PSO ESTIMATION OF ASYMETRIC THRESHOLD MODELS

Claudio Pizzi[2] and Francesca Parpinel[3]

[1] Dipartimento di Economia
Università Ca' Foscari Venezia
San Giobbe, Cannaregio 873, 30121 Venezia, Italy
[2] (e-mail: pizzic@unive.it)
[3] (e-mail: parpinel@unive.it)

ABSTRACT. The TAR model introduced by Tong may be specified in several different ways. Here we propose the case in which the delay parameter is endogenous, depending on both the past value of the series and the specific past regime. In particular we consider a system switching between two regimes each of which consider an AR(p) with respect of the value assumed by a delayed self–variable compared with an asymmetric threshold, that is the switching rule also depend on the regime in which the system is at time $t-d$.

To estimate this model we propose a procedure based on the Particle Swarm Optimization technique.

1 INTRODUCTION

The class of models proposed by Tong and Lim in the eighties is able to capture nonlinear behaviour of several different data generating process (DGP) such as jump resonance, amplitude–frequency dependency, limit cycle, subharmonics and higher harmonics. These models improve the global linear model by means of a piecewise linear model.

After the seminal paper by Tong e Lim (1980), this model captured the attention of many researchers. On one hand many articles tackle the problem to test the threshold hypothesis, see Petruccelli and Davies (1986), Tsay (1989), e Chang and Tong (1990). On the other hand, several authors face the parameters estimation problem applying the procedures to real data. For example, Potter (1995) models the nonlinear structure of evolution of US GNP, suggested by the response of output to shocks at different stages of the business cycle being asymmetric; to discover these differences he uses the SETAR model. More recently, Battaglia and Protopapas (2011), Wu and Chang (2002) use genetic algorithms, and Gonzalo and Wolf (2005) propose the subsampling techniques.

In this work we want to stretch the asymmetry characterization using a threshold. The goal of the proposed model is to capture better the asymmetric nature of some phenomenum. In finance for example, the stock market might be *bull* or *bear* and, as it is well known, the behaviour in these two *status* may be different. In particular the threshold at which the bull market begins to operate is different from that at which the bear market begins to operate.

The class of model we use are called Asymmetric Self-Exciting Threshold Autoregressive model, in short ASETAR, recalling both the asymmetric behaviour in the two regimes and the different thresholds activating the regimes.

The aim of this work is to evaluate the ability of an evolutive algorithm to improve some classical estimation procedures that we can use even in this context. In particular we want to estimate simultaneously all the parameters of the model.

In the next Section we introduce the asymmetric SETAR, that allows us to take into account the possible overlapping of the regimes, in its simplest structure. In Section 3 we briefly describe the Particle Swarm Optimization procedure and propose an identification and an estimation procedure. In Section 4 we present some resuts obtained on simulated data. We close this work emphasizing the main results and the open problems.

2 THE ASYMETRIC SETAR MODEL

The main feature of SETAR model introduced by Tong is that the thresholds define a partition of space $\mathbb{R}$, in the sense that we cross from the first regime to the second one when the delayed value y_{t-d} increases until the threshold is exceeded whereas the passage from the second regime to the first occurs when the delayed value y_{t-d} decreases until it becomes less than the threshold. In some sense, there is a symmetric behaviour with respect to the threshold. This property is very useful to study some phenomena, but sometimes might reveal unsuitable to study other ones.

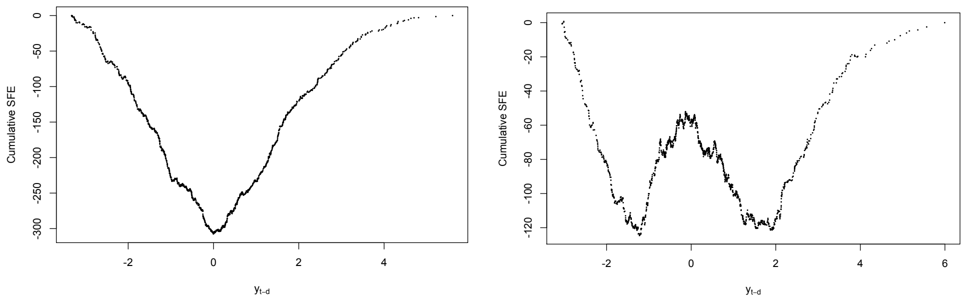
Let y_t, with $t = 1, \ldots, T$ be an observed times series and let $r_1 > r_2$ be two thresholds such that $(-\infty, r_1), (r_2, \infty)$ are two overlapping intervals. We define an Asymmetric Self-Exciting Threshold Autoregressive model, in short ASETAR$(2; p, d)$, with two regimes of order p with delay d, as follows

$$y_t = \begin{cases} a_0^{(1)} + \sum_{i=1}^{p} a_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{if } y_{t-d} \leq r_1 \text{ and } J_{t-d} = 1 \\ a_0^{(2)} + \sum_{i=1}^{p} a_i^{(2)} y_{t-i} + \varepsilon_t^{(2)} & \text{if } y_{t-d} > r_1 \text{ and } J_{t-d} = 1 \\ a_0^{(2)} + \sum_{i=1}^{p} a_i^{(2)} y_{t-i} + \varepsilon_t^{(2)} & \text{if } y_{t-d} > r_2 \text{ and } J_{t-d} = 2 \\ a_0^{(1)} + \sum_{i=1}^{p} a_i^{(1)} y_{t-i} + \varepsilon_t^{(1)} & \text{if } y_{t-d} \leq r_2 \text{ and } J_{t-d} = 2 \end{cases} \tag{1}$$

where J_{t-d} is an unobservable variable denoting the theoretical regime in which operate the system at time $t-d$, d is the delay with which the system reacts and changes regime, $r_1 > r_2$ are two thresholds depending on the regimes. This fact means that when the system work in regime 1 the activating threshold is different with respect to that working in regime 2.

Another difference in the behaviour of the SETAR and ASETAR models can be revealed applying, in both cases, the classical estimation procedures for the SETAR. In fact, in order to get the estimates, the classical techniques suggest to use the standardized forecast error (in short SFE). If we apply this techniques in the SETAR case we will find that the cumulative sum of SFE values versus n decreases until the value of the delayed variable is lesser than the threshold, and it increases in the other cases. The behaviour in the ASETAR case is quite different: there is a subset of ordered n in which the cumulative SFE behaves in a mixed way, as we can see from the graphs in Figure 1.

We emphasize again that the asymmetry does not refer to the statistical distribution of the variables but to the switching rule. Note that for some value of y_{t-d} the system may assume two different configuration. In other words in the interval (r_2, r_1) the process might assume two different structures depending on the system status.



a) SETAR(2;1,1) with $d = 1$, $r = 0$, $a_1^{(1)} = 0.8$, $a_1^{(2)} = -0.8$

b) Asymmetric SETAR with $r_1 = 1.25, r_2 = -1.25$, $d = 1$, $a_1^{(1)} = 0.8$, $a_1^{(2)} = -0.8$

Figure 1. Cumulative sum of standardized forecast errors.

3 IDENTIFICATION AND ESTIMATION

In order to identify a model on the basis of a time series observation typically we have even in the case of SETAR frameworks the problem that all the parameters present in the model are not simultaneously estimated. Some recent works try to overcome this problem, Battaglia and Protopapas (2011) propose the use of genetic algorithm to estimate a SETAR model whereas Maringer and Meyer (2008) consider the Simulated Annealing technique in the estimation of a STAR model.

With the same aim we suggest to use another evolutive approach that is the Particle Swarm Optimization (PSO). The PSO, born in the eighty, replicates the behaviour of natural flocks and swarm of animals having one specific aim or objective, preserving a swarm of particles moving around in the search space influenced by the improvements discovered by the other particles.

This general optimization procedure, due to Kennedy and Eberhart (1995) and Shi and Eberhart (1998), may be applied in very different problems.

The PSO algorithm we perform is described as follows

1. Creation of a *population* of P particles ($\mathbf{x}_i^1, i = 1, \ldots, P$) uniformly distributed over the parametric space.
2. Evaluation of each particle's position according to the objective (fitness) function.
3. If a particle's current position is better than its previous best position ($\mathbf{pb}_i$), updating it.
4. Determination of the best particle ($\mathbf{lb}_i$), according to the particle's previous best position.
5. Updating particles' velocities according to

$$\mathbf{v}_i^{\tau+1} = \omega \mathbf{v}_i^{\tau} + \varphi_1 \mathbf{U}_1^{\tau} (\mathbf{pb}_i^{\tau} - \mathbf{x}_i^{\tau}) + \varphi_2 \mathbf{U}_2^{\tau} (\mathbf{lb}_i^{\tau} - \mathbf{x}_i^{\tau})$$

   where $\mathbf{v}_i^{\tau}$ is the velocity of particle i at iteration τ, ω is an inertia weight, $\mathbf{U}_j^{\tau}$ are values from uniform random variables, φ_j are weights on the attraction towards the particles own best known position, $\mathbf{pb}_i^{\tau}$, and the swarms best known position, $\mathbf{lb}^{\tau}$.

6. Moving particles to their new positions according to

$$\mathbf{x}_i^{\tau+1} = \mathbf{x}_i^{\tau} + \mathbf{v}_i^{\tau+1}.$$

7. Going to step 2 until stopping criteria are satisfied.

The velocity equation we consider at step 5 is that proposed by Shi and Eberhart (1998) in the PSO variant called *inertia weight*. This technique has many advantage with respect to other procedures, first of all its simpleness; furthermore it does not require the gradient of the objective functions. Moreover, it starts from random position of the particles of the swarm, moving the particles in the space of m dimension according to some speed rule and the aim is to gain an optimum. On the other hand, the performance of the procedure is strongly influenced by the tuning of its behavioural parameters (ω , φ_1, φ_2), and to overcome this problem Pedersen (2010) studies and proposes some optimal combinations for them.

4 APPLICATION

In our applications we consider three data generating processes, one in the class of the linear model, one in the SETAR class and the last in the ASETAR class, from which one we simulate 100 series of 500 observations. To perform the algorithm we use 40 particles over 50 replications of it obtaining 2000 fitness evaluations. The 40 particles, with dimension equal to the number of model parameters used in the fitness function (in our case belongs to $\{1,2,5\}$), are generated by a uniform random variable defined over $[-1,1]$. The optimization function is the mean square error of the simulated series against the fitted values obtained by the model defined by each particle and the parameters procedure are tuned to $\omega = 0.7$, $\varphi_1 = 1.4$, $\varphi_2 = 1.4$. The algorithm is stopped when there are not improvements on fitness function or untill 100 iteration are reached. In order to arrange an estimate for one series we consider the mean vector of the particles getting the best fitness.

Tables 1, 2 e 3 present some results obtained performing the PSO procedure using the fitness function

$$fit = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2$$

where $y_t, t = 1,\dots,T$ is the time series and $\hat{y}_t, t = 1,\dots,T$ is the fitted series according to the model we want to estimate. In our application we consider three different models, that is three different fitness functions: AR(1), AR(2) and ASETAR(2;1,2).

The results are encouraging indeed in the case of the AR(1), see Table 2, the PSO procedure estimates correctly the parameter with either autoregressive or ASETAR fitness function. We point up that when we use the AR(2) fitness function the parameter ϕ_2 is estimated close to zero whereas when the fitness function is ASETAR the threshold parameter are estimated equal to zero and the autoregressive parameters of the two regimes are both equal to 0.78 in mean. At a glance the PSO algorithm captures the true value of the unknown parameter forcing the redundant parameter to zero. The delay parameter has no sense in this framework. More important is to comment the results for ASETAR data, reported in Table 1. We can see that the estimates of autoregressive parameters are very close to their theoretical values, just

Table 1. PSO estimation over 100 trajectory simulated by ASETAR: some statistics.

model (1) with $a_0^{(1)} = a_0^{(2)} = 0$, $a_1^{(1)} = 0.8$, $a_1^{(2)} = -0.8$, $d = 2$, $r_1 = 1$, $r_2 = -1$

Statistics	linear 1 par.	linear 2 par.		asetar 5 parameter				
	a_1	a_1	a_2	$a_1^{(1)}$	$a_1^{(2)}$	r_1	r_2	d
Min.	-0.162	-0.192	-0.180	0.648	-0.843	-1.00	0.33	1.92
1st Qu.	-0.040	-0.041	-0.034	0.762	-0.797	-0.87	0.75	1.99
Median	0.041	0.040	0.034	0.786	-0.769	-0.81	0.86	1.99
Mean	0.062	0.049	0.038	0.784	-0.769	-0.80	0.83	1.99
3rd Qu.	0.138	0.125	0.114	0.814	-0.733	-0.74	0.95	2.00
Max.	0.337	0.264	0.225	0.880	-0.649	-0.48	1.04	2.00
SD	0.129	0.116	0.093	0.040	0.042	0.10	0.15	0.01
VarWithin	4e-06	3e-07	3e-07	7e-3	9e-3	0.39	0.24	0.01

Table 2. PSO estimation over 100 trajectory simulated by AR: some statistics.

model $y_t = 0.8 y_{t-1} + \varepsilon_t$

Statistics	linear 1 par.	linear 2 par.		asetar 5 parameter				
	a_1	a_1	a_2	$a_1^{(1)}$	$a_1^{(2)}$	r_1	r_2	d
Min.	0.743	0.712	-0.112	0.658	0.652	-0.36	-0.44	1.00
1st Qu.	0.782	0.769	-0.031	0.745	0.751	-0.16	-0.14	1.03
Median	0.799	0.801	-0.003	0.796	0.802	-0.02	0.01	1.07
Mean	0.797	0.799	-0.003	0.781	0.785	-0.03	0.04	1.11
3rd Qu.	0.813	0.829	0.033	0.816	0.818	0.10	0.16	1.17
Max.	0.854	0.910	0.101	0.866	0.861	0.39	0.82	1.37
SD	0.024	0.041	0.045	0.054	0.051	0.18	0.27	0.10
VarWithin	8e-05	6e-07	6e-07	2e-3	2e-3	1.13	1.46	0.17

Table 3. PSO estimation over 100 trajectory simulated by SETAR: some statistics.

$$y_t = \begin{cases} 0.8 y_{t-1}^{(1)} + \varepsilon_t^{(1)} & \text{if} \quad y_{t-2} \leq 0 \\ -0.8 y_{t-1}^{(2)} + \varepsilon_t^{(2)} & \text{if} \quad y_{t-2} > 0 \end{cases}$$

Statistics	linear 2 par.		asetar 5 parameter				
	a_1	a_2	$a_1^{(1)}$	$a_1^{(2)}$	r_1	r_2	d
Min.	-0.343	-0.322	0.592	-0.869	0.25	-0.11	1.89
1st Qu.	-0.102	-0.095	0.755	-0.810	0.46	-0.02	1.98
Median	-0.004	-0.014	0.789	-0.787	0.52	0.00	1.99
Mean	-0.018	-0.019	0.783	-0.786	0.52	0.00	1.99
3rd Qu.	0.086	0.067	0.815	-0.765	0.58	0.02	2.00
Max.	0.304	0.217	0.890	-0.606	0.73	0.09	2.00
SD	0.129	0.116	0.051	0.042	0.09	0.03	0.02
VarWithin	3e-07	3e-07	0.019	0.011	0.303	0.02	0.01

as the estimated delay parameter. Some problems arise for the estimation of the threshold parameters, but we think that may be overcame in a more suitable tuning of the PSO procedure. When the fitness function of the PSO in not the correct one, the estimated model is just zero, so we have no estimates. This happens even for SETAR data, see Table 3. In this last case when we consider the ASETAR fitness function in PSO we obtain satisfactory estimates for the regimes and delay, the threshold is estimated with some problems. In all Tables the statistics in the last row, defined *VarWithin*, represent the variability of the particles of the PSO, and in all cases is very low as we expect.

5 CONCLUSION

The proposed technique seems to obtain interesting results in the parameters estimation; the preliminary results show robustness with respect to model miss-specification; in the cases of threshold model the technique enable us to estimate simultaneously thresholds, delay and autoregressive parameters. Furthermore we think that this model might be interesting to analyze some real time series such as the financial one that potentially show asymmetric features.

REFERENCES

BATTAGLIA, F., PROTOPAPAS, M.K., (2011): Multi-regime models for nonlinear nonstationary time series, *Computational Statistics*, *47*, 277–295.

CHAN, K.S., TONG, H., (1990): On Likelihood Ratio Tests for Threshold Autoregression, *Journal of the Royal Statistical Society. Series B (Methodological)*,*52*, 469–476.

GONZALO, J., WOLF, M. (2005): Subsampling inference in threshold autoregressive models, *Journal of Econometrics*, *127*, 201–224.

KENNEDY, J., EBERHART, R.C. (1995): Subsampling inference in threshold autoregressive models, in *In Proceedings of IEEE International Conference on Neural Networks*, 1942-1948.

MARINGER, D.G., MEYER, M., (2008), Smooth Transition Autoregressive Model-New Approaches to the Model Selection Problem, *NonliNear Dynamical Methods and Time Series Analysis*, *12-1*.

PEDERSEN, M.E.H., (2010), Good Parameters for Particle Swarm Optimization, *Hvass Laboratories*, Technical Report *HL1001*.

PETRUCCELLI, J.D., DAVIES, N. (1986): A Portmanteau Test for Self-Exciting Threshold Autoregressive-Type Nonlinearity in Time Series, *Biometrika*,*73*, 687–694.

POTTER, N. (1995): A nonlinear approach to US GNP, *Journal of applied econometrics*, *10*, 109–125.

SHI, Y., EBERHART, R. (1998): A modified particle swarm optimizer. in *Proceedings of the IEEE International Conference on Evolutionary Computation*, 69–73.

TONG, H. (2011): Threshold Models in time series analysis-30 years on, *Statistics and its Interface*, *4:2* 107-118.

TONG, H., LIM, K.S. (1980): Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *42*, 245–292.

TSAY, R.S. (1998): Testing and Modeling Multivariate Threshold Models, *Journal of the American Statistical Association*, *93*, 1188–1202.

WU, B., CHANG, C.L., (2002): Using genetic algorithm to parameters (d,r) estimation for threshold autoregressive models. *Comput. Stat. Data Anal.*, *17*, 241–264.