Qatar Univ, Sci, J. (1997), 17(1) : 27 - 32

A Mixed Markovian Time Series Model

By

B. Y. Thanon & A. M. Khalaf
Department of Mathematics, College of Science
University of Mosul

نموذج ماركوفي ممزوج للسلاسل الزمنية

يدرس نموذج مبسط للسلاسل الزمنية فيه حد التشويش ممزوج (حمي وضربي) ويتم اشتقاق الخصائص الرئيسية للنموذج المقترح كذلك تقترح خوارزمة، تعتمد على المعدل الشرطي والتباين الشرطي، لتخمين معلمات النموذج المقترح. ويطبق هذا النموذج على السلسلة الزمنية للبقع الشمسية. ويبدو من هذه الدراسة ان هذا النموذج يعطي نتائج مشجعة جداً.

ABSTRACT

We study a simple time series model the noise properties of the proposed model are derived. An algorithm based on the conditional mean and the conditional variance, for estimating the parameters of the proposed model is suggested. This model is then applied on the sunspot time series. It seems from this study that the proposed mixed model gives very encouraging results in applications.

Key Words : Conditional mean, Conditional variance, General solution, Stationarity, Sunspot time series.

INTRODUCTION

During the last few years, many classes of time series models, linear and non-linear, have been proposed (see in particular Tong [1]). For simplicity, we consider in this paper the first order case, then, most of these models can be written in the general form in which the noise term is additive.

$$x_t = g(X_{t-1} ; \theta) + Z_t \quad (1)$$

where g is known function, θ is an unknown vector parameter and $[Z_t]$ is a sequence of independent and identically distributed random variables with $E(Z_t) = 0$, $Var(Z_t) = \sigma^2_Z$ and Z_t is independent of X_{t-1}

Let $M_d(x) = E(X_t \setminus X_{t-d} = x)$ be the conditional mean of X_t given $X_{t-d} = x$. The conditional variance of X_t given $X_{t-d} = x$ is defined as :

$$V_d(x) = Var(x_t \setminus X_{t-d} = x) = E(X^2_t \setminus X_{t-d} = x) - M^2_d(x).$$

Then it is clear that for model (1) we have $M_1(x) = g(x;\theta)$ and $V_1(x) = \sigma_Z$.

If the noise term is multiplicated, the model can be written in the form

$$X_t = h(X_{t-1} ; \theta) Z_t , \quad (2)$$

where h is a known positive function. The ARCH model of Engle [2] is a special case of this model with $h(X_{t-1}; \theta) = (\theta_1 + \theta_2 X_{t-1})^{1/2}$

For model (2) we have $M_1(x) = 0$ and

$V_1(x) = h(x; \theta)$

If models (1) and (2) are combined in one model, called a mixed model then the model can be written as

$$X_t = g(X_{t-1}; \theta) + h(K_{t-1}; \theta)Z_t, \ h > 0. \quad (3)$$

For the last model we have

$M_1(x) = g(x; \underset{\sim}{\theta})$ and $V_1(x) = h^2(x; \underset{\sim}{\theta})\, \sigma_Z^2$.

It is observed practically that most real time series have non-constant conditional mean and non-constant conditional variance. Also time series with a changing conditional variance have been useful in many applications (Li and Mak [3]). Hence, model (3) seems to be very reasonable in many applications.

In this paper we consider the following model and we call it a Mixed markovian (MM) model.

$$X_t = a + bX_{t-1} + cX_{t-1} Z_t + Z_t, \quad (4)$$

where a,b and c are unknown constants. This model can be written in (3) form with

$g(x; s) = a+bx$ and $h(x; s) = cx + 1$.

This paper is devoted to study some of the theoretical properties of the MM model (4) as well as their applications.

THEORETICAL PROPERTIES

In this section we give some of the main probabilistic properties of the MM model (4). Details about the complete proofs of the following theorems are given in the appendix as well as in the M.Sc. thesis of the second author [4].

THEOREM 1 (MOMENTS) :

Let $\{X_t\}$ be generated by the MM model then the mean, variance, autocovariance function, autocorrelation function and the normalized spectral density function of X_t are, respectively given by

i. $E[X_t] = \mu_X = a/(1-b)$; $|b| < 1$

ii. $Var(X_t) = \sigma_X^2 = \gamma_o$

$$= \frac{(a^2c^2 + 2ac + 1 + b^2 - 2b - 2abc)\,\sigma_Z^2}{(1-b)^2(1 - b^2 - c^2\sigma_Z^2)} \ ; |b| < 1$$

iii. $Cov(X_t, X_{t+k}) = \gamma_k = b^{|k|}\gamma_o$; $k = 0, \pm1, \pm2, \ldots$

iv. $\varrho_k = b^{|k|}$; $k=0, \pm1, \pm2, \ldots$

v. $f(w) = \dfrac{1-b^2}{2\pi(1-2b\cos w + b^2)}$, $-\pi \quad w_z \quad \pi$

Note :

From (ii) we note that the condition for $Var(X_t) > 0$ is that $b^2 + c^2\sigma_z^2 < 1$. Also it is easy to show that

$$E[X_t^3] = \frac{a^3 + 3ab^2\mu_2 + 3ac^2\mu_2\sigma_Z^2 + abc\,\mu_2\sigma_Z^2}{[1 - b(b^2 + 3c\,\sigma_Z^2)^2]}$$

$$\frac{3ab^2\mu_1 + 4ac\mu_1^2\sigma_Z + 3b\mu_1^2\sigma_Z}{} + \frac{c\mu_1\sigma_Z + 3a\sigma_Z}{[1-b(b^2 + 3c\,\sigma_Z^2)^2]}$$

where $\mu_1 = \mu_X = a/(1-b)$ and

$\mu_2 = (X_t^2) = [a^2 + \sigma_Z^2 + (2ab + 2c\sigma_Z^2)] / (1 - b^2 - c^2\sigma_Z^2)$.

Hence the condition of existence of the third order central moment of this process is that μ_2 exists and $b(b^2 + 3c^2\sigma_Z^2) \neq 1$.

THEOREM 2 (GENERAL SOLUTION)

The general solution of the MM model is given by

$$X_t = \sum_{i=1}^{t} (a + z_i) \prod_{j=i+1}^{t} (b + cZ_j) + \prod_{i=1}^{t} (b + cZ_i) X_O.$$

THEOREM 3 (STATIONARITY)

The necessary and sufficient conditions for the process $\{X_t\}$ generated by the MM model to be second order (weakly) stationary are that

$b^2 + c^2\sigma_Z^2 < 1$ and $|b| < 1$.

THEOREM 4 (EXISTENCE)

Assume that Z_t is almost surely (a.s) not a linear function of Z_t. If $|b| < 1$ and $b^2 + c^2\sigma_z^2 < 1$ then there exists a unique strictly stationary process $\{X_t\}$, satisfying (4), and given by

$$X_t = a + Z_t + \sum_{j=1} [\prod_{R=1} (b + cZ_{t-k})] \; Z_{t-j}$$

with the infinite series being almost surely convergent.

THEOREM 5 (PROBABILITY DISTRIBUTION)

Let $\{X_t\}$ be generated by the MM model and assume that X_t has a probability density function (p.d.f) $f_{x_t}(x)$ for all values of t=1,2, ... Let Z_t has a p.d.f $f_{z_t}(z)$ for all values of t=1,2, ... Then $f_{x_t}(x)$ satisfies the following integral equation

$$f_{x_t}(x) = \int \frac{1 + cy}{ac - cx - b} \; f_{z_t}(z) \, f_{x_{t-1}}(y) \, dz$$

Note :

From **THEOREM 5** we note that the p.d.f of X_t generated by the MM model has a singularity point at $x = a - b/c$ i.e.

$$f_{x_t}(x) \to \infty \text{ as } x \to a - b/c.$$

THEOREM 6 (PREDICTOR)

Let X_t generated by the MM model than the kth-step ahead predictor of X_t , $\hat{X}_{t+k}$, is given by

$$\hat{X}_{t+k} = \frac{a(1 - b^k)}{(1 - b)} + b^k X_t \qquad ; \qquad k = 1,2, ...$$

with $\hat{X}_t = X_t$. The mean square error of prediction is given by $\sigma_k^2 = E[(X_{t+k} - \hat{X}_{t+k})^2]$

$$= \frac{(a^2c^2 + 2ac^2 + 1 + b - 2b - 2abc)\,(1 - b^{2k})\,\sigma_z^2}{(1 - b)^2\,(1 - b^2 - c^2\sigma_z^2)}$$

ESTIMATION OF PARAMETERS

Let $\{X_t \,;\, t= 1,2, ..., n\}$ be a realization from a time series $\{X_t\}$. Our problem now is to fit a MM model to the data. Because the probability distribution of $\{X_t\}$ is unknown, we cannot use the usual maximum likelihood method to obtain estimates of model parameters a, b, c and σ_z^2. Also, because the noise term in the model is non-additive, the classical least square method cannot be used in estimating in the unknown parameters. In this paper, we describe a new algorithm to obtain estimates of model parameters. This algorithm is based on the use of the conditional mean and the conditional variance.

Let $\{X_t\}$ be a completely stationary time series with a joint probability density function of X_{t-d} and X_t, $f_{t-d}X_t(x,y)$, d is a positive integer.

First we consider the following kernel estimate of $E(X_t^j \backslash X_{t-d} = x)$ (Thanoon, [5]).

$$\hat{E}[X_t^j \backslash X_{t-d} = x] = \frac{\sum_{i=1}^{n} x^j K\{(x-x_i)/S\}}{\sum_{i=1}^{n} K\{(x-x_i)/S\}} \quad ; j=1,2,...$$

where K(.) is a kernel (window) function which is a non-negative function on R with $\int_R K(u)\,du = 1$. In this paper we use the well-known Bartletts (triangle) window which is defined by $K(u) = 1 - |u|$; for $|u|$ 1 and $K(u) = 0$; otherwise s is a smoothing parameter. Here the standard deviation of the data is used as the smoothing parameter. Then the $M_d(x)$ and $V_d(x)$ can be estimated respectively as follows :

$$\hat{M}_d(x) = \hat{E}\{X_t \backslash X_{t-d} = x\} \text{ and}$$

$$\hat{V}(x) = \hat{E}\{X_t^2 \backslash X_{t-d} = x\} - \hat{M}_d^2(x)$$

Consider first the MM model and note that $M_1(x) = a + bx$ and $V_1(x) = (cx + 1)\,\sigma_z^2$. Hence, if we assume that $Y_t = X_t / \sigma_z$ (i.e. by dividing the R.H.S. of the MM model by σ_z^2), then we get

$Var(Y_t \setminus X_{t-1} = x) = (cx + 1)$. Therefore, we can assume without loss of generality that the noise variance is equal to unity.

We can estimate $\hat{a}$, $\hat{b}$ and $\hat{c}$ by setting $M_1(x) = \hat{M}_1(x)$ and $V_1(x) = \hat{V}_1(x)$ where $M_1(x) = a + bx$, $V_1(x) = (cx + 1)$ and $M_1(x)$, $V_1(x)$ are the kernel estimates of $M_1(x)$, $V_1(x)$; respectively. By solving these two equations, at some given values of x, we can obtain estimates of a, b and c, respectively $\hat{a}$, $\hat{b}$, $\hat{c}$.

The noise variance can then be estimated in the usual way

$$\hat{\sigma}_Z^2 = \sum_{t=2}^{n} (X_t - \hat{a} - \hat{b}X_{t-1})^2 / (n - 3).$$

AN APPLICATION (SUNSPOT SERIES 1700 - 1920)

Let $\{X_t\}$ denote the annual mean of Wolf's sunspot numbers for the year 1699+t = 1,2, ... , 221. The following SETAR (2;4, 12) model was fitted to the data by Tong and Lim [6] :

$$X_t = \begin{cases} 10.54 + 1.69 X_{t-1} - 1.16 X_{t-3} + 0.15 X_{t-4} + Z_t & X_{t-3} \quad 36.6 \\ 7.80 + 0.74 X_{t-1} - 0.04 X_{t-2} - 0.02 X_{t-3} + o.17 X_{t-4} - 0.23 X_{t-5} + 0.02 X_{t-6} + 0.16 X_{t-7} - 0.26 X_{t-8} + 0.32 X_{t-9} - 0.39 X_{t-10} + 0.43 X_{t-11} - 0.04 X_{t-12} + Z_t, & X_{t-3} > 36.6 \end{cases}$$

where $\hat{\sigma}_Z^2 = 254.6$ and $\hat{\sigma}_Z^2 = 66.8$

Gabr and Subba Rao [7] have fitted the following subset bi linear, SBL, model to this data set

$$X_t - 1.5012 X_{t-1} + 0.767 X_{t-9} - 6.8860 = - 0.0146 X_{t-2}Z_{t-1} + 0.0063 X_{t-8}Z_{t-1} - 0.0072 X_{t-4}Z_{t-3} + 0.0061 X_{t-4}Z_{t-3} + 0.0036 X_{t-1}Z_{t-s} + 0.0043 X_{t-2}Z_{t-4} + 0.0018X_{t-3}Z_{t-z} + Z_t,$$

where $Var(Z_t) = 143.33$

Thanoon and Sofia [8] suggested a threshold-bilinear TBL, model and fitted their model to the same data. The fitted model takes the form

$$X_t = \begin{cases} 2.8357 + 1.9767 X_{t-1} - 1.3805 X_{t-2} + 0.0945 X_{t-3} - 0.1178 X_{t-4} + 0.3654 X_{t-5} - 0.0032 X_{t-1}Z_{t-1} - 0.0593 X_{t-2}Z_{t-1} - 0.0776 X_{t-3}Z_{t-1} - 0.0502 X_{t-4}Z_{t-1} + 0.0225 X_{t-5}Z_{t-1} + Z_t & X_{t-3} \quad 36.6 \\ 7.8 + 0.74 X_{t-1} - 0.20 X_{t-3} + 0.17 X_{t-4} - 0.23 X_{t-5} + 0.02 X_{t-6} + 0.16 X_{t-7} - 0.26 X_{t-8} + 0.32 X_{t-9} - 0.39 X_{t-10} + 0.43 X_{t-11} - 0.04 X_{t-12} + Z_t & \end{cases}$$

where $Var(Z_t) = 98.2$

Applying the suggested algorithm in the last section, the following MM model is identified

$$X_t = 7.6145 + 0.7932 X_{t-1} + 0.3243 X_{t-1}Z_t + Z_t,$$

where $Var(Z_t) = 5.1949$. The following table given a comparison between these three models.

Table (1)

A comparison between the fitted models for sunspot series 1700 - 1920

Model	No. of Par.	Res. Var	NAIC
SETAR	18	153.70	5. 197
SBL	11	143.33	5.065
TBL	24	98.20	4.804
MM	3	5.19	1.675

* NAIC = [n in (Residual variance) + 2k] n, where n is the number of (effective) data and k is the number of parameters.

CONCLUSIONS AND DISCUSSIONS

The studied model in this paper gave very encouraging theoretical and applied properties, one of the main interesting theoretical properties is that their conditional mean and conditional variance are both non-constants.

The residual variance and the NAIC value obtained from the fitted MM model are much lower than those of all other fitted models to the same time series (see Figure (1)). The reduction in the residual variance from using the MM model rather than the SETAR, SEL and TBL models is 2859% and 1790% ; respectively. The same conclusion is drawn with other data sets (see the M.Sc. thesis of the second auther [4]). Hence we suggest further studies in the direction of these models. In particular, we suggest to study the generalization of the studied model in this paper which has the general form

$$X_t = a_\circ + a_1 X_{t-1} + \ldots + a_q X_{t-q} + (b_\circ + b_1 X_{t-1} + \ldots b_r X_{t-r}) Z_t,$$

where $b_\circ = 1$

APPENDIX
PROOFS AND THEOREMS

THEOREM 1

Let X_t be second order stationary than

i. $E[X_t] = \mu_x = a + b\mu_x + 0 + 0$

i.e. $\mu_x (1 - b) = a \quad \Rightarrow \mu_x = a / (1-b),$

ii. Squaring (4) and taking the expectation, we get

$$E[X_t^2] = a^2 + a^2b^2 / (1-b) + b^2\sigma_x^2 + a^2c^2 / (1-b)\,\sigma_x^2 + c^2\sigma_x^2\sigma_z^2 + \sigma_z^2 + 2a^2b / (1-b) + 2ac / (1-b)\,\sigma_x^2$$

Then $\sigma_x^2 = E(X_t^2) - \mu_x^2$.

iii. Similar to ii.

iv. $\rho_x = \gamma_k / \gamma_\circ$

v. $f(\omega) \; \frac{1}{2\pi} \left[1 + 2\sum_{k=1}^{\infty} \rho_k \cos(\omega k) \right]$

$$= \frac{1}{2\pi} \left[1 + 2\sum_{k=1}^{\infty} b_k \cos(\omega k) \right]$$

$$= \frac{1}{2\pi} \left[1 + 2\sum_{k=1}^{\infty} (be^{i\omega})^k \right]$$

$$= \frac{1}{2\pi} \left[1 + 2\,R\left(\frac{be^{i\omega}}{1 - be^{i\omega}}\right) \right]$$

$$f(\omega) = \frac{1 - 2b^2}{2\,(1 - 2b\cos\omega + b^2)} \qquad ; - \qquad \omega$$

THEOREM 2

Rewrite (4) in the form

$X_t = a + (b + cZ_t)\, X_{t-1} + Z_t$

Now $X_t = a + (b + cZ_1)\, X_\circ + Z_1$

$X_2 = a + (b + cZ_2)\, X_1 + Z_2$

$= a + (b + cZ_2)\, [\, a + (b + cZ_1)\; X_\circ + Z_1] + Z_2$

$= a + a(b+cZ_2) + (b+cZ_1)(b+cZ_2)\, X_\circ + (b+ cZ_2)Z_1 + Z_2$

in general we get

$$X_t = \sum_{i=1}^{t} a \prod_{j=i+1}^{t} (b+cZ_j) + \prod_{t=1}^{t} (b+cZ_i)\, X_\circ + \sum_{i=1}^{t} Z_i \prod_{j=i+1}^{t} (b+cZ_j)$$

And hence

$$X_t = \sum_{i=1}^{t} (a+Z_i) \prod_{j=i+1}^{t} (b+cZ_j) + \prod_{i=1}^{t} (b+cZ_i)\, X_\circ$$

THEOREM 3

The first condition that $b^2 + c^2\sigma_z^2 < 1$ is necessary for the variance to be positive (**THEOREM 1**). The second condition that $|b| < 1$ can be obtained also from **THEOREM 2.** Since

$$\prod_{j=i+1}^{t} (b+cZ_i) = (b+cZ_{i+1})(b+cZ_{i+2}) \ldots (b+cZ_t)$$
$$= b^{t-i-1} + \text{other terms}$$

and

$$\prod_{t=1}^{t} (b+cZ_i) = (b+cZ_1)(b+cZ_2) \ldots (b+cZ_t)$$
$$= b^t + \text{other terms}$$

Hence, for large t, b^{t-i-1} and b^t are covergent only if $|b| < 1$.

THEOREM 4

The representation can be obtained in a similar way to THEOREM 2. The covergence can be proved by using Jensen's inequality and the strong law of large numbers.

If we rewrite (4) in the form

$$X_t = a + bX_{t-1} + (i + cX_{t-1})\, Z_t$$

The distribution function of X_t is defined as

$$F_{x_t}(x) = p\,(X_t \quad x)$$
$$= p\,[a + bX_{t-1} + (cX_{t-1} + 1)Z_t \quad x]$$
$$= \int_y p\,[X_t \quad x \setminus X_{t-1} = y)\, f_{x_{t-1}}(y)\, dy$$
$$= \int_y p\,[a + by + (1 + cy)Z_t \quad x]\ f_{x_{t-1}}(y)\, dy$$
$$= \int_y p\,(Z_t \quad \frac{x - a - by}{1 + cy}\ \ f_{x_{t-1}}(y)\, dy$$

$$f_{x_t}(x) = dF_{x_t}(x)/dx = d/dx\ \ f_{z_t}{-1}\left(\frac{x - a - by}{1 + cy}\right) f_{x_{t-1}}(y)\, dy$$

Let $z = \dfrac{x - a - by}{1 + cy}$

hence

$$f_{x_t}(x) = \int y\ f_{z_t}(z)\ \frac{1}{1 + cy}\quad f_{x_{t-1}}(y)\ \frac{(1 + cy)}{ac\text{-}cx\text{-}b}\quad dz$$

$$f_{x_t}(x) = \int y\ \frac{1 + cy}{ac\text{-}cx\text{-}b}\quad f_{z_t}(z)\ f_{x_{t-1}}(y)\quad dz \qquad 2$$

THEOREM 6

It is well known that kth step - ahead predictor of X_t, $\hat{X}_{t+k}$ is the conditional expectation of X_{t+k} given $X_t, X_{t-1}, \ldots$

i.e. $\hat{X}_{t+k} = E\,[\,X_{t+k} \setminus X_t, X_{t-1}, \ldots\,]$.

Now $\hat{X}_{t+k} = E\,[\,X_{t+k} \setminus X_t, X_{t-1}, \ldots\,]$

$$= E\,[(a + bX_t + cX_tZ_t + Z_{t+1})/\ X_t, X_{t-1}, \ldots]$$
$$= a + bX_t$$

$$\hat{X}_{t+z} = E\,[\,X_{t+z} \setminus X_t, X_{t-1}, \ldots\,]$$
$$= E\,[(a + bX_{t+1} + cX_{t+1}Z_{t+2} + Z_{t+2} \setminus X_t, X_{t-1}, \ldots]$$
$$= a + bE\,[X_{t+1} \setminus X_t, X_{t-1}, \ldots]$$
$$= a + b\hat{X}_{t+1}$$

Hence, in general we have

$$\hat{X}_{t+k} = a + b\,\hat{X}_{t+k-1} \qquad ; k = 1, 2, \ldots$$

By successive substitution in the last equation we get :

$$\hat{X}_{t+k} = a + ab + ab^2 + \ldots + b^kX_t$$

which can be written in the form

$$\hat{X}_{t+k} = \frac{a(1 - b^k)}{(1 - b)} + b^kX_t \qquad ; k = 1, 2, \ldots$$

REFERENCES

[1] **Tong, H. 1990.** Non-linear Time Sreies : A Dynamical System Approach. London : Oxford University Press.

[2] **Engle, R. F. 1982.** Autoregressive Conditional Heteroscdasticity with Estimates of Variance of U. K. Inflation. Econometrica, 50 : 987 - 1008.

[3] **Li, W. K. and T. K. Mak 1994.** On the squared Residual Autocorrelation in Non-Linear Time Series with Conditional Heteroskedasticity. J. Time Ser. Anal., 15 : 627 - 636.

[4] **Khalaf, A. M. 1996.** A Study of a Simple Multiplicative Markovian Model with Applications. M. Sc. Thesis, Department of Mathematics, College of Science, University of Mosul, Iraq.

[5] **Thanoon, B. Y. 1996.** Kernel Estimation : A Graphical Approach in Statistical Estimation. Raf. Jour. Sci. (to appear).

[6] **Tong, H. and K. S. Lim 1980.** Threshold Autoregresion, Limit Cycles and Cyclic Data. J. Roy. Stat. Soc. B, 42 : 245 - 293.

[7] **Gabr, M. M. and T. Subba Rao 1981.** The Estimation and Prediction of Subest Bilinear Time Series Models with Application J. Time Ser. Anal., 2 : 155 - 171.

[8] **Thanoon, B. Y. and E. B. Sofia 1995.** Fitting the Sunspot Series by a Threshold Model. DIRASAT, Vol. 22, No. 1 : 67 - 82.

Figure (1) A comparison between the fitted models for sunspot series (1700 - 1920).