Computational Statistics and Data Analysis 68 (2013) 388–398

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Bayesian dynamic probit models for the analysis of longitudinal data

Refik Soyer [a], Minje Sung [b,*]

[a] Department of Decision Sciences, The George Washington University, Washington, DC, USA

[b] School of Business Administration, Ajou University, Suwon, Republic of Korea

HIGHLIGHTS

- A dynamic probit model with a first order Markov process was developed.
- Gibbs sampler using data augmentation approach and forward filtering backward sampling algorithm were presented.
- The discussion was extended to propose models to generalized models including Student t-link function.
- Model fit was compared between static and dynamic probit models.

ARTICLE INFO

Article history:
Received 18 June 2012
Received in revised form 28 January 2013
Accepted 26 April 2013
Available online 4 May 2013

Keywords:
Markov models
Bayesian inference
Longitudinal data
Dynamic linear models
Model selection

ABSTRACT

The authors consider a dynamic probit model where the coefficients follow a first-order Markov process. An exact Gibbs sampler for Bayesian analysis is presented for the model using the data augmentation approach and the forward filtering backward sampling algorithm for dynamic linear models. The authors discuss how our approach can be used for dynamic probit models as well as its generalizations including Markov regressions and models with Student link functions. An approach is presented to compare static and dynamic probit models as well as for Markov order selection in these classes of dynamic models. The developed approach is implemented to some actual data.

1. Introduction and overview

Time varying coefficient models for categorical longitudinal data have been considered by authors such as Carlin and Polson (1992), Shephard and Pitt (1997), Gamerman (1998), Kauermann (2000), and more recently by Fruhwirth-Schnatter and Fruhwirth (2007). Most of the previous works have considered logit-type state-space models. As noted by Fruhwirth-Schnatter and Fruhwirth (2007), Markov chain Monte Carlo (MCMC) approaches proposed by Shephard and Pitt (1997) and Gamerman (1998) for the analysis of these models are based on the Metropolis–Hastings algorithm which requires specification of a proposal density in high dimensions. To alleviate this, the authors proposed a data augmentation based MCMC method for analysis of dynamic logit models. A simple version of a dynamic probit model has been considered by Andrieu and Daucet (2002) where the authors used particle filtering for Bayesian analysis.

In what follows, we consider probit-type state-space models and develop an exact Gibbs sampler for Bayesian analysis of this class of models. Our approach is an extension of the data augmentation approach of Albert and Chib (1993) to dynamic

* Correspondence to: School of Business, Ajou University, San 5 Woncheon-Dong Yeongtong-Gu, Suwon, Kyunggi-Do 443-749, Republic of Korea. Tel.: +82 10 6414 1660; fax: +82 31 219 1616.

E-mail address: sungmj@ajou.ac.kr (M. Sung).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.04.016

probit models where we implement the forward filtering backward sampling algorithm of Fruhwirth-Schnatter (1994). Our approach can be easily generalized to consider t distribution link functions that can be used to consider logit type dynamic models as well as Markov regression models.

We consider a binary time series X_t and we define a dynamic probit model similar to that considered by Andrieu and Daucet (2002) as

$$\Pr\{X_t = 1 \mid \pi_t\} = \pi_t \quad \text{with } \pi_t = \Phi(\boldsymbol{F}_t \boldsymbol{\theta}_t), \tag{1.1}$$

where $\boldsymbol{F}_t$ is a $1 \times K$ covariate vector and $\boldsymbol{\theta}_t$ is a $K \times 1$ vector of regression parameters. We define the dynamic nature of the model via a state equation for $\boldsymbol{\theta}_t$ as

$$\boldsymbol{\theta}_t = \boldsymbol{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{w}_t \tag{1.2}$$

with $\boldsymbol{w}_t$'s are uncorrelated multivariate normal error vectors with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{W}_\theta$ and $\boldsymbol{G}$ is the specified transition matrix of the model. It is most common to assume that $\boldsymbol{G}$ is an identity matrix implying a *steady model* in the sense of West and Harrison (1997). Thus, in our development we consider

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{w}_t. \tag{1.3}$$

The procedure can be easily extended for a general known transition $\boldsymbol{G}$ as well as for certain cases where $\boldsymbol{G}$ is unknown.

We can extend this for longitudinal data for individuals $i = 1, \ldots, M$. In this case we write the above model as

$$\Pr\{X_{it} = 1 \mid \pi_{it}\} = \pi_{it} \quad \text{with } \pi_{it} = \Phi(\boldsymbol{F}_{it} \boldsymbol{\theta}_t), \tag{1.4}$$

and assume the same state equation (1.3) for all individuals.

In Section 2, we introduce Bayesian inference for dynamic probit models and illustrate how an exact Gibbs sampler can be used. Extensions to Markov regression models and Student-t link functions are considered in Section 3. We illustrate how marginal likelihoods can be obtained and used to compare static versus dynamic probit models in Section 4. We also discuss how the marginal likelihood computations can be easily extended for Student-t link function case and thus can be used for selecting degrees of freedom. Implementation of our approach is illustrated in Section 5 using real data from the Great Smoky Mountains Study of Costello et al. (1996).

2. Bayesian inference for the dynamic probit model

We first consider the case for the ith individual Following Albert and Chib (1993), we can define the above model by using independent latent variables Z_{it} such that

$$X_{it} = \begin{cases} 1 & \text{if } Z_{it} > 0 \\ 0 & \text{Otherwise.} \end{cases} \tag{2.1}$$

If we assume that Z_{it}'s are normally distributed with mean $\boldsymbol{F}_{it}\boldsymbol{\theta}_t$ and variance 1, that is, $Z_{it} \sim N(\boldsymbol{F}_{it}\boldsymbol{\theta}_t, 1)$, then we have the probit model

$$\pi_{it} = \Phi(\boldsymbol{F}_{it}\boldsymbol{\theta}_t). \tag{2.2}$$

Given the above setup, we can develop a Gibbs sampler for the inference using the data augmentation algorithm of Albert and Chib (1993) with the algorithm proposed by Fruhwirth-Schnatter (1994) for dynamic linear models.

Given observed data $D = \{X_{it};\ t = 1, \ldots, T\}$, we can design a Gibbs sampler using full posterior conditional distributions $p(\Theta \mid D, \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}})$ and $p(\boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}} \mid D, \Theta)$ with vectors $\Theta = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_T)$ and $\boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}} = (Z_{i1}\, Z_{i2}\, \cdots\, Z_{iT})$. In obtaining $p(\boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}} \mid D, \Theta)$, we note that Z_{it}'s are independent random variables and use

$$(Z_{it} \mid \boldsymbol{\theta}_t, X_{it} = 1) \sim N(\boldsymbol{F}_{it}\boldsymbol{\theta}_t, 1)I \quad (Z_{it} > 0),$$

and

$$(Z_{it} \mid \boldsymbol{\theta}_t, X_{it} = 0) \sim N(\boldsymbol{F}_{it}\boldsymbol{\theta}_t, 1)I \quad (Z_{it} \le 0).$$

In the implementation of the Gibbs sampler, we can directly draw from the joint posterior distribution of $p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_T \mid \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}})$ using the *forward filtering backward sampling* algorithm of Fruhwirth-Schnatter (1994) which is given in West and Harrison (1997) for Kalman filter type models. It is possible to adopt the algorithm for our case as will be discussed next.

We define $\boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{t}} = (\boldsymbol{Z}_i^{\boldsymbol{t}-1}, Z_{it}),\ t = 1, \ldots, T$ and note that, similar to the Bayesian dynamic linear models of West and Harrison (1997), using the Markov structure of our model we can write $p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_T \mid \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}})$ as

$$p(\boldsymbol{\theta}_T \mid \boldsymbol{Z}_{\boldsymbol{i}}^{\boldsymbol{T}})\, p(\boldsymbol{\theta}_{T-1} \mid \boldsymbol{\theta}_T, \boldsymbol{Z}_i^{T-1}) \cdots p(\boldsymbol{\theta}_1 \mid \boldsymbol{\theta}_2, \boldsymbol{Z}_i^{1}), \tag{2.3}$$