LEAST SQUARES PREDICTORS FOR THRESHOLD MODELS: PROPERTIES AND FORECAST EVALUATION

Alessandra Amendola[1], Marcella Niglio[1], Cosimo Vitale[1]

[1] Dipartimento di Scienze Economiche e Statistiche
Università degli Studi di Salerno
(e-mail: alamendola@unisa.it.it, mniglio@unisa.it, vitale@unina.it)

ABSTRACT: The forecasts generation from SETARMA models is presented and discussed. In particular, least squares, plug-in and combined predictors are pointed out even highlighting the main problems that arise when forecasts have to be computed for this class of threshold models. An empirical example to three stock market indexes shows the performance of the proposed predictors giving evidence in favor of the forecasts combination.

KEYWORDS: SETARMA model, forecasts generation, combination

1 Introduction

Since their introduction at the end of the '70, threshold models have been widely applied to study economic and financial time series.
The interest arisen from these models is testified by the relevant number of variants proposed in literature, with respect to the original one introduced in Tong (1978). Among them, in the present paper the attention is focused on a variant, introduced in Amendola *et al.* (2006), of the so called *Self Exciting Threshold Autoregressive Moving Average* (SETARMA) model (Tong, 1983) which is a direct generalization of the linear ARMA structure (Box & Jenkins, 1976). The k-regimes model of order $(k; p_1, \ldots, p_k; q_1, \ldots, q_k)$ can be shortly given as:

$$X_t = \sum_{i=1}^{k} [\phi_0^{(i)} + \phi_{p_i}^{(i)}(B) X_t^{(i)} + \theta_{q_i}^{(i)}(B) e_t^{(i)}] \cdot I(X_{t-d} \in R_i), \tag{1}$$

with B the backshift operator, the two polynomials $\phi_{p_i}^{(i)}(B) = \sum_{j=1}^{p_i} \phi_j^{(i)} B^j$ and $\theta_{q_i}^{(i)}(B) = 1 - \sum_{w=1}^{q_i} \theta_w^{(i)} B^w$ have no roots in common within each regime, $e_t^{(i)} = \sigma_i^2 e_t$, e_t are i.i.d. random variables, with $E[e_t] = 0$ and $E[e_t^2] = 1$, for $i = 1, \ldots, k$, $I(\cdot)$ is an indicator function, $R_i = [r_{i-1}, r_i)$ forms a partition of the real line such that $-\infty = r_0 < r_1 < r_2 < \ldots < r_k = +\infty$, r_i are the threshold values, d is the threshold delay, p_i and q_i are positive integers.
Amendola *et al.* (2006) show that, under well defined conditions on the process X_t,

model (1), with $k=2$, can be alternatively written

$$X_t = \left[c_0^{(1)} + \sum_{j=0}^{\infty} \psi_j^{(1)} B^j e_t^{(1)}\right] I_{t-d} + \left[c_0^{(2)} + \sum_{j=0}^{\infty} \psi_j^{(2)} B^j e_t^{(2)}\right] (1 - I_{t-d}), \quad (2)$$

with $I_{t-d} = 1$ if $X_{t-d} \geq r_1$ and $I_{t-d} = 0$ otherwise, $c_0^{(i)} = \frac{\phi_0^{(i)}}{1-\phi_{p_i}^{(i)}(B)} = \frac{\phi_0^{(i)}}{1-\sum_{j=1}^{p_i} \phi_j^{(i)}}$, $\frac{\theta_{q_i}^{(i)}(B)}{\phi_{p_i}^{(i)}(B)} = \sum_{j=0}^{\infty} \psi_j^{(i)} B^j$, where the weights $\psi_j^{(i)}$, properly computed, are such that $\psi_0 = 1$ and $\sum_{j=0}^{\infty} |\psi_j^{(i)}| < \infty$, for $i = 1,2$.

Starting from model (1), with $k=2$, and from the alternative form (2), exact multi-step forecasts have been derived. In particular in Section 2, the best predictor, in term of minimum mean square error, for this class of models is presented, highlighting different aspects which affect its generation further distinguishing among the *least squares* multi-step predictors and the so called *plug-in* predictors. Taking advantage of these results, the two predictors are properly combined in a scheme based on their variance of the two cited predictors. An application to three stock markets index returns is shortly presented in the last Section.

2 The SETARMA predictors

Given the time series $X_1, \ldots, X_t$ and the lead time h, it is well known that the least square predictor of X_{t+h}, denoted $X_t(h)$, is obtained from the conditional expectation $X_t(h) = E[X_{t+h}|\Omega_t]$, with $\Omega_t = \{X_1, \ldots, X_t\}$.

When a SETARMA model is involved, the forecast $X_t(h)$ is strongly subjected to the *threshold variable*, which controls the switching among regimes, and the *threshold delay*, which has relevant implications on the predictor form and on its distribution. More precisely, when $h \leq d$ the predictor is derived simply following the results of the linear time series analysis (Box & Jenkins, 1976). On the contrary, when $h > d$ the estimation of the threshold variable implies some relevant complications here briefly discussed.

When $h \leq d$, the generation of multi-step forecasts is $X_t(h) = X_t^{(1)}(h) I_{t+h-d} + X_t^{(2)}(h)(1 - I_{t+h-d})$ or, using the SETARMA representation (2), it can be even written $X_t(h) = \sum_{i=1}^{2} \left[c_0^{(i)} + \sum_{j=h}^{\infty} \psi_j^{(i)} e_{t+h-j}^{(i)}\right] I(X_{t-d} \in R_1)$, where $E(e_{t+h-j}^{(i)}|\Omega_t) = 0$, for $j = 1,2,...,h-1$ and $R_1 = [r_1, +\infty)$.

The results completely change when the lead time h is greater than the threshold delay d. In particular, given the SETARMA model, its least squares predictor implies the estimation of $E(I_{t+h-d} = 1|\Omega_t) = P(X_{t+h-d} \geq r|\Omega_t)$.

In this case, $X_{t+h-d} \notin \Omega_t$ and so I_{t+h-d} becomes a Bernoulli random variable

$$i_{h-d} = \begin{cases} 1 & \text{with} \quad P(X_{t+h-d} \geq r|\Omega_t) \\ 0 & \text{with} \quad P(X_{t+h-d} < r|\Omega_t) \end{cases} \qquad \text{for } h = d+1, d+2, \ldots \quad (3)$$

where $P(X_{t+h-d} \geq r|\Omega_t) = E[i_{h-d}|\Omega_t] = p_{(h-d)}$.
The least square predictor in this case becomes

$$X_t(h) = X_t^{(2)}(h) + p_{(h-d)} \cdot \left(X_t^{(1)}(h) - X_t^{(2)}(h)\right), \tag{4}$$

which can be even written as

$$X_t(h) = c_0^{(2)} + \sum_{j=h}^{\infty} \psi_j^{(2)} e_{t+h-j}^{(1)} + \left[c_0^{(1)} - c_0^{(2)} + \sum_{j=h}^{\infty} \left(\psi_j^{(1)} - \psi_j^{(2)}\right) e_{t+h-j}^{(2)}\right] p_{(h-d)}, \tag{5}$$

with the prediction error $e_t(h)$:

$$e_t(h) = e_t^{(2)}(h) + I_{t+h-d} \cdot [e_t^{(1)}(h) - e_t^{(2)}(h)] + [I_{t+h-d} - p_{(h-d)}] \cdot [X_t^{(1)}(h) - X_t^{(2)}(h)],$$

where $e_t^{(i)}(h) = \sum_{j=0}^{h-1} \psi_j^{(i)} e_{t+h-j}^{(i)}$ and $X_t^{(i)}(h) = \sum_{j=h}^{\infty} \psi_j^{(i)} e_{t+h-j}^{(i)}$ are the forecast errors and the prediction generated from regime i respectively ($i = 1, 2$). Finally the variance of $e_t(h)$ is

$$\begin{aligned} \sigma_e^2(h) &= \sigma_{2,e}^2(h) + p \cdot [\sigma_{1,e}^2(h) - \sigma_{2,e}^2(h)] + [p + p_{(h-d)}^2 - 2p \cdot p_{(h-d)}] \cdot \\ &\cdot [\sigma_{1,X}^2(h) + \sigma_{2,X}^2(h) - 2\sigma_{12,X}(h)]. \end{aligned} \tag{6}$$

2.1 Plug-in and combined forecasts

When $h > d$ the generation of multi step-ahead predictions can be even accomplished using a different strategy frequently used in empirical framework. The forecasts are generated treating the values predicted at the previous steps, as true values.
This implies that the conditional set Ω_t grows at each step becoming $\Omega_t(h-d) = \{X_1, \ldots, X_t, X_t(1), \ldots, X_t(h-d)\}$. In this case the predictions $X_t(1), \ldots, X_t(h-d)$, which belong to $\Omega_t(h-d)$, are treated as true values whereas the indicator function I_{t+h-d} becomes:

$$i_t(h-d) = [I_{t+h-d}|\Omega_t(h-d)] \begin{cases} 1 & \text{if } X_t(h-d) \geq r \\ 0 & \text{if } X_t(h-d) < r. \end{cases} \tag{7}$$

The *plug-in* predictor, $X_t^{PI}(h)$ is so given as

$$X_t^{PI}(h) = E[X_{t+h}|\Omega_t(h-d)] = X_t^{(2)}(h) + i_t(h-d) \cdot \left(X_t^{(1)}(h) - X_t^{(2)}(h)\right), \tag{8}$$

with forecast error:

$$e_t^{PI}(h) = e_t^{(2)}(h) + I_{t+h-d} \cdot [e_t^{(1)}(h) - e_t^{(2)}(h)] + [I_{t+h-d} - i_t(h-d)] \cdot [X_t^{(1)}(h) - X_t^{(2)}(h)],$$

and variance of $e_t^{PI}(h)$

$$\begin{aligned} \sigma_{PI,e}^2(h) &= \sigma_{2,e}^2(h) + p \cdot [\sigma_{1,e}^2(h) - \sigma_{2,e}^2(h)] + [p + i_t(h-d) - 2p \cdot i_t(h-d)] \cdot \\ &\cdot [\sigma_{1,X}^2(h) + \sigma_{2,X}^2(h) - 2\sigma_{12,X}(h)]. \end{aligned} \tag{9}$$

Amendola *et al.* (2005) show that the two predictors $X_t(h)$ and $X_t^{PI}(h)$ are both unbiased whereas the variances (6) and (9) are such that

$$\frac{\sigma^2_{LS,e}(h)}{\sigma^2_{PI,e}(h)} \geq 1 \quad \text{if} \quad X_t(h-d) \geq r \quad \text{and} \quad \frac{\sigma^2_{LS,e}(h)}{\sigma^2_{PI,e}(h)} \leq 1 \quad \text{if} \quad X_t(h-d) < r. \qquad (10)$$

Taking advantage of these results, the predictors (4) and (8) are combined as

$$X_t^C(h) = X_t^{LS}(h) + i_t(h-d)\left[X_t^{PI}(h) - X_t^{LS}(h)\right], \qquad (11)$$

where the forecasts are generated picking the "best" (in term of MSFE) predictor selected according to the indicator function $i_t(h-d)$ and conditional to the augmented set $\Omega_t(h-d)$.

3 Empirical results

In order to evaluate the forecast accuracy of the Least Squares (LS), Plug-in (PI) and Combined (C) predictors, they are applied to generate ex-post forecasts from daily Dow-Jones (DJ), Nikkey (NK) and FTSE 100 (FTSE) stock index log-returns form January, 1^{st} 2002 to June, 30^{th} 2006. The SETARMA models fitted to the data have been selected according to the minimum AIC leaving out the last 10 observations used to generate forecasts. The LS, PI and C forecasts have been compared, in terms of Root Mean Square Error, in Table 1. The combination outperforms the LS and Plug-in forecasts with the DJ and FTSE time series, whereas the RMSE of PI and C of the NK agree because the second inequality in (10) is always fulfilled for $h = 1, 2, ..., 10$.

	LS	PI	C
Dow Jones	0.86475	0.86189	0.86021
Nikkey	1.58643	1.58415	1.58415
FTSE 100	0.74898	0.74271	0.74223

Table 1. *RMSE*($\times 10^{-2}$) *of the forecasts generated from DJ, NK and FTSE time series with* $h = 10$

References

AMENDOLA, A., NIGLIO, M., & VITALE, C. 2005. Multi-step ahead predictors of SETARMA models. *Mimeo*.

AMENDOLA, A., NIGLIO, M., & VITALE, C. 2006. The autocorrelation functions in SETARMA models. *In:* KONTOGHIORGHES, E.J., & C., GATU (eds), *Optimization, Econometric and Financial Analysis*. Springer-Verlag.

BOX, G.E.P., & JENKINS, G.M. 1976. *Time series analysis, forecasting and control*. Holden-Day, San Francisco.

TONG, H. 1983. *Threshold models in nonlinear time series analysis*. Springer-Verlag.