NUMERICAL EVALUATION OF DISTRIBUTIONS IN NON-LINEAR AUTOREGRESSION

By R. Moeanaddin and Howell Tong

University of Kent at Canterbury

First version received October 1988

Abstract. We use the Chapman–Kolmogorov formula as a recursive relation for computing the m-step-ahead conditional density of a non-linear autoregressive model. We approximate the stationary marginal probability density function of the model by the m-step-ahead conditional density for sufficiently large m. An advantage of our method is its simple implementation; only one NAG subroutine is needed. We have also studied the advantage of incorporating the matrix-squaring procedure.

Keywords: Chapman–Kolmogorov formula; conditional density; eigenvalues; EXPAR models; Gauss-type formula; integral equation; matrix squaring; non-linear autoregressive models; numerical integration; SETAR models.

1. INTRODUCTION

Although the classical linear autoregressive model is very useful as a first-order approximate explanation of stochastic discrete time phenomena, the generating mechanisms of most physical phenomena are essentially non-linear. Jones (1978) has considered the first-order non-linear autoregressive model NLAR(1)

$$X_t = \lambda(X_{t-1}) + e_t \qquad (t = 0, \pm 1, \ldots) \tag{1.1}$$

where $\lambda(.)$ is a fixed real function of a real argument and $\{e_t\}$ is a sequence of independent identically distributed random variables with zero mean and constant variance and probability density function $k(.)$

The problem of evaluating the stationary marginal distribution of model (1.1) explicitly is extremely involved. Andel *et al.* (1984) have managed to obtain the explicit analytic solution for the special threshold model

$$X_t = -\alpha |X_{t-1}| + e_t \tag{1.2}$$

where $\alpha \in (0, 1)$ and $e_t \sim N(0, 1)$. The stationary marginal density function is shown to be

$$f(x) = \left\{\frac{2(1-\alpha^2)}{\pi}\right\}^{1/2} \exp\left\{-\frac{1}{2}(1-\alpha^2)x^2\right\} \Phi(-\alpha x) \tag{1.3}$$

where $\Phi(.)$ is the distribution function of the standard normal.

0143-9782/90/01 033–16 $02.50/0

In a similar vein, Chan and Tong (1986) have considered a much more general case in which there is a compact group structure associated with $\lambda(.)$ in (1.1). Note that a particularly simple rotation group is associated with (1.2).

However, in a general NLAR model, we cannot always expect some nice group structure and the explicit form of the stationary marginal distribution is usually intractable. Therefore we have to resort to numerical methods in many cases.

2. NUMERICAL SOLUTION

We shall review some of the numerical methods briefly. Later we shall employ the Chapman–Kolmogorov relation to obtain a simple recursive formula. Without loss of generality, we restrict our discussion to the first-order case, i.e. NLAR(1), which we always assume to be strictly stationary in this paper. We also assume that the stationary distribution is absolutely continuous with a density $f(.)$. Clearly,

$$f(x) = \int_{-\infty}^{\infty} k(x - \lambda(y))f(y)dy. \tag{2.1}$$

2.1 *Series expansion*

Jones (1976, 1978) has derived power series expansions for obtaining the probability density function and moments of a stationary discrete time series model (1.1). The method involves consideration of the family of time series indexed by the parameter β:

$$X_t(\beta) = a + bX_{t-1}(\beta) + \beta[\lambda\{X_{t-1}(\beta)\} - a - bX_{t-1}(\beta)] + e_t$$

where a is an arbitrary but fixed constant and $|b| < 1$. The resulting series expansion in powers of β usually requires b to be zero for practical calculation, which involves truncation. The value of the constant a affects the convergence of the series and so its value has to be chosen carefully. Some difficulties have been experienced with convergence when $\lambda(.)$ is piecewise linear and the variance of e_t is 'small' (Pemberton, 1985).

2.2 *Iterative numerical quadrature method*

Pemberton (1985) has proposed the following numerical iteration for solving the integral equation (2.1) for $f(.)$:

$$f_{n+1}(x) = C_n \int_{-\infty}^{\infty} f_n(y)k(x - \lambda(y))dy \tag{2.2}$$

with some suitable starting function $f_0(x)$ normalized so that

$$\int_{-\infty}^{\infty} f_0(x)dx = 1.$$

The multiplying factor C_n is a renormalization constant and is necessary because typically the integral on the right-hand side of (2.2) is evaluated numerically over a finite range of points. The starting function $f_0(x)$ can be the standard normal density. It is noted by Pemberton (1985) that convergence is not always achieved.

2.3 *Conditional density approach*

The idea of approximating the stationary distribution as the limit of conditional distributions was used by Jones (1976) as one of his three methods. In practice, it appears that it is not easy to use his results as complicated series expansions are again involved. We propose an alternative way of exploiting the Chapman–Kolmogorov relation.

Since we have a Markov chain over R, we recall the Chapman–Kolmorov relation

$$f(x_m \mid x_0) = \int_{-\infty}^{\infty} f(x_m \mid x_1) f(x_1 \mid x_0) dx_1 \qquad (m = 1, 2, \ldots) \tag{2.3}$$

where $f(x_s \mid x_t)$ denotes the conditional probability density function of X_s given $X_t = x_t$. Note that

$$f(x_1 \mid x_0) = k(x_1 - \lambda(x_0)). \tag{2.4}$$

In evaluating the improper integral in (2.3), we are faced with two possible difficulties: (i) that *discontinuity problem* and (ii) the *finite-range problem*. The former may arise because λ may be discontinuous as in a threshold model, in which case the integrand of (2.3) will be discontinuous over the integrating variable x_1. This type of discontinuity problem can be handled by splitting the integral into disjoint intervals over each of which the integrand is continuous. This is easily achieved for the threshold model. The latter problem is well known. To handle it, we follow Pemberton by incorporating renormalization. We adopt the following procedure.

It is well known that, if g is an integrable function over an interval, there exist some points $\xi_1, \xi_2, \ldots, \xi_N$ with corresponding weights $w_1, w_2, \ldots, w_N$ such that

$$\int g(x)dx \simeq \sum_{k=1}^{N} w_k g(\xi_k),$$

which is exact for functions g of class P_{2N-1}, the class of all functions that are polynomials of degree $2N - 1$ or less.

In practice, we can employ some Gauss-type formulae to generate an appropriate set of points $\xi_1, \ldots, \xi_N$ with corresponding weights $w_1, \ldots, w_N$. By using the recursive formula (2.3) with (2.4), a sequence of conditional densities $f(x_m \mid x_0)$ $(m = 2, 3, \ldots)$ can be calculated. As m increases, the conditional density $f(x_m \mid x_0)$ converges to the unique stationary marginal probability density function f of X_t, which is assumed to exist. Convergence is deemed to have been achieved when $\max_x |f_m(x) - f_{m-1}(x)| < \epsilon$ for a small

positive value of ϵ, e.g. $\epsilon = 10^{-6}$. A referee has suggested replacing ϵ by $\epsilon \max_x f_m(x)$. In our experience, the latter choice will improve the consistency of convergence especially for widely dispersed distribution.

In our study, the NAG routine DO1BBF is employed to generate an appropriate set of points $\xi_1, \ldots, \xi_N$ with corresponding weights $w_1, \ldots, w_N$. (This subroutine automatically calls another NAG routine, namely DO1BAX.) We renormalize the conditional density to unity at each step so as to avoid accumulating errors. Before renormalization we also check the integral of conditional density as a precaution. If the evaluated integral is not close to unity (i.e. with error 10^{-3} or more) then we change the 'scaling' parameter of the NAG routine DO1BBF to generate another set of points and so on until a more adequate set is found.

3. APPLICATION

As a check on the method, it has been used to obtain the unique stationary marginal density of the linear AR(1) process

$$X_t = \alpha X_{t-1} + e_t \qquad (|\alpha| < 1) \tag{3.1}$$

and the SETAR (2; 1, 1) process

$$X_t = -\alpha |X_{t-1}| + e_t \qquad (0 < \alpha < 1) \tag{3.2}$$

where $e_t \sim \mathrm{N}(0, 1)$, for different values of α and different numbers of points. Since the theoretical stationary probability density functions and the theoretical rth moments for models of the form (3.1) and (3.2) are available, in Tables I and II we compare the rth moments ($r = 1, 2, 3, 4$) based on the conditional density approach with $\epsilon < 10^{-6}$ (the columns headed 'Full numerical method') with the theoretical approach (columns headed 'Theoretical'). Using the same points $\xi_1, \ldots, \xi_N$ and $w_1, \ldots, w_N$ and the theoretical probability density function $f(x)$, we can also perform numerical integration to evaluate $\int_{-\infty}^{\infty} x^r f(x)dx$. The results are listed in the columns headed 'Moment derived numerically from theoretical density'. The columns headed 'Run time' refer to the computation time for the columns headed 'Full numerical method'.

It is interesting to note that $E(X_t^2)$ for model (3.2) is equal to that for the linear model (3.1). We can also check that the $E(X_t^4)$ of these two models are equal for fixed α.

As further evidence of the efficacy of the conditional density method, we have computed the density and moments of a threshold model considered by Jones (1976). This is of the form $X_t = \lambda(X_{t-1}) + e_t$, where $e_t \sim \mathrm{N}(0, 1)$ and

$$\lambda(x) = \begin{cases} 0.5 + x & \text{if } -1.5 < x \leq 1 \\ -1 & \text{otherwise} \end{cases} \tag{3.3}$$

TABLE I

AR(1)

N	α	$E(X_t)$			$E(X_t^2)$			Run time (s)	$E(X_t^3)$			$E(_t^4)$		
		Full numerical method	Moment derived numerically from theoretical density	Theoretical	Full numerical method	Moment derived numerically from theoretical density	Theoretical		Full numerical method	Moment derived numerically from theoretical density	Theoretical	Full numerical method	Moment derived numerically from theoretical density	The cal
12	−0.9	0.00086	0.00000	0.0	5.26547	5.26316	5.26316	4.9	0.00136	0.00000	0.0	83.15470	83.10154	83.1
24	−0.9	0.00000	0.00000	0.0	5.26316	5.26316	5.26316	5.6	0.00000	0.00000	0.0	83.10223	83.10243	83.
48	−0.9	0.00000	0.00000	0.0	5.26316	5.26315	5.26316	6.9	0.00000	0.00000	0.0	83.10185	83.10229	83.1
64	−0.9	0.00000	0.00000	0.0	5.26315	5.26316	5.26316	11.4	0.00000	0.00000	0.0	83.10244	83.10246	83.1
12	−0.5	0.00000	0.00000	0.0	1.33333	1.33333	1.33333	1.0	0.00000	0.00000	0.0	5.33173	5.33180	5.3
24	−0.5	0.00000	0.00000	0.0	1.33333	1.33333	1.33333	1.5	0.00000	0.00000	0.0	5.33333	5.33333	5.3
48	−0.5	0.00000	0.00000	0.0	1.33333	1.33333	1.33333	2.9	0.00000	0.00000	0.0	5.33333	5.33333	5.3
64	−0.5	0.00000	0.00000	0.0	1.33333	1.33333	1.33333	4.4	0.00000	0.00000	0.0	5.33333	5.33333	5.3
12	−0.2	0.00000	0.00000	0.0	1.04167	1.04167	1.04167	0.8	0.00000	0.00000	0.0	3.25521	3.25521	3.2
24	−0.2	0.00000	0.00000	0.0	1.04167	1.04167	1.04167	1.2	0.00000	0.00000	0.0	3.25521	3.25521	3.2
48	−0.2	0.00000	0.00000	0.0	1.04167	1.04167	1.04167	2.5	0.00000	0.00000	0.0	3.25521	3.25521	3.2
64	−0.2	0.00000	0.00000	0.0	1.04167	1.04167	1.04167	3.4	0.00000	0.00000	0.0	3.25521	3.25521	3.2

TABLE II

SETAR (2; 1, 1)

N	α	$E(X_t)$			$E(X_t^2)$			Run time (s)	$E(X_t^3)$			$E(_t^4)$		
		Full numerical method	Moment derived numerically from theoretical density	Theoretical	Full numerical method	Moment derived numerically from theoretical density	Theoretical		Full numerical method	moment derived numerically from theoretical density	Theoretical	Full numerical method	Moment derived numerically from theoretical density	Th et
24	0.9	1.64742	1.64742	1.64742	5.27112	5.26316	5.26316	5.2	19.04422	18.98873	18.98874	83.40329	83.10229	83.1
48	0.9	1.64742	1.64742	1.64742	5.26314	5.26311	5.26316	6.8	18.98840	18.98809	18.98874	83.09683	83.09399	83.1
96	0.9	1.64742	1.64742	1.64742	5.26316	5.26316	5.26316	24.3	18.98893	18.98874	18.98874	83.10239	83.10245	83.1
128	0.9	1.64742	1.64742	1.64742	5.26312	5.26315	5.26316	40.1	18.98837	18.98861	18.98874	83.09853	83.10078	83.1
24	0.5	0.46066	0.46066	0.46066	1.33333	1.33333	1.33333	1.4	1.68908	1.68908	1.68908	5.33333	5.33333	5.3
48	0.5	0.46066	0.46066	0.46066	1.33333	1.33333	1.33333	2.8	1.68908	1.68908	1.68908	5.33333	5.33333	5.3
96	0.5	0.46066	0.46066	0.46066	1.33333	1.33333	1.33333	8.1	1.68908	1.68908	1.68908	5.33333	5.33333	5.3
128	0.5	0.46066	0.46066	0.46066	1.33333	1.33333	1.33333	12.0	1.68908	1.68908	1.68908	5.33333	5.33333	5.3
24	0.2	0.16287	0.16287	0.16287	1.04167	1.04167	1.04167	1.0	0.50217	0.50217	0.50217	3.25521	3.25521	3.2
48	0.2	0.16287	0.16287	0.16287	1.04167	1.04167	1.04167	2.4	0.50217	0.50217	0.50217	3.25521	3.25521	3.2
96	0.2	0.16287	0.16287	0.16287	1.04167	1.04167	1.04167	6.1	0.50217	0.50217	0.50217	3.25521	3.25521	3.2
128	0.2	0.16287	0.16287	0.16287	1.04167	1.04167	1.04167	8.9	0.50217	0.50217	0.50217	3.25521	3.25521	3.2

The results are given in Table III. The method is feasible even for long-tailed distributions as shown in Figure 1.

We now illustrate results concerning the stationary and conditional densities etc.

EXAMPLE 1. Let $\{X_t\}$ satisfy the SETAR (2; 1, 1) model

$$X_t = \begin{cases} 1.5 - 0.9X_{t-1} + e_t & \text{if } X_{t-1} \leq 0 \\ -0.4 - 0.6X_{t-1} + e_t & \text{if } X_{t-1} > 0 \end{cases} \tag{3.4}$$

where $e_t \sim \mathrm{N}(0, \sigma^2)$. This model has a limit cycle of period 2 at $C = \{-2.826, 4.043\}$. The stationary marginal probability density function of model (3.4) is shown in Figure 2(a) for $\sigma^2 = 1$ and in Figure 2(b) for $\sigma^2 = 4$.

TABLE III

n	$E(X_t)$	$E(X_t^2)$	$E(X_t^3)$	$E(X_t^4)$	Skewness	Kurtosis	Run time (s)
16	−0.18206	1.68061	−0.62693	7.87811	0.13195	−0.14362	5.78
20	−0.18232	1.68340	−0.65533	7.81880	0.11957	−0.18397	4.19
28	−0.18216	1.68417	−0.65196	7.83041	0.12083	−0.17972	4.58
40	−0.18216	1.68416	−0.65198	7.83036	0.12082	−0.17974	5.52
56	−0.18216	1.68416	−0.65198	7.83036	0.12082	−0.17974	9.64
Jones's	−0.18216	1.68416	—	—	0.12082	−0.17974	—

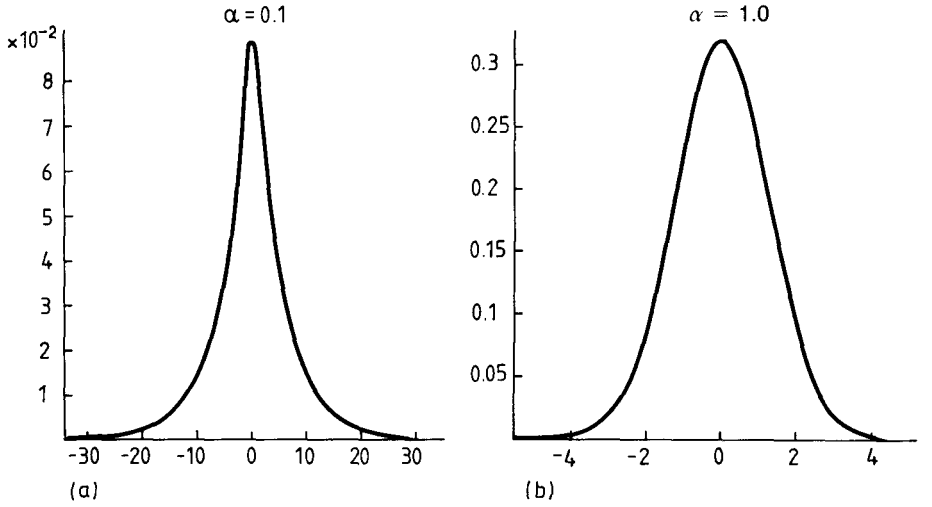


FIGURE 1. The stationary marginal probability density function of the model

$$X_t - X_{t-1} = \begin{cases} \alpha + \epsilon_t & \text{if } X_{t-1} \leq 0 \\ -\alpha + \epsilon_t & \text{if } X_{t-1} > 0 \end{cases}$$

where $\epsilon_t \sim \mathrm{N}(0, 1)$ and (a) $\alpha = 0.1$, (b) $\alpha = 1.0$.

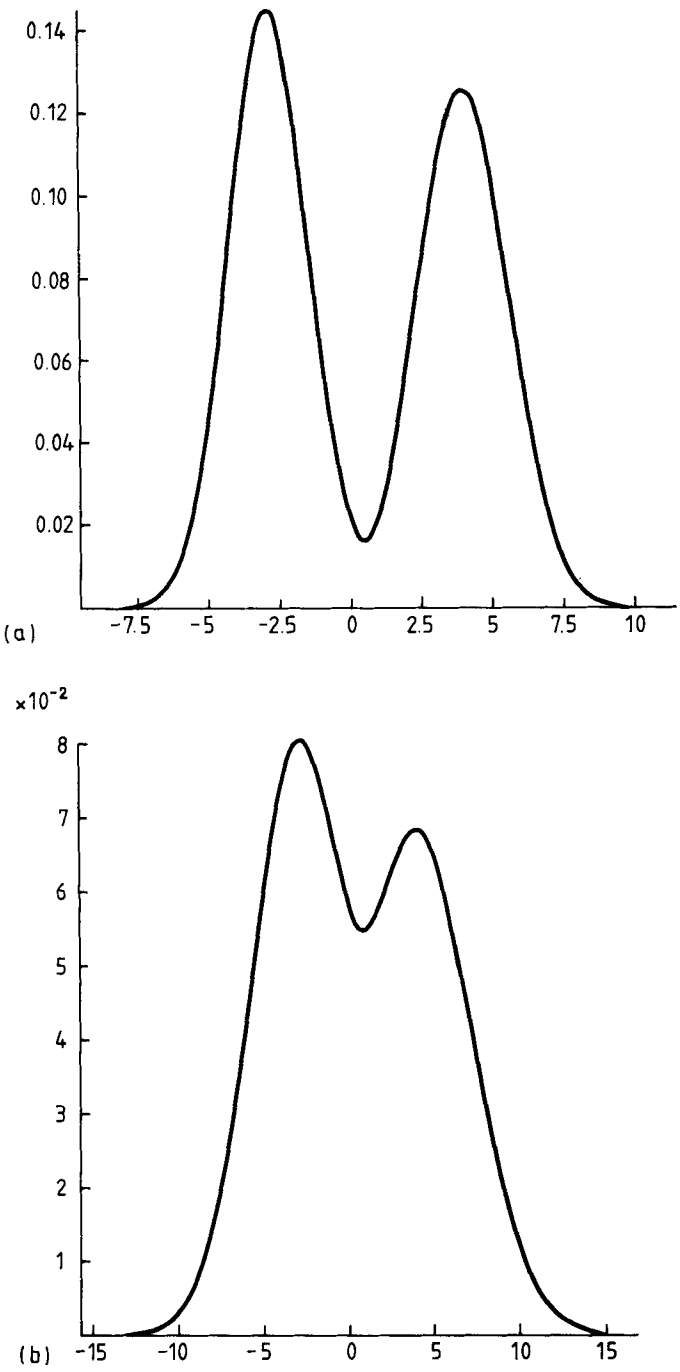


FIGURE 2. The stationary marginal probability density function of model (3.4) with (a) $\sigma^2 = 1$ and (b) $\sigma^2 = 4$.

We also plot the conditional density of X_{t+m} given $X_t = 1$ for $m = 1, 2, 3, 4, 5$ and $\sigma^2 = 1$ to highlight the non-normality and complexity of these densities for $m \geq 2$. Obviously the conditional density of X_{t+1} given X_t is normal. The results are shown in Figure 3. The joint probability density

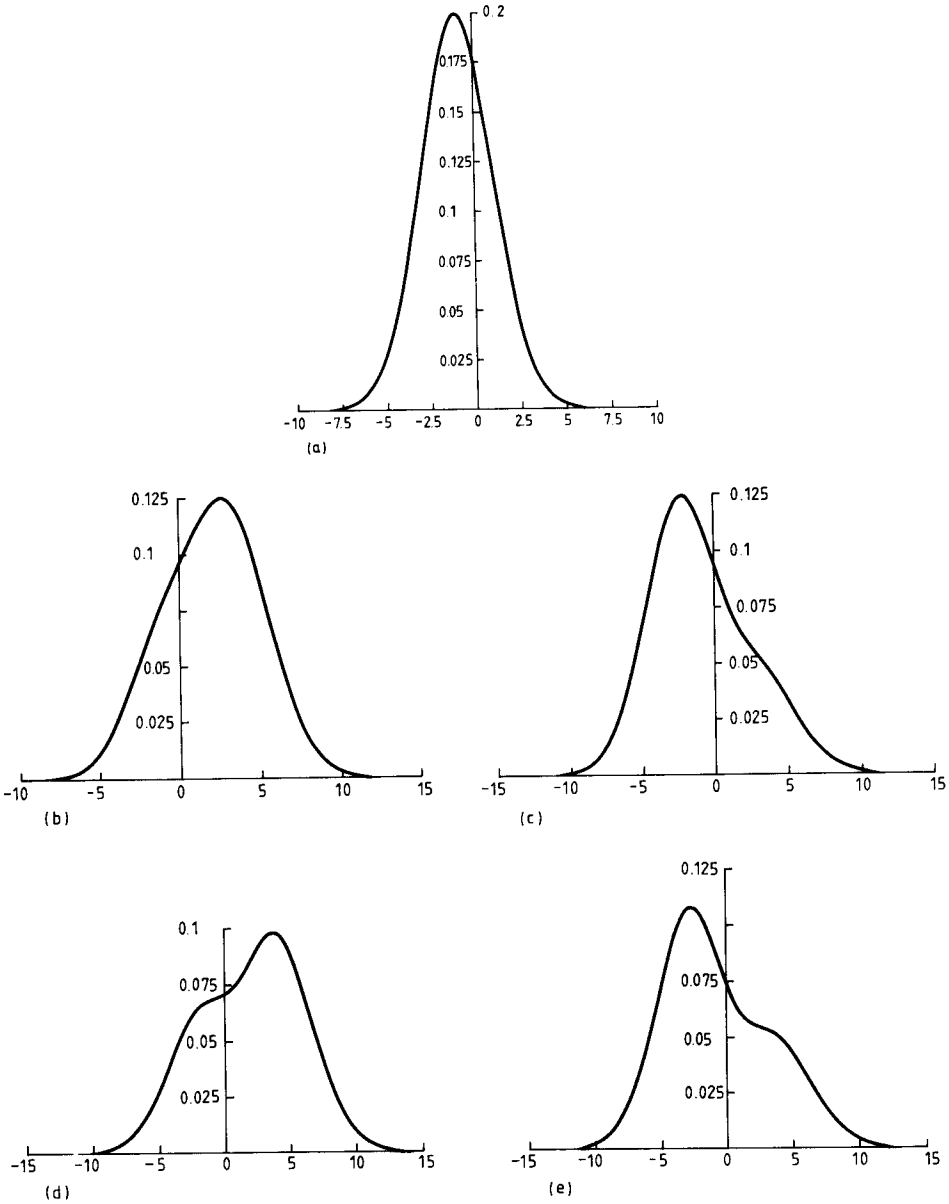


FIGURE 3. The conditional density function of $X_{t+m} | X_t = 1$ of model (3.4) with $\sigma^2 = 1$: (a) $m = 1$; (b) $m = 2$; (c) $m = 3$; (d) $m = 4$; (e) $m = 5$.

function $f_{t+1,t}(.\ ,.)$ of (X_t, X_{t-1}) can be found by using

$$f_{t+1,t}(x, y) = k(x - \lambda(y))f(y).$$

Since the process is Markovian, this can be extended to obtain the joint density of $(X_t, X_{t-1}, \ldots, X_{t-k})$ $(k \geqslant 1)$. The joint density function of (X_t, X_{t-1}) with $\sigma^2 = 1$ is given in Figure 4, and the joint density function of (X_t, X_{t-2}) with $\sigma^2 = 1$ is given in Figure 5.

EXAMPLE 2. Let

$$X_t = \{0.1 + (0.1 - 20.0X_{t-1}) \exp(-X_{t-1}^2)\}X_{t-1} + e_t \tag{3.5}$$

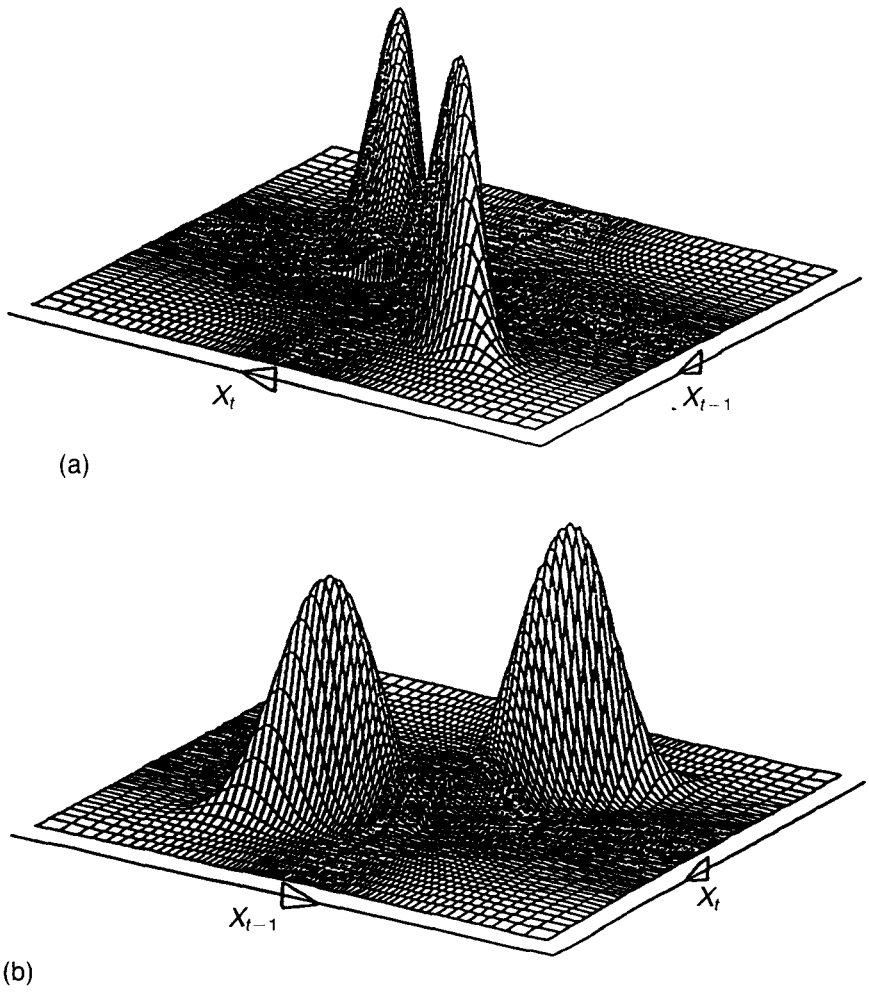


FIGURE 4. The joint density of (X_t, X_{t-1}) of model (3.4) with $\sigma^2 = 1$.

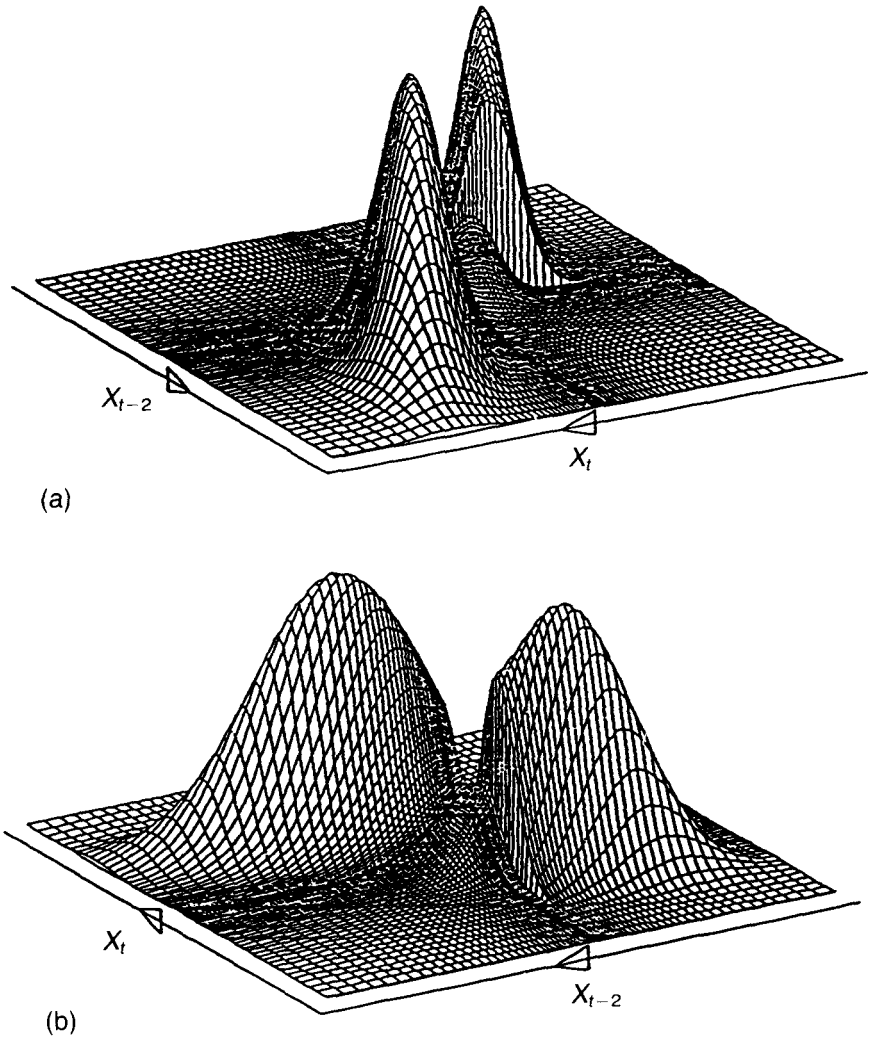


FIGURE 5. The joint density of (X_t, X_{t-2}) of model (3.4) with $\sigma^2 = 1$.

where $e_t \sim N(0, 1)$. The stationary probability density function of this model is given by Figure 6. Let

$$X_t = \lambda(X_{t-1}, X_{t-2}) + e_t \qquad (t = 0, \pm 1, \ldots) \tag{3.6}$$

where $\lambda(.\,,.)$ is a fixed real function of real arguments and $\{e_t\}$ is a sequence of independent identically distributed random variables with probability density function $k(.)$. Denote the conditional density of X_m, given X_j, X_k, by $f(x_m \mid x_j, x_k)$. It can be shown, by arguments similar to those of the first-order case, that the following recursive formula holds:

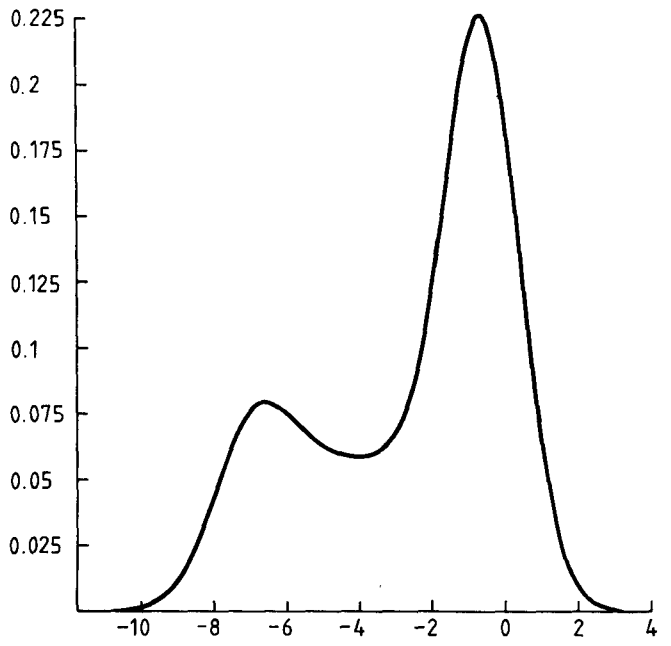


FIGURE 6. The stationary marginal probability density function of model (3.5).

$$f(x_{m+1} \mid x_0, x_{-1}) = \int_{-\infty}^{\infty} f(x_{m+1} \mid x_1, x_0) k(x_1 - \lambda(x_0, x_{-1})) dx_1. \quad (3.7)$$

EXAMPLE 3. As an application of our method for the second-order case, we consider the following SETAR(2; 2, 2) model:

$$X_t = \begin{cases} 0.62 + 1.25X_{t-1} - 0.43X_{t-2} + e_t & \text{if } X_{t-2} \leq 3.25 \\ 2.25 + 1.52X_{t-1} - 1.24X_{t-2} + e_t & \text{if } X_{t-2} > 3.25 \end{cases} \quad (3.8)$$

where $\{e_t\}$ is a sequence of independent identically distributed random variables N(0, 0.0503). Tong (1983, pp. 102–6) has given a rough and ready calculation of $E(X_t \mid X_{t-j})$ $(j = 1, 2, 3, 4)$. The results of our numerical method for the expectation of X_t given $X_{t-j} (j = 1, 2, 3, 4)$ are shown in Figure 7, where they can be compared with Tong's rough calculations. The conditional densities of $(X_t \mid X_{t-j})$ $(j = 1, 2, 3)$ are shown in Figure 8.

4. ACCELERATING BY MATRIX SQUARING

In the above experiments we have found that a judicious choice of the set of grid points x_i may increase the accuracy and improve the convergence of the

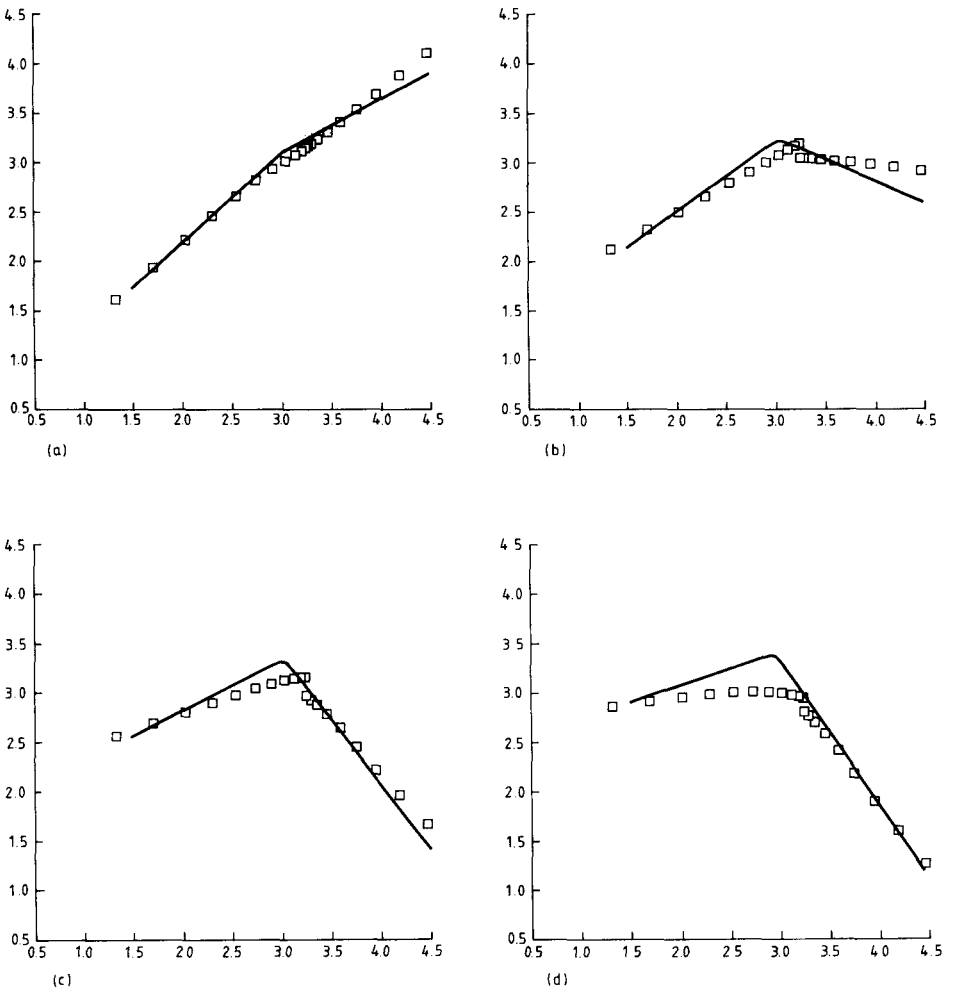
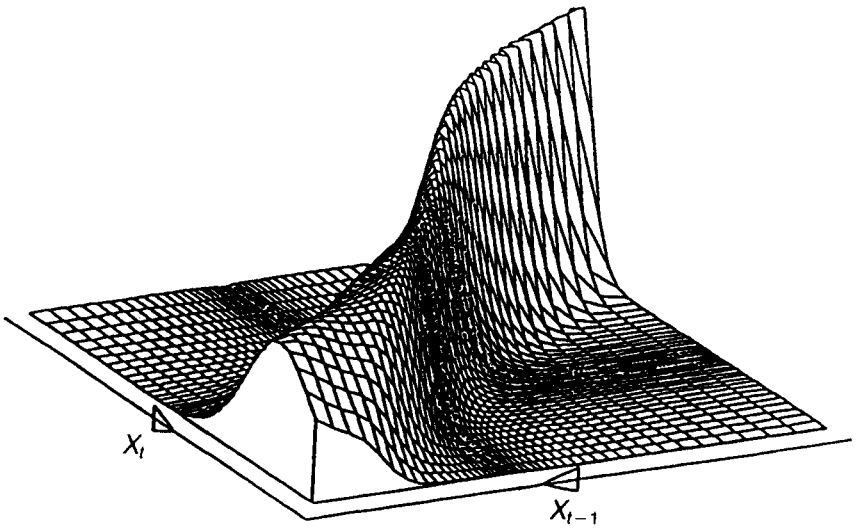


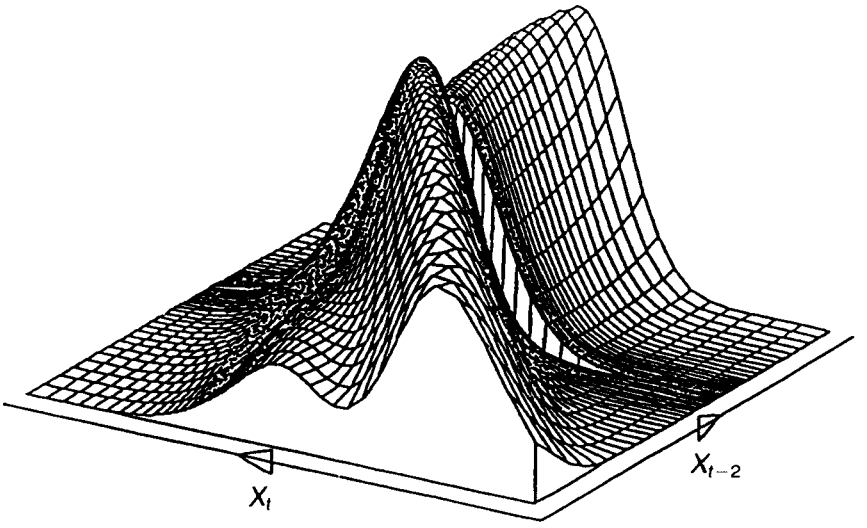
FIGURE 7. The conditional mean $E(X_t \mid X_{t-j})$ of model (3.8) together with Tong's rough calculation (□, approximation values based on the integration method; —, Tong's rough calculation): (a) $j = 1$; (b) $j = 2$; (c) $j = 3$; (d) $j = 4$.

method. Moreover, when the model has a limit point or its autoregressive function is continuous, convergence is achieved after at most 10 steps. When the model has a limit cycle and the variance of e_t is small, the convergence rate is slower. In all our experiments convergence is always achieved.

However, by comparing the numerical integration of (2.3) with the numerical eigenvalue problem, it is clear that we can consider adopting the matrix-squaring method (e.g. Wilkinson, 1965, p. 615) from the latter approach. Specifically, instead of iterating (2.3) over $m = 1, 2, 3, \ldots$, we can calculate $f(x_{2^m} \mid x_0)$ for $m = 0, 1, 2, \ldots$. Intermediate conditional densities



(a)

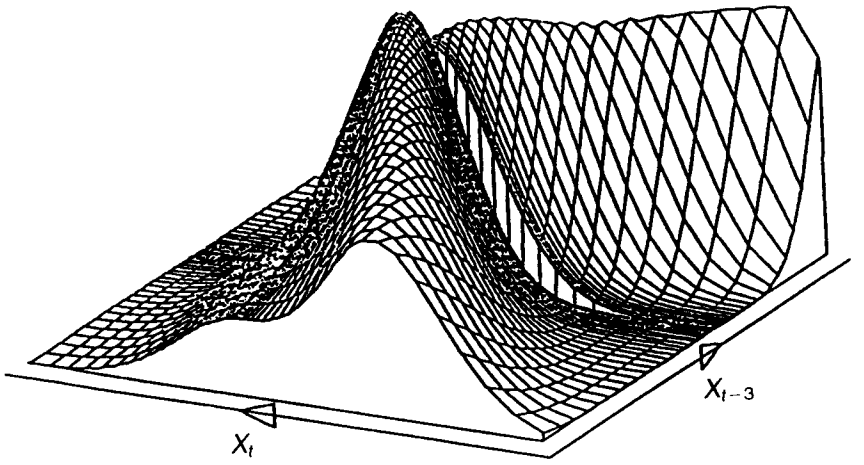


(b)

can be calculated in an obvious economical way. For example, we can calculate $f(x_6 \mid x_0)$ using

$$f(x_6 \mid x_0) = \int_{-\infty}^{\infty} f(x \mid x_2) f(x_2 \mid x_0) dx_2$$

which involves the $f(x_{2^2} \mid x_0)$ and $f(x_{2^1} \mid x_0)$ already calculated. The number of iterations is nominally reduced from m to $\log_2 m$. However, our experience



(c)

FIGURE 8. The conditional density functions of X_t given X_{t-j} of model (3.8): (a) $j = 1$; (b) $j = 2$; (c) $j = 3$.

suggests that whether there is any real reduction in the computing time depends to a large extent on the ratio $|\lambda_1/\lambda_2|$, where $|\lambda_1|$ is the largest absolute value of the eigenvalues and $|\lambda_2|$ is the next largest. Here, the eigenvalues correspond to the 'transition matrix' whose i-jth entry is $f(x_1 \mid x_0)$ with x_0 set equal to ξ_i and x_1 set equal to $\xi_j(\{\xi_1, \xi_2, \ldots, \xi_N\}$ are the points selected for the numerical integration described in Section 3). Our observations are as expected by analogy with the situation in the numerical eigenvalue problem (Wilkinson, 1965). That is, there is a substantial saving in computing time if $|\lambda_1/\lambda_2|$ is close to unity. Otherwise there need not be any saving at all. We illustrate our observations in Tables IV and V. Note that we can rescale each row of the transition matrix so that the rows sum to unity. This can be realized by multiplying the i-jth entry by w_j. In this case, $\lambda_1 = 1$.

Moreover, we have observed the following in the above experiments.

(i) The variance of e_t and the number of points N do not affect the ratio $|\lambda_1/\lambda_2|$ in the linear case. However, in the SETAR case N does not affect the ratio $|\lambda_1/\lambda_2|$ but the variance of e_t does.

(ii) In the linear case, the run times with $\text{var}(e_t) = 0.01$ are approximately the same as those with $\text{var}(e_t) = 1$.

(iii) In the non-linear case the method without matrix squaring does not converge when $\text{var}(e_t)$ is reduced, but the method with matrix squaring converges whether $\text{var}(e_t) = 1$ or $\text{var}(e_t) = 0.01$ and the run times are similar in the latter case.

(iv) In the non-linear case with $\text{var}(e_t) = 0.01$, the ratio $|\lambda_1/\lambda_2|$ is approximately equal to unity.

TABLE IV

NUMERICAL EVALUATION OF THE STATIONARY DENSITY OF THE AR(1) MODEL $X_t = 0.5X_{t-1} + e_t$

N	$\lvert\lambda_1\rvert/\lvert\lambda_2\rvert$	Run time without matrix squaring	Run time with matrix squaring
24	2.0	1.4	1.3
32	2.0	2.0	2.4
64	2.0	4.4	11.9

$e_t \sim N(0, 1)$.

TABLE V

NUMERICAL EVALUATION OF THE STATIONARY DENSITY OF

$$X_t = \begin{cases} 1.5 - 0.9X_{t-1} + e_t & \text{if } X_{t-1} \leqslant 0 \\ -0.4 - 0.6X_{t-1} + e_t & \text{if } X_{t-1} > 0 \end{cases}$$

N	$\lvert\lambda_1\rvert/\lvert\lambda_2\rvert$	Run time without matrix squaring	Run time with matrix squaring
24	1.02	76.9	1.2
32	1.02	85.0	2.5
64	1.02	160.0	20.9

$e_t \sim N(0, 1)$.

ACKNOWLEDGEMENT

We are most grateful to a referee whose comments led us to an investigation reported in Section 4.

REFERENCES

ANDEL, J., NETUKA, I and ZVARA, K (1984) On threshold autoregressive processes. *Kybernetika* 20, 89–160.

CHAN, K. S. and TONG, H. (1986) A note on certain integral equations associated with non-linear time series analysis. *Probab. Theory Relat. Fields* 73, 153–58.

JONES, D. A. (1976) Non-linear autoregressive processes. Ph.D. Thesis, University of London.

—— (1978) Non-linear autoregressive processes. *Proc. R. Soc. London, Ser. A* 360, 71–95.

PEMBERTON, J (1985) Contributions to the theory of non-linear time series models. Ph.D. Thesis, University of Manchester.

TONG, H. (1983) *Threshold Models in Non-linear Time Series Analysis. Lecture Notes in Statistics*, Vol. 21. New York: Springer Verlag.

WILKINSON, J. H. (1965) *The Algebraic Eigenvalue Problem*. London: Oxford University Press.