Study of Non linear time series models using R

A001 Ved Deshpande

Nilkamal School of Mathematics, Applied Statistics and Analytics

Contents

1 Introduction

In the modern world of fluctuating economies, uncertain events and various unpredictable situations in the financial domain, to draw a pattern in the financial time data is getting difficult. With the current financial time series data, only the use of linear models is not sufficient as the data may itself have non-linearity. In such situations working with the linear time series models becomes futile.
To study or draw patterns in such non-linear data, we have to make use of non-linear time series models to draw inferences out of the non-linear time series data.

$$x_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

where x_t is the time series data, ψ_i belongs to real numbers with $\psi_0 = 1$ and white noise as a_t which follows identical independently distributed random variable with a well define distribution. E(a_t)=0 and Var(a_t)=σ_a^2. Any time series or Stochastic process which does not follow the above equation is said to be a non-linear stochastic process or non-linear time series. To define a model for non-linear time series we have,

$$x_t = g(F_{t-1}) + \sqrt{h(F_{t-1})\epsilon_t}$$

where F_{t-1} is the collection of linear combination of elements in $[x_{t-1}, x_{t-2}, ...]$ and $[a_{t-1}, a_{t-2}, ...]$ and $\epsilon_t = \frac{a_t}{\sigma_t}$.

There are many non-linear models to analyse the non-linear time series data namely, Bi-linear model, threshold auto-regressive model and Markov switching model. The idea that these model propose is the conditional mean evolve over time according to some parametric nonlinear function. In this paper, we will majorly focus on Markov switching model and apply it on a given US unemployment data set.

2 Problem Statement

The financial time series data in the modern day can be very vivid and we cannot always apply simple linear models for forecasting , we may have to use non linear methods for the given non linear data and try to fit a corresponding suitable non-linear model to our data. The aim is to study and understand various different non-linear time series models and apply one of them to a given data set and extract information out of it.

3 Theory

3.1 Markov Switching Auto-regressive Model

For a hidden two-state Markov chain, a time series x_t is said to follow Markov switching Auto-regressive model if it satisfies

$$x_t = \begin{cases} c_1 + \sum_{i=1}^{p} \phi_{1,i} x_{t-i} + a_1 t, & \text{if } s_t{=}1 \\ c_2 + \sum_{i=1}^{p} \phi_{2,i} x_{t-i} + a_2 t, & \text{if } s_t{=}2. \end{cases}$$

where the transition probabilities are given as

$$P(s_t = 2|s_{t-1=1}) = w1, \quad P(s_t = 1|s_{t-1=2}) = w2$$

A small w_i means that the model tends to stay longer in state i and 1 w$_i$ is the expected time duration of the process to stay in state i. This model uses the hidden Markov chain to tell the transition from one conditional mean function to another. The idea of Markov switching as a tool to perform model comparison and selection between non nested nonlinear time series models.Finally, the MSA model can easily be generalized to the case of more than two states. The computational intensity involved increases rapidly,

4 Data Analysis

For a given US unemployment dataset, we try to fit a Markov chain Auto-regressive model and try to draw the inferences out of it.

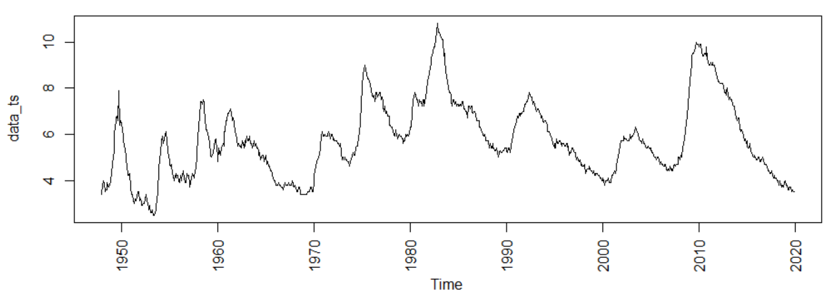

Figure 1: Time series Plot

The time series plot of the data tells us that there is no linear trend which is there and therefore we have to make use of non-linear approach to the given data.

5 Observation

When the Markov chain auto-regressive model is applied to the given data with 2 and 3 regimes, we get that for the 2 regimes the AIC of the model was 2317.143 and for 3 regimes the model AIC turned out to be 1779.035. So model with the three regimes is a better model.

The summary statistics of the model is given as follow:

```
Markov Switching Model

Call: msmFit(object = model, k = k, sw = sw, p = p)

       AIC       BIC    logLik
  1779.035  1848.173  -883.5173

Coefficients:

Regime 1
---------
                 Estimate  Std. Error  t value   Pr(>|t|)
(Intercept)(S)     5.8636      0.0680   86.229 < 2.2e-16 ***
time(S)           -0.0019      0.0001  -19.000 < 2.2e-16 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1

Residual standard error: 0.4536845
Multiple R-squared: 0.5536

Standardized Residuals:
          Min              Q1             Med              Q3
Max
-0.8779650374  -0.0015348190   0.0000312294   0.0316755213   0.8192793424

Regime 2
---------
                 Estimate  Std. Error  t value   Pr(>|t|)
(Intercept)(S)     6.3896      0.2123   30.097 < 2.2e-16 ***
time(S)            0.0028      0.0004    7.000  2.56e-12 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05    .    0.1

Residual standard error: 1.02517
Multiple R-squared: 0.2289

Standardized Residuals:
          Min              Q1             Med              Q3
```

```
Max
−1.487894507  −0.083802252  −0.005920802  −0.001640117   3.257124552

Regime 3
---------
                 Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S)     3.3165     0.0575  57.678 < 2.2e−16 ***
time(S)            0.0037     0.0001  37.000 < 2.2e−16 ***
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1

Residual standard error: 0.4960282
Multiple R−squared: 0.788

Standardized Residuals:
          Min             Q1            Med             Q3
Max
−1.0595442128  −0.0039671836   0.0001258661   0.0246440644   0.9900086670

Transition probabilities:
           Regime 1   Regime 2   Regime 3
Regime 1 0.96757316 0.01386148 0.01639057
Regime 2 0.01261539 0.97430005 0.01137781
Regime 3 0.01981146 0.01183847 0.97223162
```

5.1 Plots

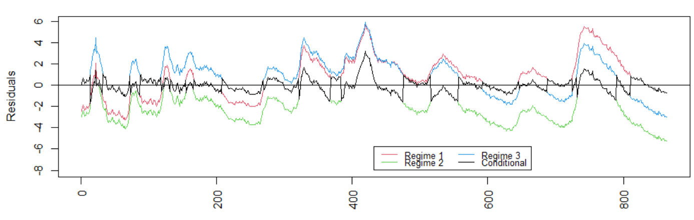

Figure 2: Residual v/s order graph

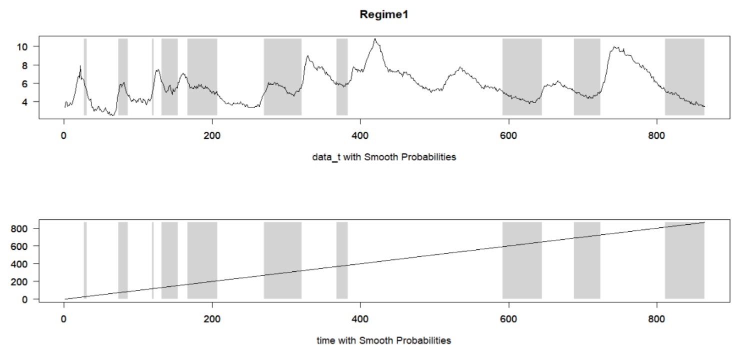

Figure 3: Regime 1

s

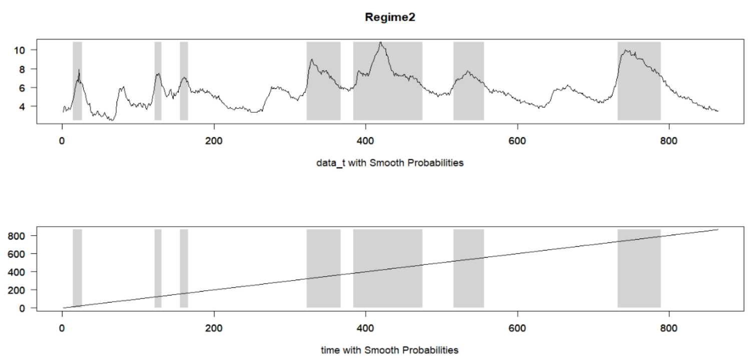

Figure 4: Regime 2

6 Result and conclusion

The model which we have fit to our US unemployment data has a regime where the intercept is coming out to be very significant and in the other regime the auto correlation variable is very significant too. In both, the R-squared have high values. Finally, the transition probabilities matrix has high values which indicate that is difficult to change from on regime to the other. The model detect perfectly the periods of each state.

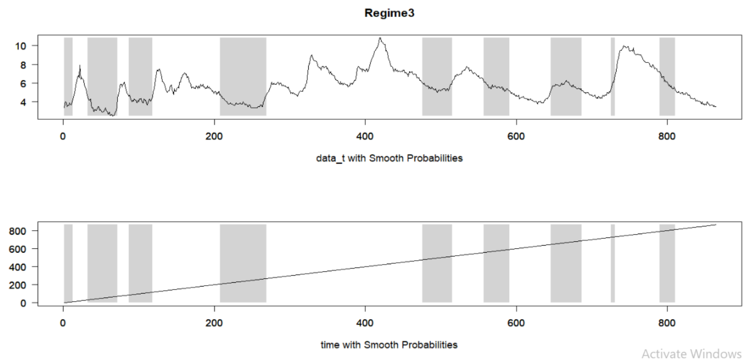

Figure 5: Regime 3

7 References

- Analysis of financial Time-series by Ruey S. Tsay

Appendix

Original Program:

```
library(MSwM)
model_lm=lm(data_t~month,data=data_main)
model_markov=msmFit(data_t~time,k=2,data = data_main,sw=c(TRUE,T,T))
summary(model_markov)
model_markov1=msmFit(data_t~time,k=3,data = data_main,sw=c(TRUE,T,T))
summary(model_markov1)

AIC(model_markov)
AIC(model_markov1)

plotProb(model_markov,which=1)
plotProb(model_markov,which=2)
plotProb(model_markov,which=3)

AIC(model_markov1)
plot(model_markov1)
plotDiag(model_markov1)
plotReg(model_markov1,regime = 1)
plotReg(model_markov1,regime = 2)
plotReg(model_markov1,regime = 3)
```