Say, we want to model a financial time series, which is known to exhibit regime switching behaviour, i.e. periods of high and low volatility. So we decide to use regime switching models. The key challenges that need to be addressed are:

- Finding the evidence in favour of using regime switching models over linear models
- Estimating the number of regimes in the time series
- Determining which periods belong to which regime
- Estimating the model parameters for each separate regime in a unified framework
- Identifying the times at which different regimes apply, and being able to switch between regimes at the right times

Threshold Autoregression

- Threshold Autoregression (TAR) model were proposed by Tong (1978) and further discussed by Tong and Lim (1980)

- The TAR model with two regimes, with an AR(1) model structure is given by:

$$y_t = \begin{cases} \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_{t1} \; if \; y_{t-d} < r \\ \beta_0 + \beta_1 y_{t-1} + \varepsilon_{t2} \; if \; y_{t-d} \geq r \end{cases}$$

where r is a threshold variable that divides the domain into two regimes, while y_{t-d} is the state determining variable, whereas d is a delay variable (integer valued)

- In each distinct regime, the time series y_t is modelled by a AR(1) model. Although the model structure remains same (AR) in each regime, the model coefficients are different

Threshold Autoregression

- The general form of TAR model is given by:

$$y_t = \alpha_0^{(i)} + \sum_{j=1}^{p^{(i)}} \alpha_j^{(i)} y_{t-j} + \varepsilon_t^{(i)}$$

$$i = 1,2,\dots,k; -\infty = r_0 < r_1 < \cdots < r_k = \infty$$

The k partitions, r_i, correspond to each regime and form a non-overlapping partition of the real line

d is the delay parameter (or threshold lag)

$p^{(i)}$ is the AR model order in the i^{th} regime

$\varepsilon_t^{(i)}$ is a sequence of i.i.d. normal random variables with zero mean and $\sigma^{(i)}$ standard deviation

TAR Applications

TAR applications:

• Tong and Lim (1980) showed that the TAR model is capable of producing the asymmetric, periodic behaviour exhibited in the annual sunspot data and the Canadian Lynx data

• Tyssedal and Tjostheim (1988) applied the TAR model to daily closing prices of IBM between 1959 and 1962

• Tong (1990) analysed the Hang Seng Index from 1984 to 1987

• Pope and Yadav (1990) employed a TAR model to characterise mispricing of the FTSE 100 index futures

• Cao and Tsay (1993) investigated monthly volatility

• Tiao and Tsay (1994) employed a TAR model to investigate the cyclical properties of US GNP

• Gao and Wang (1999) analysed the non-linear dynamics of the S&P 500 index

- **Hansen (2011)** provides an exhaustive review of the applications of TAR in modelling interest rates, price movements, stock returns and exchange rates (*Statistics and its interface*, 123-127; 2011)

Self-exciting TAR (SETAR) model

- SETAR models are a subclass of the Threshold Autoregressive (TAR) models
- When the threshold variable (r) is an observation in the time series, the TAR model is described as 'self-exciting', hence the name SETAR
- A self-exciting threshold autoregressive (SETAR) model is a piecewise linear model
- SETAR provides local linear approximations using distinct AR models while switching between different regimes based on the value of the delay variable
- In a financial setting, the return generating mechanism depends on the value of the price in previous period
- The key characteristics of a SETAR model include time irreversibility, asymmetric limit cycles and jumps
- A major advantage of this model is that the parameters can be estimated using least squares

SETAR model

A SETAR model composed of N_r regimes is denoted as SETAR $(N_r; p_1, p_2, \dots, p_{N_r})$, and is represented as:

$$y_t = \alpha_0^{(i)} + \sum_{j=1}^{p^{(i)}} \alpha_j^{(i)} y_{t-j} + \varepsilon_t^{(i)}$$

where $\alpha_j^{(i)}$ is the j^{th} autoregressive coefficient for a given regime index i that obeys $r_{i-1} \leq y_{t-d} < r_i$, d is the delay order, while r_i for $i = 1, 2, \dots, N_r - 1$ are the *thresholds* that divide the time series into N_r different regimes, $p^{(i)}$ is the AR model order for regime i, and $\varepsilon_t^{(i)}$ is a NID process with mean zero and variance $\sigma^{2(i)}$. The values of thresholds are chosen from observations in the time series, i.e. thresholds are restricted to be *endogenous*, and the choice of AR model coefficients depends on the thresholds r_i and the *delay order* d

SETAR model

- Each regime is governed by an AR model
- The AR models are separated by the lagged observation value y_{t-d} and the value of the threshold r
- The magnitude of delay is given by the delay order d
- The choice of regime is based on the value of a threshold variable, whereby the role of the threshold variable is to force the model to switch between different regime models
- Note that in TAR, the threshold variable is exogenous, while in SETAR, the threshold is restricted to be endogenous
- Both TAR and SETAR are piecewise linear AR models
- From the perspective of economic modelling, the condition $y_{t-d} \leq r$ may correspond to a regime of recession, while $y_{t-d} > r$ may correspond to regime of growth
- For financial time series, the same conditions may represent periods of high and low volatility

SETAR estimation

- The SETAR model requires the estimation of:
  - thresholds r_i
  - delay order d
  - autoregressive model parameters $\alpha_j^{(i)}$
  - model orders $p^{(i)}$

- Common to make a subjective judgement about number of regimes N_r (growth/recession). One could also estimate N_r using cross-validation

SETAR estimation

- Hansen (1996) discusses SETAR parameter estimation
- Vary d and r over a grid and select parameters which minimizes RSS
- d takes on integer values 0, 1, 2,…
- r is usually varied in the range of the 15th and 85th percentiles of distribution of the time series
- To estimate model coefficients, split the sample into N_r regimes and perform OLS separately
- If p is unknown, use AIC and minimize:

$$N_L \ln\hat{\sigma}_L^2 + N_U \ln \hat{\sigma}_U^2 + 2(p_L + 1) + 2(p_U + 1)$$

where N_L and N_U are number of observations in the lower and upper regime, $\hat{\sigma}_L^2$ and $\hat{\sigma}_U^2$ are the corresponding variances of the residuals, while p_L and p_U are the model orders within each regime

SETAR model estimation

- **Model**:

$$y_t = \begin{cases} \alpha_{0U} + \sum_{j=1}^{p^{(1)}} \alpha_{jU} y_{t-j} + \varepsilon_t^{(1)}, if \;\; y_{t-d} \le r & where \;\; \varepsilon_t^{(1)} \sim \text{NID}(0, \sigma_U) \\ \alpha_{0L} + \sum_{j=1}^{p^{(2)}} \alpha_{jL} y_{t-j} + \varepsilon_t^{(2)}, if \;\; y_{t-d} > r & where \;\; \varepsilon_t^{(2)} \sim \text{NID}(0, \sigma_L) \end{cases}$$

PARAMETER ESTIMATES FOR SETAR(2;1,1) MODEL AND
CORRESPONDING STANDARD ERRORS (IN PARENTHESIS) FOR US GNP

Parameter	*Estimate (Standard Error)*
Lower Regime	
α_{0L}	0.4507 (0.1803)
α_{1L}	0.2806 (0.1506)
σ_L	1.2359
N_L	52
Upper Regime	
α_{0U}	0.6166 (0.1081)
α_{1U}	0.3159 (0.0769)
σ_U	0.8952
N_U	145
r	0.2808
d	2
Total Parameters	8

SETAR Applications

SETAR applications:

- Boero and Marrocu (2002) – Euro effective exchange rate
- Tiao and Tsay (1994), Potter (1995), Clements and Krolzig (1998) – US gross national product
- Watier and Richardson (1995) – epidemiological time series
- Feng and Liu (2002) – Canadian gross domestic product (GDP)
- Gooijer and Bruin (1998) provide a review on SETAR processes

- *Smooth transition* variants of regime switching models include: Smooth Transition Autoregression (STAR), Logistic Smooth Transition Autoregression (LSTAR) and Exponential Smooth Transition Autoregression (ESTAR)