Journal of Computational and Applied Mathematics 296 (2016) 613–624

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Covariance matrix and transfer function of dynamic generalized linear models

Guangbao Guo [a,b,*], Wenjie You [c], Lu Lin [a], Guoqi Qian [d]

[a] School of Mathematics, Shandong University, Jinan 250100, China
[b] Department of Statistics, Shandong University of Technology, Zibo 255000, China
[c] School of Electronic and Information Engineering, Fujian Normal University, Fuqing 350300, China
[d] School of Mathematics and Statistics, The University of Melbourne, Parkville VIC 3010, Australia

ARTICLE INFO

Article history:
Received 23 May 2015
Received in revised form 17 September 2015

MSC:
62J12
62M10
60-08
37C30

Keywords:
Dynamic generalized linear models
Covariance matrix
Transfer function

ABSTRACT

Statistical inference for dynamic generalized linear models (DGLMs) is challenging due to the time varying nature of the unknown parameters in these models. In this paper, we focus on the covariance matrix and the transfer function, the two key components in DGLMs. We first establish some convergence results for the covariance matrix estimation. We then provide an in-depth study of the transfer function on its stability and Fourier transformation, which is necessary for parameter estimation in DGLMs. Implications of our results on estimation in DGLMs are illustrated in the paper through a simulation study and a real data example. Our understanding on DGLMs has substantially improved though this study.

1. Introduction

West et al. [1] developed an extension of dynamic models by allowing the response observations to be non-Gaussian and to follow a probability distribution in the exponential family. This extension results in the so-called dynamic generalized linear models (DGLMs). Details about DGLMs can be found in e.g. [2–9]. DGLMs have brought ample opportunities, along with many challenges for statistical inference as well, for performing advanced statistical regression analysis for time-series data that contain regressors having time-varying effects. It is our objective in this paper that we undertake to understand some of these opportunities and challenges and to establish new results for DGLMs in regard to their covariance matrices and transfer functions.

A DGLM connecting a response time series $\mathbf{Y}_T = (y_1, \dots, y_T)$ with a regressor time series $\mathbf{X}_T = (X_1, \dots, X_T)$ and another $n \times 1$ deterministic regressor vector time series $\mathcal{F}_T = (\mathbf{F}_1, \dots, \mathbf{F}_T)$ can be formulated as follows:

$$y_t \sim f(y_t, \eta_t) \propto \exp\{y_t\eta_t - b(\eta_t)\}, \quad \text{(observation equation)} \tag{1}$$

$$g(\mu_t) = \eta_t = \upsilon(\delta)X_t + \mathbf{F}_t'\boldsymbol{\alpha}_t, \quad \upsilon(\delta)X_t \sim N_1(0, V),\ t = 1, \dots, T, \tag{2}$$

* Corresponding author at: School of Mathematics, Shandong University, Jinan 250100, China.
E-mail address: ggb11111111@163.com (G. Guo).

http://dx.doi.org/10.1016/j.cam.2015.10.015
0377-0427/

$$\boldsymbol{\alpha}_t = \mathbf{G}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim N_n(\mathbf{0}, \mathbf{W}_t), \quad \text{(transition equation)} \tag{3}$$

$$\upsilon(\delta) = \frac{\beta(\delta)}{\rho(\delta)}, \quad \text{(transfer function)}. \tag{4}$$

Here $f(y_t, \eta_t)$ is the probability density function (pdf) of Y_t given η_t that is assumed to belong to an exponential family with natural parameter η_t and scale parameter equal 1 for ease of presentation. The conditional mean $\mu_t = E(y_t|\eta_t) = \dot{b}(\eta_t) = db(\eta_t)/d\eta_t$ can be derived from $f(y_t, \eta_t)$, and is linked with the mean predictor by the link function $g(\cdot)$. In the current setting the natural parameter η_t is treated as the mean predictor which equals $\upsilon(\delta)X_t + \mathbf{F}_t'\boldsymbol{\alpha}_t$. Hence $g(\cdot) = \dot{b}^{-1}(\cdot)$ meaning $g(\cdot)$ is the canonical link. Effects of the evolution of $\mathbf{F}_t$ on μ_t are specified by the time-varying $n \times 1$ random vector parameter $\boldsymbol{\alpha}_t$ that has an autoregressive structure given by the transition equation, in which $\mathbf{G}$ is an $n \times n$ matrix and $\boldsymbol{\xi}_t$ follows an n-dimensional multivariate normal distribution $N_n(\mathbf{0}, \mathbf{W}_t)$ with mean vector of zero and covariance matrix of $\mathbf{W}_t$. We assume $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_T$ are mutually independent of each other. We also assume $y_1, \ldots, y_T$ are conditionally independent of each other given $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T$ and $\mathbf{X}_T$. Finally, $\upsilon(\delta) \equiv \upsilon_0 + \upsilon_1\delta + \upsilon_2\delta^2 + \cdots$ is the transfer function on the lag operator δ (i.e. $\delta X_t = X_{t-1}$), and measures the cumulative effect of the time series $\{X_t\}$ on the mean predictor η_t. Note the coefficients $\upsilon_0, \upsilon_1, \upsilon_2, \ldots$ are referred to as the impulse responses and express the instantaneous effect of X_t on η_t at present and future times. The transfer function $\upsilon(\delta)$ in DGLM is specified as the ratio of two polynomials on δ: $\beta(\delta) = \beta_0 + \beta_1\delta + \cdots + \beta_{m-1}\delta^{m-1} + \beta_m\delta^m$ and $\rho(\delta) = \rho_0 + \rho_1\delta + \cdots + \rho_{r-1}\delta^{r-1} + \rho_r\delta^r$. The roots of $\rho(\delta) = 0$ are called the poles of $\upsilon(\delta)$; the roots of $\beta(\delta) = 0$ are called the zeros of $\upsilon(\delta)$. Stability of the transfer function $\upsilon(\delta)$ is determined by whether these poles and zeros are outside the unit circle.

DGLMs as described above generalize both the dynamical regression models and generalized linear models (GLMs) from different perspectives, making them flexible and popular in applications. Although they can be further generalized by, e.g. including additional covariates with non-time-varying effects in the mean predictor equation (2) and extending X_t to be vector-valued, we will use DGLMs as in their current form for ease of presentation. Our primary focuses in this paper are then on deriving useful features of the covariance matrix and the transfer function in DGLM for assisting statistical inference.

The rest of the paper is organized as follows. In Section 2, we investigate the convergence properties of the covariance matrix in DGLM in general and also, as an illustration, in a special situation where the DGLM is canonical. In Section 3, we introduce some measures for approximating the transfer function in the DGLMs. The approximation performance is then studied based on using Fourier transformation. In Section 4 we provide a simulation study and a real data example to complement the results in Sections 2 and 3. Finally we provide some conclusion remarks in Section 5.

2. Convergence of covariance matrix estimation

2.1. Covariance matrix

Conditional covariance matrix $\mathbf{C}_t = var(\boldsymbol{\alpha}_t|\mathbf{Y}_t)(\mathbf{Y}_t = (y_1, \ldots, y_t))$ is a key component in DGLM and its efficient estimation plays a pivotal role there. When using a Bayesian approach, the posterior distributions involved in the definition of $\mathbf{C}_t$ may be multi-modal sometimes. In such situations enhanced MCMC methods such as parallel tempering (PT) and Population MCMC (Pop-MCMC) are often used to accelerate generating these difficult posterior distributions, see e.g. [7,10] for details.

For DGLM presented in Section 1, the posterior distribution of $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T$ is

$$\pi_T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T|\mathbf{Y}_T) \propto \prod_{t=1}^{T} f(y_t|\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t) f_{\alpha_0}(\boldsymbol{\alpha}_0|\mathbf{C}_0) f_{\alpha_1,\ldots,\alpha_t}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_t|\mathbf{C}_1, \ldots, \mathbf{C}_t) f_{C_t}(\mathbf{C}_t).$$

Denote the temperature sequence used in MCMC as $\{T_i\}_{i=1}^T$ satisfying $T_1 = 1 < T_2 < \cdots < T_i < \cdots < T_{T-1} < T_T$. Then PT is a parallel MCMC with T chains, each having a different stationary distribution $\pi_i(\cdot) \propto \pi_T(\cdot)^{\frac{1}{T_i}}$ defined on $\mathcal{B}(\mathcal{X})$, $i = 1, 2, \ldots, T$, where $\mathcal{B}(\mathcal{X})$ is a σ-algebra of the state space $\mathcal{X}$ of $(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T)$. Then PT is such a method that swaps the tth values in every pair of chains i and j according to the acceptance probability:

$$M_q(\alpha_t^{(i)}, \alpha_t^{(j)}) = \min\left\{1, \quad \frac{\pi_i(\alpha_t^{(j)})\pi_j(\alpha_t^{(i)})}{\pi_j(\alpha_t^{(j)})\pi_i(\alpha_t^{(i)})} \cdot \frac{q(\alpha_t^*|\alpha_t^{**})}{q(\alpha_t^{**}|\alpha_t^*)}\right\},$$

where $\alpha_t^* = (\alpha_t^{(1)}, \ldots, \alpha_t^{(i)}, \ldots, \alpha_t^{(j)}, \ldots, \alpha_t^{(N)})$ and $\alpha_t^{**} = (\alpha_t^{(1)}, \ldots, \alpha_t^{(j)}, \ldots, \alpha_t^{(i)}, \ldots, \alpha_t^{(N)})$ and $q(\cdot|\cdot)$ is the proposal density.

Pop-MCMC may be used to simulate the posterior distribution $\pi_T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_T|\mathbf{Y}_T)$ in the form of T Markov chains. Usually the joint stationary distribution for these T chains is of the form

$$\pi(\alpha) = \prod_{i=1}^{T} \pi_i(\alpha_i).$$