

# A Note on the Invertibility of Nonlinear ARMA models

Kung-Sik Chan

Department of Statistics & Actuarial Science  
University of Iowa, Iowa City, IA 52242, U.S.A.

Email: kung-sik-chan@uiowa.edu

Howell Tong

Department of Statistics  
London School of Economics, London WC2A 2AE, U.K.

Email: howell.tong@gmail.com

April, 2009

## Abstract

We review the concepts of local and global invertibility for a nonlinear auto-regressive moving-average (NLARMA) model. Under very general conditions, a local invertibility analysis of a NLARMA model admits the generic dichotomy that the innovation reconstruction errors either diminish geometrically fast or grow geometrically fast. We derive a simple sufficient condition for a NLARMA model to be locally invertible. The invertibility of the polynomial MA models is revisited. Moreover, we show that the Threshold MA models may be globally invertible even though some component MA models are non-invertible. One novelty of our approach is its cross-fertilization with dynamical systems.

**Keywords:** Attractor, Dynamical system, Nonlinear time series, Polynomial MA model, Subadditive ergodic theory, Threshold MA model.

# 1 Introduction

Despite the growing literature on nonlinear time series analysis (Priestley, 1988, Tong, 1990, Franses and van Dijk, 2000, Chan and Tong, 2001, Fan and Yao, 2003, Small, 2005, and Gao, 2007), the general framework makes use of nonlinear autoregressive models. In contrast, nonlinear moving-average (NLMA) models are relatively under-explored. Part of the problem contributing to the slow development, both empirical and theoretical, on NLMA models is due to the difficulty in establishing the invertibility of a NLMA model.

Here, we focus on nonlinear autoregressive moving-average (NLARMA) models, and discuss two concepts of invertibility for these models. We illustrate these concepts with the polynomial MA models and the threshold MA models.

## 2 Nonlinear Auto-Regressive Moving-average Models

The linear moving-average (MA) model of order  $q$  is characterized by the feature that it has memory of  $q$  lags. Recall that an MA( $q$ ) process  $\{Y_t\}$  is defined by the equation:

$$Y_t = \mu + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i},$$

where  $\mu$  is the mean of  $Y_t$  and the innovations  $\{\varepsilon_t\}$  are white noise, i.e. uncorrelated random variables of zero mean and finite (identical) variance  $\sigma^2$ . For simplicity, we shall assume  $\mu = 0$ . It is well known that the autocorrelation function (ACF) of an MA( $q$ ) process has a cut-off after lag  $q$ , i.e.  $\text{corr}(Y_t, Y_{t-\ell}) = 0$  for  $\ell > q$ . In other words, the process is uncorrelated with its  $q+1$ -th or higher lags. If the innovations are, furthermore, jointly independent, then the process is independent of its  $q+1$ -th or higher lags, in which case the process is said to have finite memory. The natural question of developing nonlinear models with finite memory has received some attention in the literature.

Here, we review some nonlinear time-series models of finite memory. Henceforth, we shall assume that  $\{\varepsilon_t\}$  is an independent and identically distributed sequence of random variables with zero mean and finite variance. To begin with,

we note that any model of the following form is of finite memory:

$$Y_t = \varepsilon_t + h(\varepsilon_{t-1}, \dots, \varepsilon_{t-q}; \theta), \quad (1)$$

where  $h(\cdot; \theta)$  is a known function for known parameter  $\theta$ . The linear MA( $q$ ) model is obtained by setting  $h$  to be a linear function. Similarly, a nonlinear finite-memory model, also known as a nonlinear moving-average (NLMA) model, can be obtained by setting  $h$  to be some parametric nonlinear functions. Clearly, any nonlinear moving-average model is stationary. However, as in the case of linear moving-average models, the issue of invertibility is pivotal. Invertibility refers to the feasibility of reconstructing the innovations from the observations, assuming that the true model is known. Given the parameter  $\theta$ , Eqn. (1) can be inverted to define the residuals

$$\hat{\varepsilon}_t = Y_t - h(\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-q}; \theta), \quad (2)$$

where the initial values are generally set as  $\hat{\varepsilon}_{1-k} = 0$ , the mean of the innovations, for  $k = 1, \dots, q$ . The polynomial moving-average model (Robinson, 1977) is obtained by letting  $h$  be a polynomial. For example,

$$Y_t = \varepsilon_t + \beta \varepsilon_{t-1}^2$$

is a simple quadratic MA(1) model. However, it has been noted that polynomial MA models are generally non-invertible (e.g. Granger and Andersen, 1978b), which makes them not suitable for prediction purpose, and also makes it hard to carry out model diagnostics. We shall elaborate on the concept of invertibility in the following sections.

Several interesting mixed nonlinear ARMA (in short NLARMA) models that may be invertible have been proposed in the literature. An NLARMA( $p, q$ ) model is defined by a stochastic difference equation of the following form:

$$Y_t = \varepsilon_t + h(Y_{t-1}, \dots, Y_{t-p}, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}; \theta), \quad (3)$$

A sub-class of the NLARMA models belongs to the family of bilinear models (Granger and Andersen, 1978a, Subba Rao, 1981, Guegan and Pham Dinh, 1987 Priestley, 1988, Tong, 1990); they are linear in both past lags of the process and past lags of the innovation, e.g.

$$Y_t = \varepsilon_t + \theta Y_{t-1} \varepsilon_{t-1}$$

is a simple bilinear model where  $\theta$  is a parameter. Some results on the invertibility of sub-classes of bilinear models have been derived and surveyed in the aforementioned works.

Recently, Ling and Tong (2005) re-visited the Threshold MA (TMA) model, which specifies that the process switches from one linear MA model to another linear MA model whenever some lag of the process exceeds one of the threshold values. A simple TMA model of order one and with one threshold takes the following form:

$$Y_t = \begin{cases} \varepsilon_t - \theta_{1,1}\varepsilon_{t-1} & \text{if } Y_{t-d} \leq r \\ \varepsilon_t - \theta_{2,1}\varepsilon_{t-1} & \text{otherwise,} \end{cases} \quad (4)$$

where the  $\theta$ 's and  $r$  are parameters;  $d$  is a positive integer parameter known as the delay parameter. A more general TMA model will be considered later. It is straightforward to generalize the TMA model to Threshold ARMA (TARMA) model by replacing the linear MA sub-models to linear ARMA sub-models. Ling and Tong (2005) gave some sufficient conditions for the invertibility of a TMA model of order one and with multiple thresholds. For example, they considered the following model:

$$Y_t = \{\phi_0 + \sum_{j=1}^k \psi_j I(r_{j-1} < Y_{t-1} \leq r_j)\} \varepsilon_{t-1} + \varepsilon_t, \quad (5)$$

where  $I(A)$  is the indicator function of the event  $A$ , and  $-\infty = r_0 < r_1 < \dots < r_k = \infty$  are the  $k$  thresholds. Let  $F_y(\cdot)$  denote the cumulative distribution function of  $Y$ . They established the following theorem, which gives an almost necessary and sufficient condition:

**THEOREM 1.**  *$\{Y_t\}$  is invertible if  $\prod_{j=1}^k \{|\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})}\} < 1$  and is not invertible if  $\prod_{j=1}^k \{|\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})}\} > 1$ .*

Note that the case with  $\prod_{j=1}^k \{|\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})}\} = 1$  is undecided but they conjectured non-invertibility. Note also that the MA coefficients of intermediate linear MA sub-models also feature in the invertibility condition. However, for TMA models of higher order, they were only able to give some rather restrictive sufficient conditions.

### 3 Global and Local Invertibility

In this section, we elaborate on the concept of invertibility for NLARMA models. For conciseness, we focus on the case of an NLMA model defined by (1) for which the innovations may be estimated by the residuals defined by (2), but note that all results in this section and the next can be extended to the case of NLARMA models. On the other hand, the innovations satisfy a similar difference equation:

$$\varepsilon_t = Y_t - h(\varepsilon_{t-1}, \dots, \varepsilon_{t-q}; \theta),$$

so that the reconstruction errors  $W_t = \hat{\varepsilon}_t - \varepsilon_t$  satisfy the equation

$$W_t = h(\varepsilon_{t-1}, \dots, \varepsilon_{t-q}; \theta) - h(W_{t-1} + \varepsilon_{t-1}, \dots, W_{t-q} + \varepsilon_{t-q}; \theta), \quad (6)$$

which is generally a random-coefficient stochastic difference equation for  $\{W_t\}$ . Invertibility requires that the reconstruction errors  $\{W_t\}$  approach zero in some sense, e.g., in probability. Conditions for invertibility are then simply conditions for the solutions of the difference equation (6) to approach 0 as  $t \rightarrow \infty$ , in probability. For linear MA models, the necessary and sufficient condition for invertibility is well known. Let

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}.$$

Then

$$W_t - \theta_1 W_{t-1} - \dots - \theta_q W_{t-q} = 0,$$

in which case the condition of invertibility is that all roots of the characteristic equation

$$1 - \theta_1 x - \dots - \theta_q x^q = 0$$

lie outside the unit circle; see, e.g., Box, Jenkins and Reinsel (1994) and Cryer and Chan (2008).

However, for the NLMA models, general conditions for invertibility seem difficult to obtain, as it is generally difficult to derive necessary and sufficient conditions for zero to be a global attractor for all solutions of (6). Before proceeding further, we note that we can vectorize (6) into a first-order vector equation for the case that  $q > 1$ . Let  $\mathbf{W}_t = (W_t, W_{t-1}, \dots, W_{t-q+1})^T$ , and  $\boldsymbol{\varepsilon}_t = (\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1})^T$ .

Then

$$\begin{aligned}\mathbf{W}_t &= F(\mathbf{W}_{t-1}; \boldsymbol{\varepsilon}_{t-1}, \theta) \\ &= (h(\varepsilon_{t-1}, \dots, \varepsilon_{t-q}; \theta) - h(W_{t-1} + \varepsilon_{t-1}, \dots, W_{t-q} + \varepsilon_{t-q}; \theta), W_{t-1}, \dots, W_{t-q+1})^T.\end{aligned}\tag{7}$$

(For an NLARMA model defined by (3),  $F$  is a function of  $(\mathbf{W}_{t-1}; Y_{t-1}, \dots, Y_{t-p}, \boldsymbol{\varepsilon}_{t-1}, \theta)$ .)

Clearly  $0 = F(0; \boldsymbol{\varepsilon}, \theta)$  for all  $\boldsymbol{\varepsilon}$  so that the origin is an equilibrium point (for the dynamical model defined by (7)). Then, invertibility is equivalent to the origin  $0 \in R^q$  being an asymptotically global attractor, in probability. It is often hard to study the global nature of the origin. Hence, a weaker form of invertibility has been studied in the literature (e.g. Granger and Anderson, 1978b) which requires the origin to be locally and asymptotically stable. In other words, local invertibility concerns whether the innovations can be asymptotically recovered if the initial conditions are approximately correct. Local invertibility can be assessed by linearizing  $F$  around the origin. Let  $\dot{F} = \frac{\partial F}{\partial \mathbf{W}}$  evaluated at  $\mathbf{W} = 0$  and  $\boldsymbol{\varepsilon}_{t-1}$  be concisely denoted as  $F_t$ . Then, the local asymptotic stability of the origin can be inferred from the stability of the origin for the random-coefficient linear stochastic difference equation (c.f. Grobman and Hartman Theorem on p. 237 of Chan and Tong, 2001):

$$\mathbf{W}_t = F_t \mathbf{W}_{t-1}.\tag{8}$$

We now illustrate these concepts with the simple quadratic MA(1) model:

$$Y_t = \varepsilon_t - \beta \varepsilon_{t-1}^2,$$

where  $\beta \neq 0$ . It is straightforward to show that the reconstruction errors  $W_t$  satisfy the stochastic equation:

$$W_t = W_{t-1}(\beta W_{t-1} + 2\beta \varepsilon_{t-1}).\tag{9}$$

It can be shown that  $W_t$  diverges to infinity with positive probability if  $|W_0| > w_0$  where  $w_0 = 2/|\beta|$  with positive probability. Subject to the latter condition, the claim of transience of  $\{\mathbf{W}_t\}$  can be justified as follows. Let  $\gamma = E|\varepsilon_t|$ . Markov inequality implies that the event  $|W_t| \geq 2^t w_0$  for all  $t$  occurs with probability not smaller than  $Prob(|W_0| > w_0) \times \prod_{t=1}^{\infty} (1 - \frac{2\gamma|\beta|}{|\beta|2^t w_0 - 2})$ , which is positive. The

preceding argument can be adapted to show that polynomial MA models of degree greater than 1 are non-invertible under similar conditions, along similar lines as employed in Chan and Tong (1994, Theorem 2) for the transience of polynomial AR models. It should be noted that Granger and Anderson (1978b) claimed that ‘a set of models which is not invertible for any non-zero values of its parameters consists of non-linear moving averages’. However, we are unaware of any rigorous proof of the claim to-date. Later, we shall state, in Theorem 2, precise conditions under which a NLMA model is locally invertible/non-invertible with rigorous proof. For now, the preceding analysis reveals that the non-invertibility of the polynomial MA model is associated with an unbounded support for the initial reconstruction error. Below, we conduct a local analysis that shows that the polynomial MA model may be invertible for the case of innovations with sufficiently small support around the origin.

Next, we illustrate the concept of local invertibility with the preceding simple quadratic MA(1) model. Linearizing (9) around  $W_t = 0$  yields the following linear model:

$$W_t = 2\beta\varepsilon_{t-1}W_{t-1},$$

where the coefficient is random and equals  $2\beta\varepsilon_{t-1}$ . The solution of the equation is trivial, it being

$$W_t = W_0(2\beta)^t \prod_{s=1}^t \varepsilon_{s-1}.$$

In particular

$$|W_t| = |W_0| \exp\left(n \sum_{s=1}^n \{\ln |(2\beta| + \ln |\varepsilon_{s-1}|)\}/n\right).$$

Hence if  $\ln |2\beta| + E \ln |\varepsilon| < 0$ , then the law of large numbers implies that the origin is asymptotically stable so that the model is locally invertible. If  $\ln |2\beta| + E \ln |\varepsilon| > 0$ , then the origin is locally unstable so that the model is not locally invertible. The case  $\ln |2\beta| + E \ln |\varepsilon| = 0$  is delicate and requires further analysis that will not be pursued here. This example shows that the conclusions concerning invertibility from a local analysis and a global analysis can differ. Furthermore, the global non-invertibility is predicated on the condition that the initial reconstruction errors can be arbitrarily large, with positive probability. On the other hand, the local

analysis suggests that if the errors are of sufficiently small bounded support, then the model can be invertible if the initial conditions respect the bounded support condition for the innovations.

## 4 Dichotomy of Local Invertibility Analysis

Recall that a linear MA( $q$ ) model is invertible if and only if (or iff for short) all the roots of the characteristic equation lie outside the unit circle. This result follows from a stability analysis of the difference equation for the reconstruction errors which satisfy the equation (with  $F$  being a companion matrix whose first row equals  $(\theta_1, \theta_2, \dots, \theta_q)$ ):

$$\begin{aligned}\mathbf{W}_t &= F\mathbf{W}_{t-1} \\ &= F^t\mathbf{W}_0,\end{aligned}$$

and the fact that the asymptotic behavior of  $F^t$  depends solely on the largest eigenvalue of  $F$  in magnitude which is smaller than 1 iff the root condition alluded to above holds. Indeed, let  $\lambda_1$  be the largest eigenvalue of  $F$  in magnitude. It is well known that  $\lambda_1$  is less than 1 in magnitude iff all roots of the characteristic equation are outside the unit circle, in which case, for almost all initial  $\mathbf{W}_0$  with respect to the Lebesgue measure,  $|\mathbf{W}_t| \sim \lambda_1^t$ , meaning that the ratio of the two terms is bounded, and hence the reconstruction errors vanish geometrically fast. (Here  $|\mathbf{W}_t|$  denotes the Euclidean norm of the vector  $\mathbf{W}_t$ .) On the other hand,  $\lambda_1$  is larger than 1 in magnitude iff some root of the characteristic equation is inside the unit circle, in which case the reconstruction errors grow exponentially in magnitude, and hence the model is non-invertible. If  $|\lambda| = 1$ , then the model is still non-invertible since the reconstruction errors preserve their magnitude. However, the last case happens with zero Lebesgue measure. Thus, the generic situation is the dichotomy that the reconstruction errors of a linear MA model either vanish geometrically fast or they grow exponentially fast.

It turns out that this dichotomy holds for a local invertibility analysis for any nonlinear MA model. To see this, note that (8) entails that

$$\mathbf{W}_t = \left(\prod_{s=1}^t F_s\right)\mathbf{W}_0. \tag{10}$$



Recall that  $F_s = \dot{F}(0, \boldsymbol{\varepsilon}_s; \theta)$  is a function of  $\boldsymbol{\varepsilon}_s = (\varepsilon_s, \dots, \varepsilon_{s-q+1})^T$ . Under very general conditions, a product of random matrices of the above form asymptotically behaves like the power of some constant matrix. Specifically, Theorem C of Cohen (1988) states that if  $E(\max(0, \log \|F_1\|)) < \infty$ , then

$$\lim_{t \uparrow \infty} t^{-1} \log \left\| \prod_{s=1}^t F_s \right\| = \xi$$

with probability 1, where  $\|\cdot\|$  denotes a matrix norm for which  $\|AB\| \leq \|A\|\|B\|$  for any matrices  $A$  and  $B$ , and  $-\infty \leq \xi < \infty$  is a constant; furthermore,  $\xi = \lim t^{-1} E(\log \|\prod_{s=1}^t F_s\|)$ . The preceding result of Cohen follows from the general subadditive ergodic theory of Kingman (1973). The determination of  $\xi$  is, however, a generally hard problem, except that for the scalar case, i.e.  $q = 1$ ,  $\xi = E(\log \|F_1\|)$ , by the independence of the  $F_s$ 's. Otherwise, only in rare cases does  $\xi$  admit a closed-form expression. Finally, we note that the preceding result on the asymptotic behavior of the product of the random matrices holds if  $F_s$  is a function of a stationary ergodic process; such an extension is useful for the invertibility analysis of a NLARMA model.

In particular, if we take  $\|\cdot\|$  to be the spectral norm (the maximum eigenvalue in magnitude), the preceding result implies that there exists a constant  $\xi$  such that the local reconstruction errors  $|\mathbf{W}_t| \sim (\exp \xi)^t$ , as  $t \rightarrow \infty$  for almost all initial reconstruction error (w.r.t. the Lebesgue measure). Hence, the model is locally invertible iff  $\xi < 0$ . Therefore, the necessary and sufficient conditions for local invertibility of a nonlinear MA model hinge on deriving conditions for  $\xi$  to be less than 0. As mentioned earlier, the determination of  $\xi$  is generally a hard problem. Nevertheless, simple sufficient conditions for  $\xi < 0$  can be obtained by noting that for any fixed positive integer  $m$ ,

$$\lim t^{-1} E(\log \left\| \prod_{s=1}^t F_s \right\|) \leq m^{-1} E(\log (\|F_1 F_2 \cdots F_m\|)).$$

To see this, let  $t = mk + r$  where  $k$  and  $r$  are integers and  $0 \leq r < m$ . Recall the matrix norm has the property that  $\|AB\| \leq \|A\|\|B\|$  for any matrices  $A$  and  $B$ . It follows from this property and stationarity that

$$t^{-1} E(\log \left\| \prod_{s=1}^t F_s \right\|) \leq t^{-1} k E(\log (\|F_1 F_2 \cdots F_m\|)) + t^{-1} E^r(\log \|F_1\|)$$

from which the claimed result can be obtained by passing to the limit. In particular, we obtain the following result.

**THEOREM 2.** *The nonlinear MA( $q$ ) model defined by (1) is locally invertible if  $E(\log \|F_1\|) < 0$  where  $\|\cdot\|$  is a matrix norm, e.g. the spectral norm. For the case  $q = 1$ , the model is locally non-invertible if  $E(\log \|F_1\|) > 0$*

The non-invertibility result for  $q = 1$  stated above follows trivially from the fact that  $\xi = E(\log \|F_1\|)$ , in the scalar case. So far, we assume that the underlying process is an NLMA process, but the preceding theorem can be extended readily to the case of a stationary NLARMA model defined by (3), for which  $F_s$  is a function of  $Y_{s-1}, \dots, Y_{s-p}, \varepsilon_{s-1}$  and  $\theta$ . Furthermore, if  $h$  in (3) is conditionally linear in the innovations given the  $Y$ 's, then the local invertibility analysis is equivalent to global invertibility analysis.

## 5 Threshold MA Model Revisited

Ling and Tong (2005) studied the Threshold MA (TMA) model, which is a piecewise linear MA model. For simplicity, consider the simple case of the two regimes:

$$Y_t = \varepsilon_t - I(Y_{t-d} \leq r) \sum_{j=1}^q \theta_{1,j} \varepsilon_{t-j} - I(Y_{t-d} > r) \sum_{j=1}^q \theta_{2,j} \varepsilon_{t-j} \quad (11)$$

where the  $\theta$ 's are parameters and  $r$  the unknown threshold parameter and  $d$  is a positive integer known as the delay parameter. Intuitively, the transition of a TMA process switches between two MA processes where the MA process indexed by the parameter vector  $(\theta_{1,1}, \dots, \theta_{1,q})^T$  is in operation if the process at lag  $d$  is below the threshold  $r$ , otherwise the MA process indexed by the parameter vector  $(\theta_{2,1}, \dots, \theta_{2,q})^T$  is operational. For the TMA model, the reconstruction errors satisfy (8) with  $F_t$  being a companion matrix with its first row equal to  $(\theta_{1,1}, \dots, \theta_{1,q})I(Y_{t-d} \leq r) + (\theta_{2,1}, \dots, \theta_{2,q})I(Y_{t-d} > r)$ . The remark below Theorem 2 then implies that the TMA model is invertible if the spectral norms of the two sub-MA processes are less than 1. That is, a TMA model is invertible if all the roots of the two characteristic equations

$$1 - \sum_{j=1}^q \theta_{i,j} x^j = 0, i = 1, 2$$

lie outside the unit circle. Much stronger results on the invertibility can be drawn for the first-order TMA models from Theorem 2. For the first order TMA model satisfying (5), it can be checked that  $\exp(\xi) = \prod_{j=1}^k \{|\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})}\}$ , so that Theorem 1 follows from Theorem 2.

However, a TMA process may be globally invertible even if one of its MA sub-processes is non-invertible. Indeed, this can be seen by the following argument. Suppose the MA sub-process in the lower regime ( $Y_{t-d} \leq r$ ) is invertible with spectral norm  $s_1 < 1$ , but the MA process in the upper regime is non-invertible with spectral norm  $s_2 > 1$ . Consider a net of TMA models with its two MA sub-model parameters constant but with variable threshold indexed by  $r \in R$ . As  $r \rightarrow \infty$ , the stationary probability of the upper regime, say  $p_r$ , approaches 0. However,  $E(\log \|F_1\|) = (1 - p_r) \log(s_1) + p_r \log(s_2) \rightarrow \log(s_1) < 0$ , as  $r \rightarrow \infty$ . Hence, for all such TMA models with sufficiently large threshold, they must be invertible. The argument can be generalized to TMA models with more than two regimes. Thus, the TMA model may be globally invertible even though it is locally non-invertible over some regimes! To delineate completely regions of invertibility in the parameter space requires a necessary and sufficient condition. The problem is largely solved for the TMA model of order 1 but it remains open for the higher order cases.

## 6 Conclusion and Acknowledgements

In this note, we illustrate the use of the subadditive ergodic theory for investigating the local invertibility of a NLARMA model, after linking the problem with the stability of an attractor in a dynamical system. This approach yields simple sufficient conditions for invertibility for the Threshold MA models. We conjecture that more sophisticated analysis via this approach may yield weaker conditions for invertibility. Moreover, it is of interest to develop empirical approaches to assess invertibility.

We gratefully acknowledge partial support from the US National Science Foundation (CMG-0620789 for KSC), the Hong Kong Research Grants Committee (GRF HKU 7036/06P for HT) and the University of Hong Kong (Distinguished visiting professorship for HT) .

## References

- [1] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*, 2nd ed. New York: Prentice-Hall.
- [2] Chan, K. S. and Tong, H. (2001). *Chaos: A Statistical Perspective*, New York: Springer Verlag.
- [3] Cohen, J.E. (1988). Subadditivity, generalized products of random matrices and operations research. *Siam Review*, **30**, 1, 69-86.
- [4] Cryer, J.D. and Chan, K.S. (2008). *Time Series Analysis: With Applications in R*, 2nd ed. New York: Springer Verlag.
- [5] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer Verlag.
- [6] Franses, P.H. and van Dijk, D. (2000). *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press.
- [7] Gao, J. (2007) *Nonlinear Time Series: Semiparametric and Nonparametric Methods*. Boca Raton: Chapman & Hall/CRC Press.
- [8] Granger, C.W. and Andersen, A.P., (1978a). *An Introduction to Bilinear Time Series Models*. Göttingen: Vandenhoeck and Ruprecht.
- [9] Granger, C.W. and Andersen, A.P., (1978b). On the invertibility of time series models. *Stoch. Proc. & Appl.*, **8**, 87-92.
- [10] Guegan, D. and Pham Dinh Tuan (1987). Minimalité et inversibilité des modèles bilinéaires à temps discret. *C. R. Acad. Sci. Paris*, **394**, 159-62.
- [11] Kingman, J.F.C. (1973). Subadditive ergodic theory, *Ann. Probability*, **1**, 883-909.
- [12] Ling, S. and Tong, H. (2005) Testing a linear MA model against threshold MA models. *Ann. Statistics*, **33**, 2529-2552
- [13] Priestley, M.B. (1988). *Non-linear and Non-stationary Time Series Analysis*. San Diego: Academic Press.

- [14] Robinson, P. (1977). The estimation of a nonlinear moving average model. *Stoch. Process. Appl.*, **5**, 81-90.
- [15] Small, M. (2005). *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*. World Scientific: Singapore.
- [16] Subba Rao, T., 1981. On the theory of bilinear models. *J. Roy. Statist. Soc.*, **B 43**, 244-255.
- [17] Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*. Oxford: Clarendon Press.