Econométrie de la finance

Chapitre 3 : Les modèles non liléaires univariés

Mohamed Essaied Hamrita

Octobre 2021

Introduction

- Dans ce chapitre nous introduisons quelques modèles non linéaires univariés et nous discutons leurs propriétes statistiques.
- Les modèles introduits sont : le modèle TAR (Threshold AR), le modèle Markov switching (MSM), smooth threshold autoregressive (STAR) models, et time-varying parameter models (TVM)

Le modèle TAR

- Le modèle TAR est proposé par (Tong 1978) et largement utilisé depuis la publication de l'article de (Tong et Lim 1980).
- Nous commençons par un modèle TAR simple à deux régimes, puis nous discutons le modèle TAR à régimes multiples.
- **Définition :** Une série temporelle $\{X_t\}$ suit un modèle TAR d'ordre p avec la variable seuil X_{t-d}, s'il vérifie la relation suivante :

$$X_t = \begin{cases} \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \sigma_1 \varepsilon_t \ , \ \text{si} \ \ X_{t-d} \leq r \\ \theta_0 + \sum_{i=1}^{p} \theta_i X_{t-i} + \sigma_2 \varepsilon_t \ \ , \ \text{si} \ \ X_{t-d} > r \end{cases}$$

  avec $\varepsilon_t \overset{iid}{\sim} (0,1)$, ϕ_i et θ_i des valeurs réelles telles que $\phi_i \neq \theta_i$ pour quelques valeurs de i. d est entier positif et r représente la valeur seuil (threshold).

Le modèle TAR à deux régimes peut s'écrire sous la forme suivante :

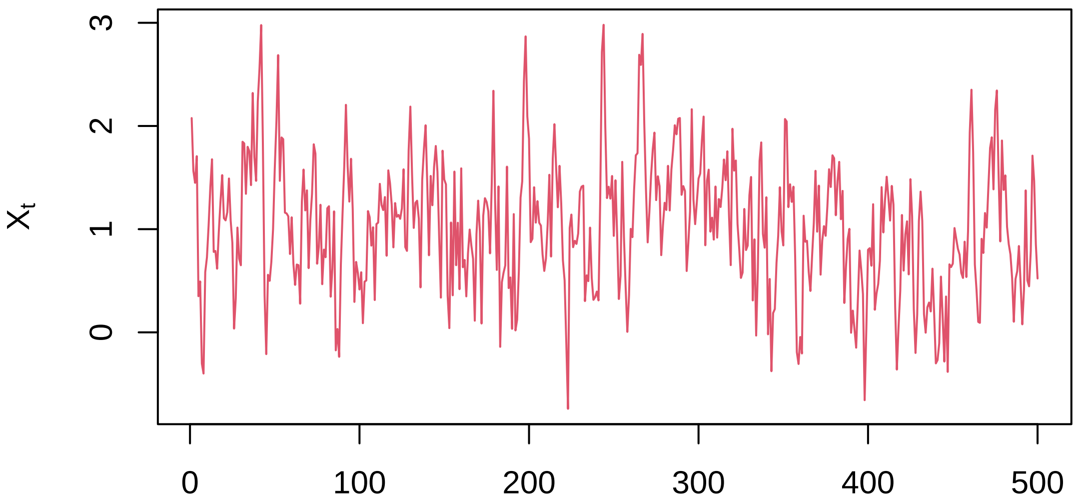
$$X_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \sigma_1 \varepsilon_t + I(X_{t-d} > r) \left(\beta_0 + \sum_{i=1}^{p} \beta_i X_{t-i} + \gamma \varepsilon_t \right)$$

où $I(X_{t-d} > r) = 1$ si $X_{t-d} > r$ et 0 sinon. $\beta_i = \theta_i - \phi_i$ pour $i = 0, 1, \dots, p$ et $\gamma = \sigma_2 - \sigma_1$

Exemple : Soit le processus $\{X_t\}$ définit par :

$$X_t = \begin{cases} 0.2 + 0.45 X_{t-1} + 0.2 \varepsilon_t \ , \ \text{si} \ \ X_{t-4} \leq -1 \\ 0.4 + 0.6 X_{t-1} + 0.5 \varepsilon_t \ \ , \ \text{si} \ \ X_{t-4} > 1 \end{cases}$$

```
library(TSA);
set.seed(12345)
x=tar.sim(n=500,Phi1=c(0.2,0.45), Phi2=c(0.4,0.6),p=1,
          d=4,sigma1=0.2,sigma2=0.5, thd=-1)$y
plot(x, type="l", col=2, xlab="", ylab=expression(X[t]))
```



Propriétés statistiques

Considérons le modèle TAR(1) à deux régime :

$$X_t = \begin{cases} \phi_1 X_{t-1} + \sigma_1 \varepsilon_t \ , \ \text{si} \ \ X_{t-1} \leq 0 \\ \theta_1 X_{t-1} + \sigma_2 \varepsilon_t \ \ , \ \text{si} \ \ X_{t-1} > 0 \end{cases}$$

La skeleton de ce modèle est

$$f(X_{t-1}) = \begin{cases} \phi_1 X_{t-1} \ , \ \text{si} \ \ X_{t-1} \leq 0 \\ \theta_1 X_{t-1} \ \ , \ \text{si} \ \ X_{t-1} > 0 \end{cases}$$

Puisque x_0 est supposée un réel quelconque, alors pour que la série X_t soit stable, il faut que $\phi_1 < 1$, $\theta_1 < 1$ et $\phi_1 \theta_1 < 1$.

Pour $d > 1$, (Chen et Tsay 1991) ont montré que les conditions de stabilité du modèle est : $\phi_1 < 1$, $\theta_1 < 1$ et $\phi_1^{s(d)} \theta_1^{t(d)}$ où $s(d) = t(d) = 1$ si $d = 1$ et $s(d) = 1$, $t(d) = 2$ si $d = 2$.

Estimation du modèle TAR

Le modèle TAR à deux régime peut être estimé soit par la méthode du maximum de vraisemblance, soit par la méthode des MCO. Ici, on donne la méthode des MCO. La matrice des variables explicatives peut être écrite de la forme suivante :

$$X_t(r) = (X_1' \, I(X_{t-d} \leq r) \ \ X_2' \, I(X_{t-d} > r)))$$

d'où, le modèle peut se réécrire comme :

$$X_t = X_t'(r)\alpha + \varepsilon_t \ \ \text{où} \ \ \alpha = (\phi' \ \ \theta')'$$

Pour r donnée, on obtient :

$$\hat{\alpha}(r) = \left(X_t'(r)X_t(r)\right)^{-1} X_t'(r)X_t$$

Estimation du paramètre r : L'estimation du vecteur α se fait pour r fixé. Or, r est généralement inconnu. Donc, ce paramètre sera estimé en faisant varier le paramètre r, soit une séquence de n valeurs. Pour chaque valeur de r, on calcul la somme carré des erreurs du modèle estimé, puis on retient la valeur du r qui correspond à la somme des carrés des erreurs la plus faible.

Estimation du paramètre d : Dans la pratique, le paramètre d est inconnu et on doit l'estimer. Ce paramètre sera estimé avec le vecteur α en imposant $d \in \{1, 2, \dots, \overline{d}\}$.

Test statistique TAR :

Une question importante est posée : quand le modèle TAR est statistiquement significative relativement à un modèle linéaire. Il s'agit de tester : $H_0 : \ \phi = \theta$. Puisque la valeur du seuil r est inconnue, un tel test devient délicat.

Pour plus amples discussions, voir (Chan 1990) et (Hansen 1997)

Prévision

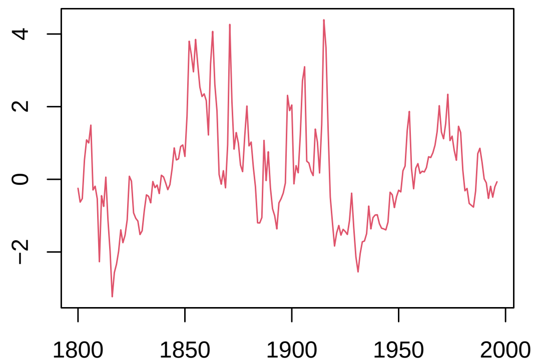
Pour $d \geq 1$ donné, la prévision à une étape est donnée par :

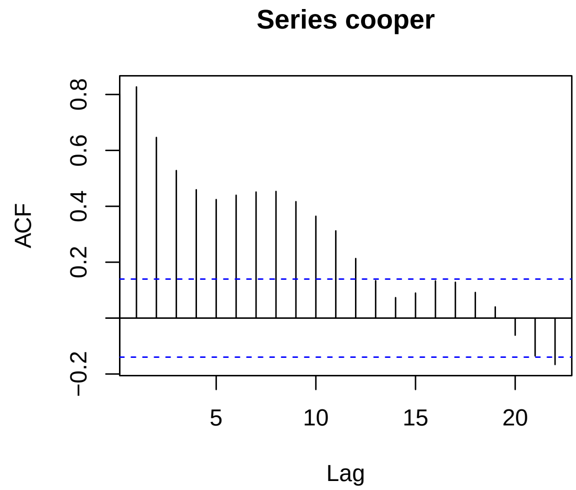
$$X_{T+1} = \begin{cases} \phi_0 + \displaystyle\sum_{i=1}^{p} \phi_i X_{T+1-i}, \quad e_{T+1} = \sigma_1 \varepsilon_{T+1}, \quad \text{si } X_{t-d} \leq r \\ \theta_0 + \displaystyle\sum_{i=1}^{p} \theta_i X_{T+1-i}, \quad e_{T+1} = \sigma_2 \varepsilon_{T+1}, \quad \text{si } X_{t-d} > r \end{cases}$$

Application

On veut modéliser le prix du Cooper. On considère la série des prix annuels du Cooper allant de 1800 à 1996 (Hyndman et Yang 2018). La série est ajustée après avoir éliminer la tendance. La figure suivante montre l'évolution des prix et la fonction d'auto-corrélation simple.

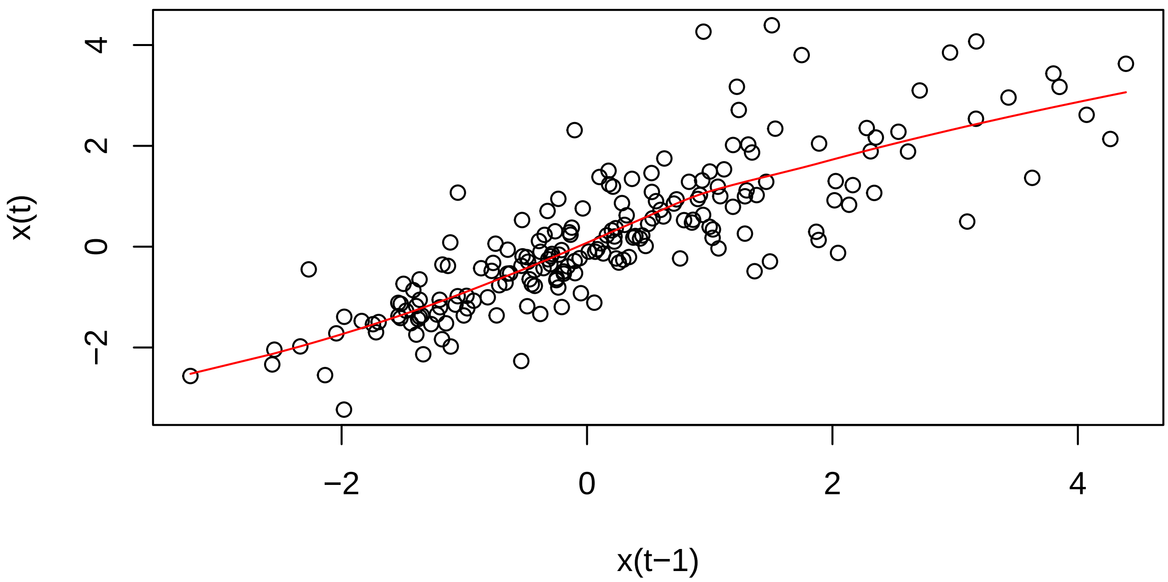
```
cooper=scan("https://raw.githubusercontent.com/Hamrita/TSFin/main/Chap3/copper.txt")
par(mfrow=c(1,2))
plot(1800:1996,cooper, type="l", col=2, xlab="", ylab="")
acf(cooper)
```





Un premier examen de non linéarité se fait en représentant x_t en fonction de x_{t-1}

```
y <- cooper[2:197]; x <- cooper[1:196]
m1 <- loess(y~x)  ## local smoothing
sx <- sort(x,index=T)  ## sorting the threshold variable
ix <- sx$ix ## index for order-statistics
 plot(x,y,xlab='x(t-1)',ylab='x(t)')
lines(x[ix],m1$fitted[ix],col="red")
```



On remarque bien que la dépendance entre x_t et x_{t-1} est non linéaire.

Tests de linéarité

```
library(nonlinearTseries)
tests=nonlinearityTest(cooper, F)
tests$TarTest
```

```
$percentiles
[1] 24.72527 75.27473

$test.statistic
[1] 66.50607

$p.value
[1] 1.896341e-06
```

On a p-value qui est inférieur à 5%, donc on rejette l'hypothèse nulle de linéarité du modèle.

Estimation

```
library(tsDyn)
mod=setar(cooper,m=3,d=1,mL=2,mH=2)
ss=summary(mod)
ss$coef
```

```
           Estimate  Std. Error    t value     Pr(>|t|)
const.L  0.11254434  0.06677038  1.6855430 9.351685e-02
phiL.1   0.98690602  0.10805163  9.1336526 9.634921e-17
phiL.2  -0.02297053  0.09154768 -0.2509133 8.021508e-01
const.H  0.62130257  0.34080242  1.8230580 6.985800e-02
phiH.1   0.80837079  0.14826866  5.4520680 1.526336e-07
phiH.2  -0.33916240  0.10905935 -3.1098883 2.158115e-03
```

```
ss$AIC
```

```
[1] -94.53252
```

```
ss$thCoef
```

```
      th
1.221279
```

Références

Chan, K. S. 1990. « Testing for Threshold Autoregression ». *The Annals of Statistics* 18 (4) : 1886-94. https://doi.org/10.1214/aos/1176347886.

Chen, Rong, et Ruey S. Tsay. 1991. « On the Ergodicity of Tar(1) Processes ». *The Annals of Applied Probability* 1 (4) : 613-34.

Hansen, Bruce. 1997. « Inference in TAR Models ». *Studies in Nonlinear Dynamics & Econometrics* 2 (1) : 1-16.

Hyndman, Rob, et Yangzhuoran Yang. 2018. « tsdl : Time Series Data Library. v0.1.0 ».

Tong, Howell. 1978. « Pattern Recognition and Signal Processing. NATO ASI Series E : Applied Sc ». In, édité par C Chen. Sijthoff & Noordhoff, Netherlands.

Tong, Howell, et K. S. Lim. 1980. « Threshold Autoregression, Limit Cycles and Cyclical Data ». *Journal of the Royal Statistical Society : Series B (Methodological)* 42 (3) : 245-68. https://doi.org/10.1111/j.2517-6161.1980.tb01126.x.