010). Our general statistical objective is a parsimonious model, and Forest, Belki, and Suchower (1998) formulate a one-parameter model for ratings, not, however, originating from the Markov process model. Here, we claim three properties that enable the formulation of a one-parameter Markov process model. First, the Merton model (Merton, 1974) for an asset value suggests that a firm can only migrate from one rating state to an adjacent rating state, that is, up or down. All other transition intensities must be zero and observations of multiple class transitions are attributable to discontinuous observation and disregarded here as missing data. Second, a rating should be constructed so as to be cardinal, and not only ordinal. If changing rating classes does not depend on the specific state, i.e. is state-stationary, we will see that the rating can be equipped with a simple metric. Third, transition intensities should be time-stationary.

We assume the first claim and take it into account in the data analysis. The second restriction, namely whether rating class changes are class-specific, is our primary research question. We study a likelihood ratio test on the null hypothesis of cardinality. The formalization of the null hypothesis is a constant transition intensity for all rating classes and the alternative

* Corresponding address: Faculty for Economic and Social Sciences, University of Rostock, 18051 Rostock, Germany. Tel.: +49 381 4984428; fax: +49 381 4984401.

E-mail address: rafael.weissbach@uni-rostock.de (R. Weißbach).

URL: http://www.wiwi.uni-rostock.de/vwl/statistik/ (R. Weißbach).

1226-3192/$ – see front matter
doi:10.1016/j.jkss.2011.05.001

hypothesis, of an only ordinal rating, is formalized by letting each rating class have its two specific transition intensities in the directions of upgrade and downgrade. The third property will be studied briefly.

Maximum likelihood estimation for the generator of the homogeneous Markov process dates back to Albert (1962). We study the consistency and asymptotic normality of the estimator for the state-specific intensity and the state-stationary intensity. The results originate from the representation of the score statistic as martingale transforms that arise from the transition counts between rating states. The martingale limit theorem by Rebolledo (1980) suggests studying the predictable covariation process with inequalities by Lenglart and Gill. The time-stationarity can be studied by generalizing to piecewise time-stationarity. Through an argument relating to the profile likelihood, the likelihood ratio test statistics that compare on the one hand state-stationary versus state-specific and on the other time-stationary versus piecewise stationary are both asymptotically χ^2-distributed.

Our application is credit risk, in particular, the rating transition intensities in an internal rating system, loosely related to the expert-rating discussed by Kiefer (2010). We show that our data may not be modeled significantly by one parameter, even if time-stationarity is imposed by transformation of the time. However, the model appears to be close to reality, simulation studies foster such an impression.

2. Models

Consider the time-continuous discrete-state Markov processes $\mathbf{X} = \{X_t, t \in [0, T]\}$ defined on a probability space $(\Omega, \mathfrak{F}, P)$. The ordered states $1, \ldots, k$, e.g. rating classes, end in an absorbing state k (e.g. bankruptcy). We denote X_t as the state of an asset at time t, after a certain origin. Denote by $m_h(t) = P(X_t = h)$ the unconditional probability of state h at time t. The data are transition histories $\mathbf{X}_i = \{X_t^i, t \in [0, T]\}$ for each of the $i = 1, \ldots, n$ assets within a sample.

2.1. State-stationarity

The homogeneous, i.e. time-stationary, process is determined by the infinitesimal generator of the process $\mathbf{Q} = (q_{hj})_{h,j=1,\ldots,k}$ with transition intensities

$$q_{hj} = \lim_{u \to 0^+} \frac{P(X_u = j \mid X_0 = h)}{u}.$$

Note that $q_{hh} = -\sum_{j=1, j \neq h}^{k} q_{hj}$ and $q_{kj} = 0$. If transition to any other than the adjacent class is impossible, $\mathbf{Q}$ is determined by elements on the first off-diagonals. It is useful to collect the indices for all non-zero intensities in set $\mathcal{I}_1 = \{(h, j) : h = 1, \ldots, k-1; j = 1, \ldots, k; |h - j| = 1\}$ and to define set $\mathcal{I}_2 = \mathcal{I}_1 \setminus \{(1, 2)\}$.

Definition 1. Let the intensities on $[0, T]$ be

$$q_{hj} = \begin{cases} q & \text{if } (h, j) = (1, 2) \\ q + \gamma_{hj} & \text{if } (h, j) \in \mathcal{I}_2 \end{cases}$$

with $q > 0$ and $\gamma_{hj} \in (-q, \infty)$.

We denote by *state-stationarity* the one-parameter case with restriction $\gamma_{hj} = 0$ for all $(h, j) \in \mathcal{I}_2$. By defining the mapping $(h, j) \mapsto |h - j|q$ the set of states is equipped with a metric on $\mathcal{I}_1$. In the unrestricted case with $\gamma_{hj} \neq 0$, the same mapping is not a metric.

We have no intention to analyze on asset level so that, compared to the analysis of all transition histories $\mathbf{X}_1, \ldots, \mathbf{X}_n$, there is no loss of information when using the vector of initial ratings $X_0^1, \ldots, X_0^n$ together with the processes

$$N_{hj}(t) = \#\{s \in [0, t], i = 1, \ldots, n | X_{s-}^i = h, X_s^i = j\}, \quad t \in [0, T], (h, j) \in \mathcal{I}_1$$

counting the number of transitions from state h to j until time t in the entire sample. Additionally, let the processes $Y_h(t)$ denote the number of assets in state h at time t. For large samples, this constitutes a clear reduction in the number of random processes. We impose two additional assumptions.

(A1) For fixed t and $n \to \infty$ in probability ($\xrightarrow{P}$)

$$\frac{Y_h(t)}{n} \xrightarrow{P} m_h(t).$$

(A2) The counting processes N_{hj} must follow a multiplicative intensity model, i.e. with collection of q and γ_{hj} in vectors $\boldsymbol{\gamma} := (\gamma_{21}, \gamma_{23}, \ldots, \gamma_{k-1,k})' \in \mathbb{R}^{2k-4}$ and $\boldsymbol{\theta} := (q, \boldsymbol{\gamma}')' \in \mathbb{R}^{2k-3}$ they have the intensity process

$$\lambda_{hj}(t; \boldsymbol{\theta}) = Y_h(t) q_{hj}, \quad (h, j) \in \mathcal{I}_1.$$

Due to the law of large numbers, assumption (A1) is fulfilled if the Markov processes are independent. Independence is also a sufficient condition for (A2).