

The Univariate Distribution Function for a particular Bilinear Model

La Funzione di Distribuzione Univariata per un particolare Modello Bilineare

Francesco Giordano ¹

Dipartimento di Scienze Economiche e Statistiche

Università di Salerno

giordano@unisa.it

Riassunto: Questo lavoro si pone nell'ambito dei modelli bilineari per serie storiche. In particolare si affronta il problema di individuare la funzione di distribuzione univariata per una particolare rappresentazione del modello. L'obiettivo è quello di valutare la non linearità vista in termini di distribuzione per la quale si individua un'espansione in serie. La performance di tale espressione approssimata si verifica mediante un esperimento di simulazione.

Keywords: Nonlinearity, Bilinear Model, Edgeworth series.

1. Introduction

In recent years the bilinear models for time series gained a growing interest. The class of these models are difficult for their probabilistic structure. So there are many papers which concern specific bilinear models.

The idea of this paper comes from the COLS estimators introduced by Grahm (1995). Studying the properties of these estimators it can happen to simulate from a bilinear model. A way to improve the quality of the simulations is to generate from the distribution function instead of difference equation.

Consider the following bilinear model:

$$X_t = \epsilon_t + b\epsilon_{t-1}X_{t-2} \quad (1)$$

where b is an unknown parameter, $\epsilon_t \sim N(0, \sigma^2)$, $\forall t$, and, without loss of generality, $\sigma^2 = 1$.

This model is the simplest case of Type I class of Standardized bilinear models Grahm (1995). Model (1) can be easily confused with a White Noise because it has zero mean, constant variance and zero autocovariances. Further applying a parametric model to the observed time series it can happen that, from the Ljung-Box test, the residuals seem to be a White Noise even if some peaks of the series can be well captured by model (1). Anyway the stationary univariate distribution of X_t in (1) is different from the Normal one (assigned to the residuals ϵ_t).

The aim of this work is to derive an approximate univariate distribution function for a stationary X_t in (1). The second section shows the main results and the approximate distribution of X_t . A simulation experiment is presented in the third section in order

¹This paper is partially supported by MURST 2000: *Modelli stocastici e metodi di simulazione per l'analisi dei dati dipendenti*.

to evaluate the previous results. Finally some concluding remarks are given in the final section.

2. The approximate distribution function

The presentation of the main result needs to state some preliminaries given in the following lemma:

Lemma 1 *If X_t is defined as in (1) and X_t is a strong stationary process, then there exists an absolutely continuous distribution function. Besides, the density function $f(\cdot)$ of X_t is symmetric, positive, continuous and unimodal.*

Proof

From model (1) and given that X_t is a strong stationary process, it follows that its characteristic function, $H(u)$, is:

$$H(u) = e^{-\frac{1}{2}u^2} \int_{\mathbb{R}} e^{-\frac{1}{2}b^2u^2y^2} dF(y) \quad (2)$$

where $F(\cdot)$ is the stationary distribution function for X_t .

By the uniform bound of $e^{-\frac{1}{2}b^2u^2y^2}$, then

$$\int_{\mathbb{R}} |H(u)| du = \sqrt{2\pi} \int_{\mathbb{R}} \frac{1}{\sqrt{1+b^2y^2}} dF(y) \leq \sqrt{2\pi} < \infty$$

which implies that a density function $f(\cdot)$ exists.

Further $H(u) = 1$ if and only if $u = 0$, then it is impossible that $f(\cdot)$ can be an arithmetic function at some points over $\mathbb{R}$.

Given that the characteristic function is real then $f(\cdot)$ is symmetric.

By the Inversion Theorem, $f(\cdot)$ is

$$f(x) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(1+b^2y^2)}} e^{-\frac{x^2}{2(1+b^2y^2)}} f(y) dy \quad (3)$$

and so $f(\cdot)$ is defined by a second kind integral equation. If $f(x_0) = 0$ then

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(1+b^2y^2)}} e^{-\frac{x_0^2}{2(1+b^2y^2)}} f(y) dy = 0$$

the last equation can be satisfied only for $f(x) = 0, \forall x \in \mathbb{R}$, which is impossible.

Given that the integrand function in (3)

$$g(x, y) = \frac{1}{\sqrt{2\pi(1+b^2y^2)}} e^{-\frac{x^2}{2(1+b^2y^2)}}$$

has the first partial derivative with respect to x , $\forall x$, and $\frac{\partial g(x, y)}{\partial x}$ is uniformly dominated by $1/\sqrt{2\pi}$, then

$$\frac{df(x)}{dx} = -x \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(1+b^2y^2)}^3} e^{-\frac{x^2}{2(1+b^2y^2)}} f(y) dy$$

So $f(\cdot)$ is continuous and it has a maximum only at $x = 0$.
It follows that $f(\cdot)$ is unimodal. $\diamond$

Therefore, by lemma 1 the expression (3) is the true univariate density function for X_t in (1). To find an approximate closed form for (3) it can be stated the following theorem:

Theorem 1 *If X_t is defined as in (1), X_t is strong stationary and ergodic process and $|b| < \sqrt{2/5}$, then there exists an approximate univariate distribution function*

$$\tilde{F}(x) = \Phi(x) - cx\phi(x) \quad (4)$$

with $c = \frac{b^2}{2(1-b^2)}$.

Proof

Starting from (2) and, by Taylor expansion, we have

$$\tilde{H}(u) = e^{-\frac{1}{2}u^2} \int_{\mathbb{R}} \left(1 - \frac{1}{2}b^2u^2y^2\right) f(y)dy \quad (5)$$

By the Inversion Theorem we have the result in (4). The expression (4) must have the density function of lemma 1 and this is true if $|b| < \sqrt{2/5}$. $\diamond$

Remark 1 Investigating the Taylor expansion in (5), it is easy to verify that (4) underestimate the true function.

Remark 2 $|b| < \sqrt{2/5}$ implies that the sixth moment exists, that is $E(X_t^6) < \infty$.

Remark 3 The distribution (4) can be seen as an Edgeworth series at the second term. So theorem 1 states that there is an Edgeworth expansion for the stationary univariate distribution function for X_t defined in (1).

The approximate distribution function (4) depends on the unknown parameter b . If we consider a consistent estimator for b , such as the COLS shown in Grahn (1995), say $\hat{b}_{COLS}$, then $\hat{\tilde{F}}(x) = \Phi(x) - \hat{c}_{COLS}x\phi(x) \xrightarrow{a.s.} \tilde{F}(x)$.

This convergence can be proved by $\hat{b}_{COLS} \xrightarrow{a.s.} b$ Grahn (1995) and

$$\hat{c}_{COLS} = \frac{\hat{b}_{COLS}^2}{2(1-\hat{b}_{COLS}^2)} \xrightarrow{a.s.} c \text{ because it is a continuous function of } b.$$

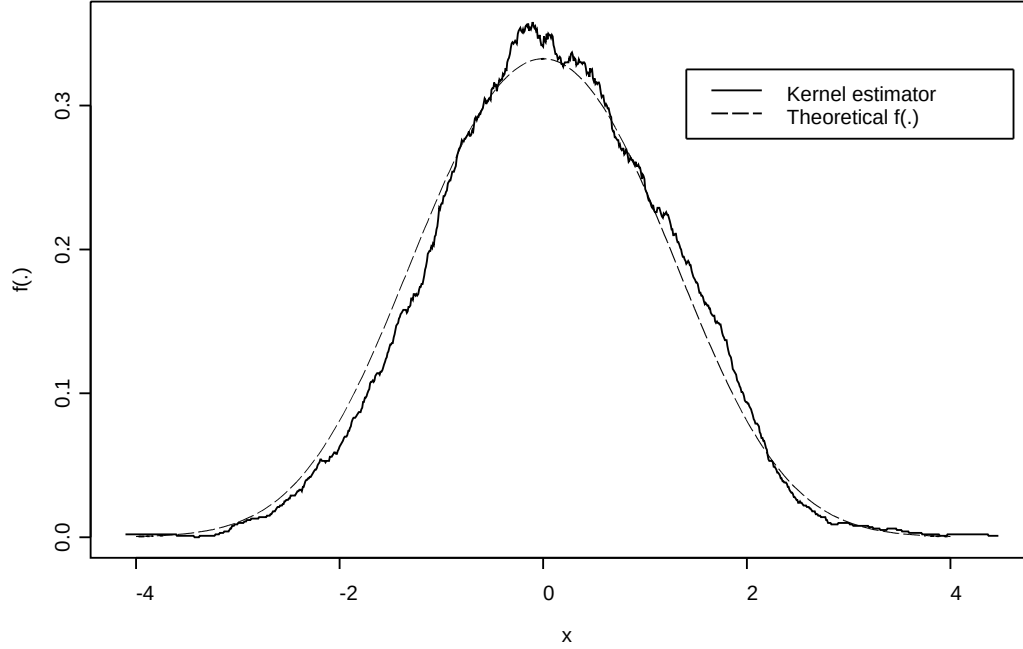
The previous results show that there exists a distribution function for X_t and its approximation is (4). Nevertheless there is the crucial problem of the existence of moments. If we suppose that all moments exist, then the parameter b must be zero. In this case the model (1) reduces to a White Noise.

3. Empirical results

Taking into account model (1) the experiment is drawn by generating 1000 series of length $n=300$ observations. Figure 1 reports the Kernel density estimate and the density function computed from (4) using the value of b to 0.5.

Other values of $|b| \leq 0.5$ also lead to graphs where the Kernel estimate curve is nearly over the theoretical one. Hence, in these cases the proposed approximation (4) works well. Instead, when $|b|$ tends to $\sqrt{2/5}$ the fitting is not as good as in the previous case.

Figure 1: Kernel estimate and theoretical $\tilde{f}(\cdot)$, with $b = 0.5$



4. Concluding remarks

The expression (4) can be used to estimate quantiles both directly or by means of Conish-Fisher method. Besides, knowing the distribution of X_t , it is possible to compute the unconditional expected length of one-step ahead prediction interval Kabaila and He (2001). The result of the second section gives the conditions for the Edgeworth expansion of the univariate stationary distribution of X_t . It is important to note that the process X_t has a dependent structure with zero autocovariances.

The approximation (4) can be improved and extended to a more general bilinear model even if the most important constraint is the existence of some moments for X_t .

A topic for future development is the derivation of a method which improves the approximation without requiring further conditions on the moments of X_t .

Finally another important subject is to understand the measure of the error using the approximation (4).

References

- Grahn T. (1995) A conditional least squares approach to bilinear time series estimation, *Journal of Times Series Analysis*, 16, 509–529.
- Kabaila P. and He Z. (2001) On prediction intervals for conditionally heteroscedastic processes, *Journal of Time Series Analysis*, 22, 725–731.
- Tong H. (1990) *Non Linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.