ON THE STATIONARY PROBABILITY DENSITY FUNCTION OF BILINEAR TIME SERIES MODELS: A NUMERICAL APPROACH

R. Moeanaddin and H. Tong

Department of Statistics, Shiraz University, Islamic Republic of Iran
Mathematical Institute, University of Kent at Canterbury, U.K.

Abstract

In this paper, we show that the Chapman-Kolmogorov formula could be used as a recursive formula for computing the m-step-ahead conditional density of a Markov bilinear model. The stationary marginal probability density function of the model may be approximated by the m-step-ahead conditional density for sufficiently large m.

Introduction

Bilinear stochastic processes in discrete time series ɜre introduced by Granger and Andersen [1]. Aspects of obabilistic structure of bilinear models such as ɪtionarity, ergodicity, etc. have been considered by fferent authors [2, 3, 4, 6, 7, 8, 11]. However, the ɪestion of distribution of a stationary bilinear process has ɪt been discussed much in the literature. This is perhaps ɪe to the analytic intractability of the problem. For ɪample, Wang Shou-Ren *et al.* [15] have shown that the ɜtionary marginal probability density function of Markov linear model

$$X_t = aX_{t-1} + bX_{t-1}e_t + e_t \tag{1}$$

here $e_t \sim N(0,1)$ and $a^2 + b^2 < 1$, is the solution of the ɪtegral equation

$$f_b(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{f_b(s)}{|1+bs|} \exp\left[-\frac{1}{2}\left(\frac{x-as}{1+bs}\right)^2\right] ds \tag{2}$$

for $x \neq -\frac{a}{b}$ and $f_b(x) = \infty$ for $x = -\frac{a}{b}$, which does not seem to admit any analytic solution.

On the other hand, Moeanaddin and Tong [14] have succeeded in using the Chapman-Kolmogorov formula to obtain a sequence of conditional (probability) densities and the stationary marginal probability density function of the class of non-linear autoregressive models. In this paper, we shall show that a similar method may be used for computing a sequence of conditional densities of a Markov bilinear model. Then we may approximate the stationary marginal probability density function (if it exists) of the process by the m-step-ahead conditional density for sufficiently large m.

In the absence of any theoretical progress, it seems that a numerical approach is the only way out for the evaluation of stationary marginal probability densities of bilinear models. On the other hand, for complex bilinear models such as those with high orders, the computational burden due to the curse of dimensionality will lessen the attractiveness of the method. For such cases, the simulation approach might be a more practical solution, in which a long realization of the time series is generated by inputting pseudo-random numbers to the bilinear system. It is well known that extreme care is needed to ensure the 'randomness' of these numbers [10]. It is also difficult to assess the accuracy of the results thus obtained.

ɪeywords: Bilinear models; Chapman-Kolmogorov formula; ɪonditional density; Gauss-type formula; Integral equation; ɪarkovian representation; Matrix squaring; Numerical integration

We note that besides being of interest in its own right, the stationary distribution is needed if we want to evaluate the expected Fisher Information of the parameter estimates of a bilinear model. Moreover, the sequence of conditional densities are of practical importance in their own right. This is another important advantage of our method.

Markov Bilinear Models

A general bilinear process $\{X_t;\ t=0, \pm 1, \ldots\}$ can be defined by

$$X_t = a_0 + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=0}^{q} b_j e_{t-j} + \sum_{k=0}^{Q} \sum_{l=1}^{P} b_{kl} e_{t-k} X_{t-l},\ b_0=1 \quad (3)$$

where $\{e_t\}$ is a sequence of independent and identically distributed random variables with zero mean and constant variance.

Let us start with the simple Markov bilinear model (1). For this model, a sufficient condition for stationarity is that

$$E\{|a + b_{et}|\} < 1$$

or alternatively $a^2+b^2\sigma^2<1$ [12].

Using the Chapman-Kolmogorov relation, it can be shown that

$$h(x_m/x_0) = \int_R h(x_m/x_1) h(x_1/x_0) dx_1, \quad (4)$$

where $h(x_s|x_r)$ denotes the conditional density of X_{t+s} given $X_{t+r}=x_r$. In particular for $e_t \sim N(0, \sigma^2)$, the one-step-ahead conditional density of X_{t+1} given $X_t=x_0$, i.e. $h(x_1 mid x_0) \rightarrow h(x_1, x_0)$ is known precisely, namely

$$h(x_1|x_0) = \frac{1}{\sigma|1+bx_0|\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\left(\frac{x_1 - a x_0}{1+bx_0}\right)^2\right]. \quad (5)$$

In evaluating the improper integral of the form (4), we are faced with the difficulty of integrating over R. Since it is not easy to find an analytic solution for $h(x_m|x_0)$, we therefore have to use a suitable numerical method for computing this integral efficiently. Note that $h(x_1|x_0)$ and $h(x_m|x_1)$ are density functions and $h(x_1|x_0)$ tend to zero as $|x_1|$ tends to infinity. We adopt the following procedure: it is well known that if f is integrable over an interval, then there exist some points $x_1, x_2, \ldots, x_n$ with corresponding weights $\omega_1, \omega_2, \ldots, \omega_n$ such that $\int f(x)\,dx \approx \sum_{k=1}^{n} \omega_k f(x_k)$. In the Gauss-Hermite formula, this integral is exact for functions f being the product of a polynomial of degree $2n-1$ or less and the Gaussian density. Needless to say, when numerical integration is employed, care must be taken to handle the accumulation of rounding errors.

In practice we can employ the Gauss-Hermite or Gauss-Legendre formula to generate an appropriate set of points $\xi_1, \ldots, \xi_N$ with corresponding weights $\omega_1, \ldots, \omega_N$ [9]. By using the recursive formula (4) with (5), a sequence of conditional densities $h(x_m|x_0), m=2,3,\ldots$ can be calculated (Recall that $h(x_1|x_0)$ is known precisely once the noise density is known). As m increases, the conditional density $h(x_m|x_0)$ converges to f, the unique stationary marginal probability density function of X_t, which is assumed to exist. Convergence is deemed to have been achieved when $\max_{\xi_i} |f_m(\xi_i) - f_{m-1}(\xi_i)| < \epsilon$ for a small positive value of ϵ, e.g. $\epsilon = 10^{-6}$. Here $f_m(x) = h(x_m = x|x_0)$.

In our study, NAG routine D01BBF is employed to generate a set of points $\xi_1, \ldots, \xi_N$ with corresponding weights $\omega_1, \ldots, \omega_N$, We renormalize the conditional density to unity at each step, so as to avoid accumulating errors. Before re-normalization we also check the integral of conditional density as a precaution. If the evaluated integral is not close to one (i.e. with error $\geq 10^{-2}$), then we would change the parameters of the NAG routine or increase the number of points to generate another set of points and so on until a more adequate set is found. By starting with an arbitrary N and then by increasing the number of points, N, and integrating over a wider range and seeing a systematic convergence on the r-th moments ($r = 1,2,3,4$), we may assess the accuracy of the results.

As a check, the method has been applied to the Markov bilinear model (1) with $e_t \sim N(0,1)$, for different values of a and b. Since the theoretical rth, ($r=1,2,3,4$) moments for the models are available in this case, in Table 2.1 we compare the conditional density approach with $\epsilon < 10^{-6}$ and $N = 64$ (column marked numerical) with the theoretical (column marked theoretical).

Specifically, Tong [12] has shown that for model (1) with $e_t \sim N(0,1)$ the moments are given by

$\mu_1 = E(X_t) = 0$

$\mu_2 = E(X_t^2) = (1-a^2-b^2)^{-1}$

$\mu_3 = E(X_t^3) = 6ab\mu_2/\{1-a(a^2+3b^2)\}$

$\mu_4 = E(X_t^4) = \{12b(a^2+b^2)\mu_3 + 6(a^2+3b^2)\mu_2+3\}/\{1-a^4-6a^2b^2-3b^4\}$

The results can be summarized as follows:

Table 2.1

	a=b=0.1		a=b=0.3		a=-0.5, b=0.5	
μ_r	Theoretical	Numerical	Theoretical	Numerical	Theoretical	Numerical
μ_1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
μ_2	1.0204	1.0204	1.2195	1.2195	2.0000	2.0003
μ_3	0.0615	0.0615	0.7383	0.7383	-2.0000	-2.0002
μ_4	3.2495	3.2496	6.6513	6.6512	24.0000	23.8919

The stationary marginal probability density function of model (1) with $a = -0.5$, $b = 0.2$ and $N = 64$, which is approximated by the conditional density approach, is shown in Figure 2.1.

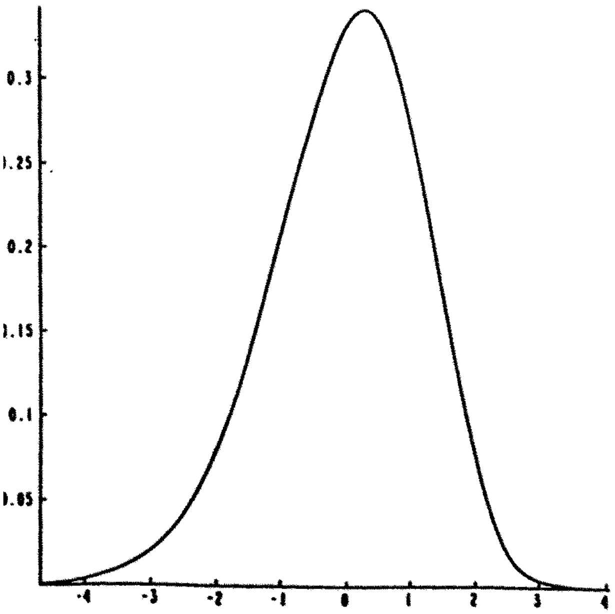


Figure 2.1. The stationary marginal probability density function of the model
$X_t = -0.5X_{t-1} + (1+0.2X_{t-1})e_t$, $e_t \sim N(0,1)$.

Superdiagonal Models

Consider the following simple superdiagonal bilinear model

$$X_t = \alpha X_{t-1} + \beta X_{t-2} e_{t-1} + e_t, \tag{6}$$

where e_t's are independent and identically distributed random variables with zero mean and constant variance. The condition $\alpha^2+\beta^2\sigma^2<1$ ensures the stationarity of the model (6).

To the best of our knowledge, no theoretical result is available for the conditional distribution of X_{t+m} given X_t and the stationary distribution of model (6). It is well known that the Markovian representation of model (6) takes the following form [13].

$$\xi_t = A\xi_{t-1} e_t + C e_t$$
$$X_t = H\xi_{t-1} + e_t \tag{7}$$

where

$$\xi_t = \begin{pmatrix} \xi_t^{(1)} \\ \xi_t^{(2)} \end{pmatrix} = \begin{pmatrix} X_t \\ \alpha X_t + \beta X_{t-1} e_t \end{pmatrix}$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & \alpha \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix}, C = \begin{bmatrix} 1 \\ \alpha \end{bmatrix}$$

and $H = (0,1)$.

Using Chapman-Kolmogorov relation, we have

$$h(\xi_m \mid \xi_0) = \int_{R^2} h(\xi_m \mid \xi_1) h(\xi_1 \mid \xi_0) d\xi_1, \tag{8}$$

where $h(\xi_s \mid \xi_r)$ denotes the conditional density of the vector process ξ at time $t+s$ given its value at time $t+r$, and $d\xi_1 = d\xi_1^{(1)} d\xi_1^{(2)}$. It is not difficult to show that for $e_t \sim N(0, \sigma^2)$

$$h(\xi_1 \mid \xi_0) = \frac{1}{\left|\alpha+\beta\xi_0^{(1)}\right|} \phi(\xi_1^{(1)} - \xi_0^{(2)}) \phi \frac{\xi_1^{(2)} - \alpha\xi_0^{(2)}}{\alpha+\beta\xi_0^{(1)}} \tag{9}$$

where ϕ (.) is the density function of e_t. In our numerical integration method, the integral equation (8) is approximated by the following summation

$$h(\xi_m \mid \xi_0) = h(\xi_m^{(1)}, \xi_m^{(2)} \mid \xi_0^{(1)} \xi_0^{(2)}) =$$

$$\Sigma_i \Sigma_j h(\xi_m^{(1)}, \xi_m^{(2)} \mid \xi_i^{(1)}, \xi_j^{(2)}) h(\xi_i^{(1)}, \xi_j^{(2)} \mid \xi_0^{(1)} \xi_0^{(2)}) \omega_i \omega_j$$

where $\xi_i^{(1)}, \xi_j^{(2)}, \omega_i, \omega_j$ are the points and their corresponding weights generated by the NAG routine.

Consequently, the stationary joint density function of

$$\xi_t = (\xi_t^{(1)}, \xi_t^{(2)})' = \begin{pmatrix} X_t \\ \alpha X_t + \beta X_{t-1} e_t \end{pmatrix}, \text{ say } \pi(\xi_t^{(1)}, \xi_t^{(2)}),$$

is approximated by $h(\xi_m \mid \xi_0)$ for sufficiently large m. Then $\int_R \pi(\xi_t^{(1)}, \xi_t^{(2)}) d\xi_t^{(2)}$ gives the stationary marginal probability density function of X_t.

Figure 3.1 shows the stationary marginal probability density function of mode

$$X_t = 0.3X_{t-1} + 0.4\, X_{t-2} e_{t-1} + e_t, \; e_t \sim N(0,1) \tag{10}$$

obtained by the above method.

Tong and Moeanaddin [14] have shown that the conditional variance of the error of an m-step non-linear least squares predictor is not necessarily a monotonic non-decreasing function of m. To illustrate this point, we consider the model

$$X_t = \beta X_{t-2} e_{t-1} + e_t, \; e_t \sim N(0,1). \tag{11}$$

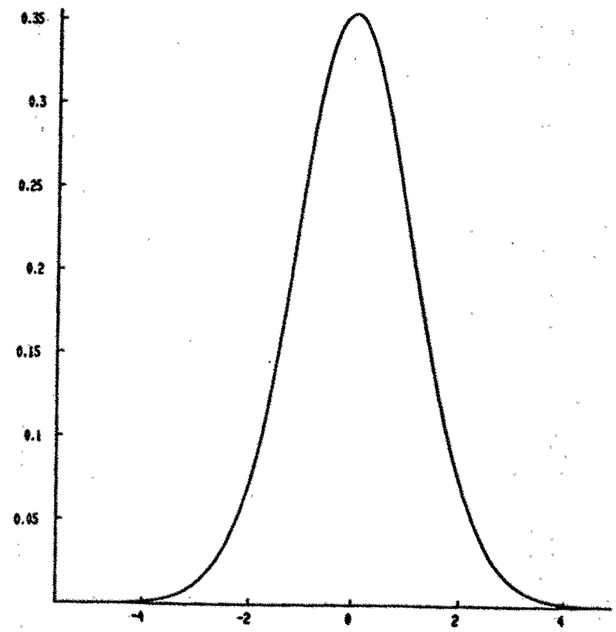


Figure 3.1. The stationary marginal probability density function of the model
$X_t = 0.3X_{t-1} + 0.4X_{t-2}e_{t-1} + e_t$, $e_t \sim N(0,1)$.

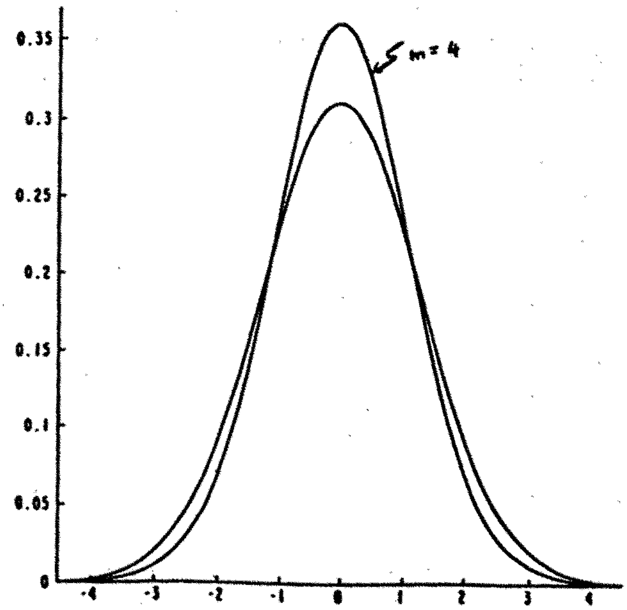


Figure 3.2. Conditional density function of X_{t+2}, and X_{t+4} given X_t=2 for the model
$X_t = 0.4X_{t-2}e_{t-1} + e_t$, $e_t \sim N(0,1)$.

To calculate the conditional variance of m-step-ahead error term, we only need to evaluate the conditional variance of X_{t+m}. Figure 3.2 shows the conditional densities of X_{t+2} and X_{t+4} given $X_t = 2$ for model (11) with $\beta = 0.4$. It is clear that the conditional density of $(X_{t+4} \mid X_t = 2)$ has a smaller variance than that of $(X_{t+2} \mid X_t = 2)$.

It is well known that for bilinear models, the expectation of X_t^m does not exist for all m, except when all parameters tend to zero. Our numerical experiences show that for model (6), the conditional densities and the stationary marginal distribution are symmetric and unimodal. Although the conditional density of X_{t+2} given X_t of model (6) with α=0 is normal, the conditional density converges to a density with tail probability heavier than normal as m increases.

Discussion

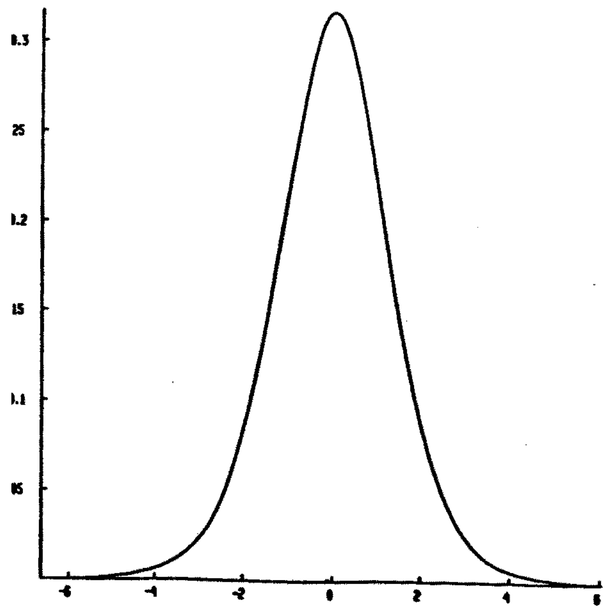
In principle, the extension of our method to the general class of bilinear models of the form (3) with $l > k$ is straightforward as it has a Markovian representation. In practice, to obtain the conditional densities, we have to be able to calculate the one-step-ahead conditional density. For some bilinear models we may encounter difficulties, e.g. when q is high. Note that for the general bilinear model, we usually need a high dimensional state space for the Markovian representation. This leads to numerical multiple integration with all the attendant implications. However, in the absence of any theoretical results, our method seems to provide a practical way of generating a sequence of conditional densities and approximating the stationary marginal density. Our method tends to experience fewer practical difficulties and the accuracy may be more controllable.

Next, the method of matrix squaring can sometimes be used to accelerate the convergence for some non-linear models [14]. However, for the two bilinear models considered in this paper, the method of matrix squaring does not turn out to be advantageous, since the conditional densities converge to the stationary marginal density in less than 15 steps anyway, without matrix squaring.

Finally, we remark that our approach could be used for calculating the stationary marginal probability density function and the conditional densities of random coefficient autoregressive (RCA) and ARCH models. As an example, consider the RCA model

$$X_t = (b_0 + b_1(t)) X_{t-1} + e_t, \tag{12}$$

where b_0 is a constant and $b_1(t) \sim N(0, \gamma^2)$, $e_t \sim N(0, \sigma^2)$, $b_1(t)$, e_t being independent. It is clear that $h(x_1 \mid x_0) \sim N(b_0 x_0, x_0^2\gamma^2 + \sigma^2)$. The stationary marginal probability density function of model (12) with $b_0 = 0.5$, $\gamma^2 = 0.25$, and $\sigma^2 = 1$, is shown in Figure 4.1.



'igure 4.1. The stationary marginal probability density function ıf the RCA model
$X_t = (0.5 + b_1(t)) X_{t-1} + e_t$,
where $b_1(t) \sim N(0, 0.25)$ and $e_t \sim N(0,1)$ and $b_1(t)$ and e_t are ndependent.

References

1. Granger, C.W.J. and Andersen, A.P. *An introduction to bilinear time series models.* Gottingen, Vandenhoek and Ruprecht, (1978).
2. Guegan, D. Etude d'un Modèle Nonlinèaire, le modèle super-diagonal d'ordre. *C.R. Acad. Sci. Paris*, **293**, 95-98, (1981).
3. Guegan, D. Different representation for bilinear model. *J. Time Series Anal.*, **8**, 389-408, (1987).
4. Guegan, D. and Pham, Dinh Tuan. Minimalitè des modèles bilineares á temps discret. *C.R. Acad. Sci. Paris*, **304**, 159-162, (1987).
5. Moeanaddin, R. and Tong, H. Numerical Evaluation of Distributions in Non-linear Autoregression. *J. Time Series Anal.*, **11**, 33-48, (1989).
6. Pham, Dinh Tuan. Bilinear Markovian representation and bilinear models. *Stoch. Proc. and Appl.*, **20**, 295-306, (1985).
7. Pham, Dinh Tuan. The mixing property of bilinear and generalized random coefficient autoregressive models. *Ibid.*, **23**, 291-300, (1986).
8. Pham, Dinh Tuan and Lanh, Tat Tran. On first-order bilinear time series models. *J. Appl. Prob.*, **18**, 617-627, (1981).
9. Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vettering, W.T. *Numerical recipes*, p. 125. Cambridge, Cambridge University Press, (1986).
10. Ripley, B.D. *Stochastics simulation.* New York, J. Wiley, (1987).
11. Subba Rao, T. and Gabr, M.M. *An introduction to bispectral analysis and bilinear time series models.* Lecture Notes in Statistics, **24**, New York, Springer-Verlag, (1984).
12. Tong, H. A note on a Markov bilinear stochastic process in discrete time. *J. Time Series Anal.*, **2**, 279-284, (1981).
13. Tong, H. *Nonlinear time series: a dynamical system approach.* Oxford, Oxford University Press, (1990).
14. Tong, H. and Moeanaddin, R. On multi-step non-linear least squares prediction. *Statistician*, **37**, 101-110, (1988).
15. Wang, S.R., An, H.Z. and Tong, H. On the distribution of a simple stationary bilinear process. *J. Time Series Anal.*, **4**, 209-216, (1983),