Statistics and Probability Letters 98 (2015) 59–64

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

On the degrees of freedom in MCMC-based Wishart models for time series data

Yuewen Xiao [a], Yu-Cheng Ku [b,*], Peter Bloomfield [c], Sujit K. Ghosh [c]

[a] Department of Statistics, School of Management, Fudan University, Shanghai, China

[b] Research and Analytics, Enterprise Risk Management, Fannie Mae, Washington, DC, USA

[c] Department of Statistics, North Carolina State University, Raleigh, NC, USA

ARTICLE INFO

Article history:
Received 29 March 2014
Received in revised form 5 December 2014
Accepted 11 December 2014
Available online 21 December 2014

Keywords:
Degrees of freedom
Gyndikin's theorem
Markov chain Monte Carlo
Sensitivity analysis
Wishart distribution

ABSTRACT

The Wishart distribution has long been a useful tool for modeling covariance structures. According to Gyndikin's theorem, the degrees of freedom (df) for a Wishart distribution can be any real number belonging to the Gyndikin set, either integer-valued or fractional. However, the fractional-df versioned Wishart distribution has received only limited attention, which may lead to inaccurate implementation in practice. This paper shows by a numerical example that, when implementing Markov chain Monte Carlo (MCMC) methods in Wishart models for time series data, the lack of attention to the fractional df where necessary can result in seriously biased posterior estimation due to the compounding errors caused by the time dependency assumption. We further conduct a sensitivity analysis to explain why the seemingly small difference between the integer-valued df and the fractional df leads to very different outcomes.

1. Introduction

The (inverse) Wishart distribution has long been a useful tool for modeling covariance structures, especially in the Bayesian inference, largely due to its conditional conjugacy property. In financial volatility modeling, Philipov and Glickman (2006a,b), Asai and McAleer (henceforth AM, Asai and McAleer, 2009), and Ku et al. (2014) proposed a class of time-varying covariance multivariate stochastic volatility (MSV) models, in which the evolution of the covariance matrices is described by inverse Wishart processes so that Bayesian Markov chain Monte Carlo (MCMC) approaches make model inference feasible. As the inference of this class of Wishart models is conducted with MCMC, the estimation process relies heavily on the random samples generated from a Wishart distribution. Throughout the paper, we denote the Wishart distribution as $W_p(k, \boldsymbol{S})$, where $\boldsymbol{S}$ is a $p \times p$ scale matrix and the scalar k is the degrees of freedom (df).

According to Gyndikin's theorem (see e.g., Graczyk et al., 2003), any k belonging to the Gyndikin set $\{1, 2, \dots, p-1\} \cup (p-1, \infty)$, can be a valid df for a $p \times p$ Wishart matrix, although the singular case where $k \in \{1, 2, \dots, p-1\}$ is used relatively less frequently in real applications. A number of methods satisfying Gyndikin's theorem have been proposed to generate random Wishart matrices, such as Bartlett's decomposition (Anderson, 2003) and its variant, the Odell–Feiveson algorithm (Odell and Feiveson, 1966). For non-singular Wishart matrices, $k \in (p-1, \infty)$ means that the df can be either integer-valued or fractional. However, in practice, the fractional df does not usually receive enough attention. Sometimes the use of the Wishart distribution is only focused on the integer case, where $\{k \geq p,\ k \in \mathbb{N}\}$, e.g., Härdle and Simar (2003),

* Corresponding author. Tel.: +1 202 752 4881; fax: +1 240 699 4114.
E-mail address: yku2@ncsu.edu (Y.-C. Ku).

http://dx.doi.org/10.1016/j.spl.2014.12.012
0167-7152/

Johnson and Wichern (2007) and Philipov and Glickman (2006a,b). The lack of attention to the fractional df is also seen in some statistical software packages, such as OX (Doornik, 2007), whose built-in Wishart matrix generator does not take fractional df into account. A natural question is then raised. What is the effect of using an integer df when in truth the df of the Wishart distribution is not integer valued? This question reflects an "implementation issue" which may not be noticed by the practitioners.

Consider the following case. Suppose we need to estimate a Wishart model using MCMC methods. In the estimation procedure, the df k is sampled from a continuous posterior distribution, which means that k can be any real number belonging to the Gyndikin set. Now, if we take the integer df for granted, then at this step when k is sampled, we will "integerize" k (e.g., taking the floor value, which is the method used by OX) even though it is drawn from a continuous density. In this case, we actually distort k from its derived posterior distribution, but unfortunately, we may not realize the distortion, as we believe that k should be integer-valued. We show by a numerical example that, in such a situation where k has been unconsciously distorted, the estimation process may fail to converge and, hence, the resulting inference can be seriously biased. We further provide a theoretical sensitivity analysis to study why the seemingly small difference between the integer k and the fractional $k' \in (k, k+1)$ can lead to thoroughly different results.

The remainder of the paper is organized as follows: Section 2 introduces Gyndikin's theorem and offers a simulation example to address the issue discussed above. Section 3 presents a sensitivity analysis to explain why the MCMC based methods may fail when the df is not correctly specified. Section 4 concludes with additional remarks.

2. Effects of mistakenly integerizing the DF in Wishart variate sampling

2.1. A review of Gyndikin's theorem

Let $\boldsymbol{V}$ be a $p \times p$ positive definite matrix and $\boldsymbol{\Theta}$ be an appropriate symmetric $p \times p$ matrix. Suppose we have identically and independently distributed random vectors $\boldsymbol{X}_i \sim \mathrm{N}_p(\boldsymbol{0}, \boldsymbol{V})$, the scatter matrix $\boldsymbol{W} = \frac{1}{2}\sum_{i=1}^{k} \boldsymbol{X}_i\boldsymbol{X}_i'$ has the Laplace transform $\mathrm{E}\left[e^{\mathrm{tr}(\boldsymbol{\Theta W})}\right] = |\boldsymbol{I}_p - \boldsymbol{V\Theta}|^{-k}$. Based on Graczyk et al. (2003) and Tourneret et al. (2005), the real-matrix version of Gyndikin's theorem states that, the set of values for k such that $|\boldsymbol{I}_p - \boldsymbol{V\Theta}|^{-k}$ is the Laplace transform of a positive measure equals $\mathcal{G}_s = \{1, 2, \ldots, p-1\} \cup (p-1, \infty)$, where $\mathcal{G}_s$ is called the Gyndikin set. Thus, by Gyndikin's theorem, we can generate Wishart variates with any real df k belonging to the Gyndikin set. In the special case where $k \in \{1, 2, \ldots, p-1\}$, the Wishart matrix is almost surely singular and its probability density function is defined in an appropriate space (Srivastava, 2003 and Díaz-García, 2007).

2.2. A numerical illustration

2.2.1. The model and the Bayesian inference

Suppose that we apply MCMC methods to estimate a Wishart model of df k. At the l-th iteration we draw the sample $k^{(l)}$ from a continuous conditional posterior distribution $p(k|\bullet)$. In this step, if we believe that k can only be integer-valued, then we will integerize $k^{(l)}$ to generate random Wishart matrices for subsequent sampling steps. Below we construct an example to investigate the impact of mistakenly integerizing the sample $k^{(l)}$.

To keep our illustration simple, we consider the following model, which is a special case of AM's dynamic correlation MSV (DCMSV2). Following a suggestion from one of the reviewers, we assume that the shock to return $\boldsymbol{\epsilon}_t$ is *observed* and has the following dynamics

$$\boldsymbol{\epsilon}_t \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma}_{\epsilon,t}), \quad \boldsymbol{\Sigma}_{\epsilon,t} = (diag\boldsymbol{Q}_t)^{-\frac{1}{2}}\boldsymbol{Q}_t(diag\boldsymbol{Q}_t)^{-\frac{1}{2}},$$

where the augmented covariance matrix $\boldsymbol{Q}_t$ follows an inverse Wishart process

$$\boldsymbol{Q}_t^{-1}|k, \boldsymbol{S}_{t-1} \sim W_p(k, \boldsymbol{S}_{t-1}), \quad \boldsymbol{S}_{t-1} = \frac{1}{k}\boldsymbol{Q}_{t-1}^{-d/2}\boldsymbol{A}\boldsymbol{Q}_{t-1}^{-d/2}, \qquad \boldsymbol{Q}_0 = \boldsymbol{I}_p.$$

Notice that the correlation matrix $\boldsymbol{\Sigma}_{\epsilon,t}$ is obtained by standardizing the covariance matrix $\boldsymbol{Q}_t$. The parameters k and $\boldsymbol{S}_{t-1}$ are the df and the scale matrix of the Wishart distribution, respectively. The "basic" matrix $\boldsymbol{A}$ is assumed to be symmetric and positive definite. The scalar parameter d accounts for the persistence of the process $\{\boldsymbol{Q}_t^{-1}\}$. The matrix power operation $\boldsymbol{Q}_{t-1}^{-d/2}$ is defined by a spectral decomposition.

To sample the parameters $\{\boldsymbol{A}, d, k\}$, we use the MCMC procedure developed by AM. The full conditional posterior density of the augmented variable $\boldsymbol{Q}_t^{-1}$ is obtained as

$$p\left(\boldsymbol{Q}_t^{-1}|\bullet\right) \propto W_p\left(\boldsymbol{Q}_t^{-1}|\hat{k}, \hat{\boldsymbol{S}}_{t-1}\right) \times f\left(\boldsymbol{Q}_t^{-1}\right), \tag{1}$$

where $\hat{k} = k + 1$ and $\hat{\boldsymbol{S}}_{t-1} = \left(\boldsymbol{S}_{t-1}^{-1} + \boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'\right)^{-1}$ denote the updated values, and the function $f\left(\boldsymbol{Q}_t^{-1}\right)$ is a remainder term.