Package MTS

Jules CORBEL & Paul GUILLOTTE

04/02/2019

Nous nous intéresserons dans ce document à la mise en place de modèles VAR à l'aide du package **MTS** afin de prédire la masse salariale trimestrielle. Un modèle VAR, pour Vecteur AutoRégressif, a pour objectif de capturer les interdépendances entre les différentes séries temporelles à notre disposition. Ainsi, chaque variable est expliquée par ses propres valeurs passées ainsi que par les valeurs passées des autres variables du modèle.

Visualisation des séries

Même contenu que pour le package vars

Transformation des séries

Même contenu que pour le package vars

Calcul de l'ordre p

Afin de mettre en place une modélisation VAR, nous devons dans un premier temps nous intéresser à l'ordre p du modèle VAR. L'ordre p correspond à l'ordre de l'opérateur de retard, c'est-à-dire le nombre de valeurs du passé qui ont un impact sur la valeur à un instant défini. Dans le package **MTS**, la fonction utilisée est VARorder, qui comme VARselect utilise les critères d'AIC, BIC et HQ afin de déterminer l'ordre du processus. Toutefois, le critère FPE n'est pas présent, ce qui nous conforte dans notre idée qu'il n'est pas très utile à notre étude. Les valeurs des différents indicateurs nous font retenir un ordre p=3.

```
selec <- VARorder(cbind(MSEStaTrain, PIBStaTrain, SMICStaTrain, TCHOFStaTrain), maxp=10)
```

```
## selected order: aic =  10
## selected order: bic =  0
## selected order: hq =  3
## Summary table:
##        p      AIC      BIC       HQ    M(p) p-value
##  [1,]  0 -31.0096 -31.0096 -31.0096  0.0000  0.0000
##  [2,]  1 -30.9214 -30.5096 -30.7546 19.5070  0.2432
##  [3,]  2 -31.5822 -30.7587 -31.2487 80.4006  0.0000
##  [4,]  3 -31.9307 -30.6954 -31.4305 51.9850  0.0000
##  [5,]  4 -32.0595 -30.4124 -31.3925 32.9649  0.0075
##  [6,]  5 -32.0057 -29.9469 -31.1721 18.3312  0.3049
##  [7,]  6 -31.9908 -29.5202 -30.9904 19.8690  0.2262
##  [8,]  7 -31.9619 -29.0796 -30.7948 17.8037  0.3355
##  [9,]  8 -31.8948 -28.6007 -30.5609 14.4236  0.5672
## [10,]  9 -32.1777 -28.4719 -30.6771 32.5182  0.0086
## [11,] 10 -32.2875 -28.1699 -30.6202 21.3890  0.1640
```

Estimation du modèle

Dans le package **MTS**, la fonction utilisée pour construire des modèles VAR est *VAR*, qui prend en entrée la série temporelle multivariée et l'ordre du processus. On affiche ci-dessous les résultats renvoyés par la fonction sur le modèle

```
modele<-VAR(cbind(MSEStaTrain, PIBStaTrain, SMICStaTrain, TCHOFStaTrain), p=3, output=F)
print("Coefficients du modèle")
```

```
## [1] "Coefficients du modèle"
```

```
coeff<-modele$coef[2:5,]
colnames(coeff)<-rownames(coeff)<-c("MSE", "PIB", "SMIC", "TCHOF")
coeff
```

```
##               MSE           PIB        SMIC      TCHOF
## MSE   -0.56278414 -0.0168586446 -0.02625032 -0.3580146
## PIB    0.64149028  0.1814898925  0.48689607 -6.7304243
## SMIC   0.00298799 -0.0002301839 -0.53899891  0.1622353
## TCHOF -0.00732917  0.0011657405 -0.01560563 -0.2811625
```

```
print("Erreurs standard des coefficients")
```

```
## [1] "Erreurs standard des coefficients"
```

```
se<-modele$secoef[2:5,]
colnames(se)<-rownames(se)<-c("MSE", "PIB", "SMIC", "TCHOF")
se
```

```
##              MSE         PIB       SMIC     TCHOF
## MSE   0.09953303 0.013493075 0.20448277 0.8112122
## PIB   0.78217596 0.106034738 1.60691891 6.3748752
## SMIC  0.04600804 0.006237024 0.09451989 0.3749738
## TCHOF 0.01283070 0.001739379 0.02635967 0.1045725
```

```
matrix(data=c(modele$aic, modele$bic, modele$hq), nrow=1, byrow=T, dimnames=list(NULL, c("AIC", "BIC",
```

```
##            AIC       BIC        HQ
## [1,] -31.79783 -30.56255 -31.29762
```

Afin de vérifier que les résidus du modèle ne présentent ni d'auto-corrélation (corrélation dans une série entre deux temps différents), ni de corrélation croisée (corrélation entre deux séries différentes).

```
LjungBox(modele$residuals, 1:10)
```

```
##  lags  statistic  df   p-value
##     1   6.566401  16 0.9807419
##     2  11.053772  32 0.9997889
##     3  17.856956  48 0.9999783
##     4  38.659275  64 0.9949102
##     5  60.230726  80 0.9515649
##     6  77.266279  96 0.9196279
##     7  98.322899 112 0.8182994
##     8 117.417567 128 0.7384538
##     9 136.629432 144 0.6563374
##    10 149.286954 160 0.7174559
```

L'hypothèse nulle d'indépendance des résidus est conservée, ce qui nous permet de valider notre modèle.

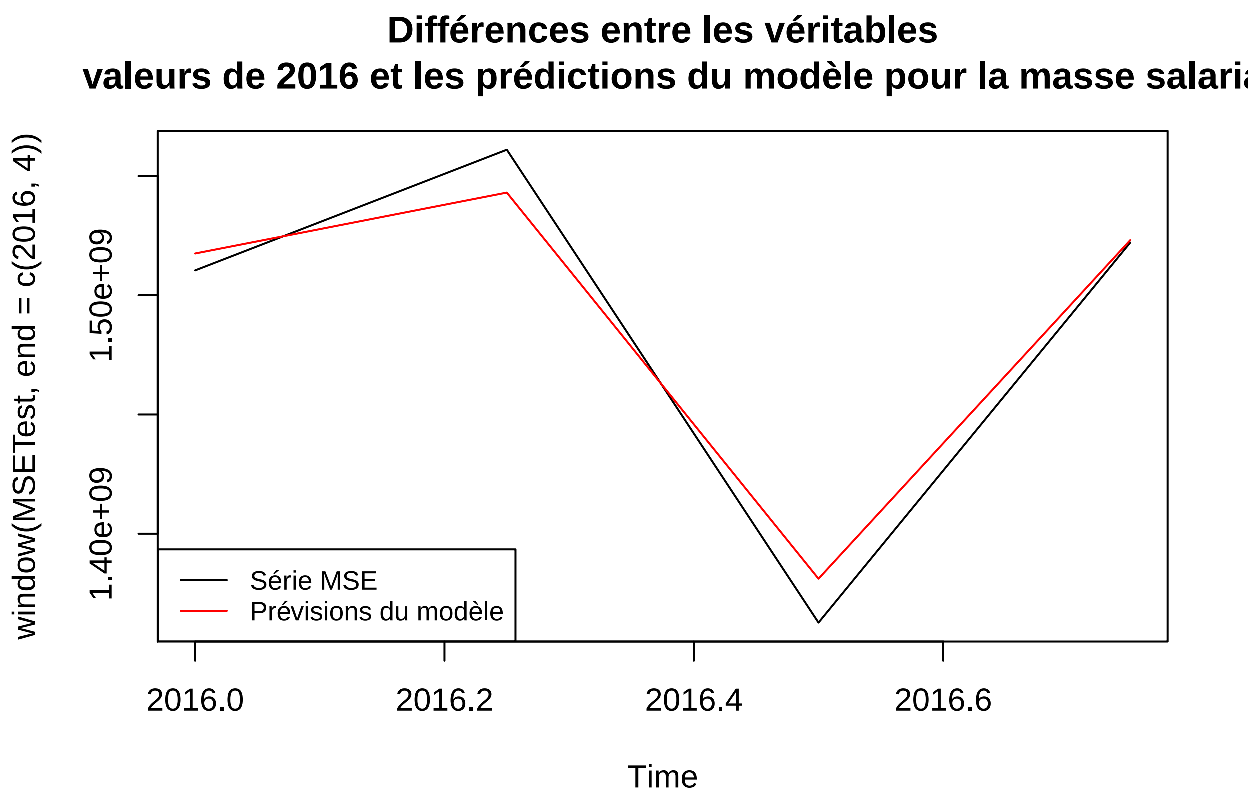
Prévisions

Maintenant que nous avons estimé l'ordre des différents modèle VAR, et que nous avons explicité l'estimation des modèles, nous cherchons désormais à trouver celui dont les prédictions sont les plus proches de la réalité.

Après avoir comparé tous les modèles possibles (7 : 3 modèles avec deux variables, 3 modèles avec trois variables et un modèle avec les quatre variables), nous nous apercevons que le meilleur en terme de prédictions est le modèle prenant en compte le SMIC et le PIB, avec un ordre égal à 4.

```
#SMIC
VARorder(cbind(MSEStaTrain, SMICStaTrain, PIBStaTrain), maxp=10)
modele<-VAR(cbind(MSEStaTrain, SMICStaTrain, PIBStaTrain), p=4)
pred <- VARpred(modele, 4)
```

```
plot(window(MSETest, end=c(2016,4)), main="Différences entre les véritables
     valeurs de 2016 et les prédictions du modèle pour la masse salariale")
#Reconstruction de la variable stationnaire
recon <- ts(pred$pred[,1], start=2016, frequency=4) * MSETrendTest * MSESeasonalTest
lines(recon, col = "red")
legend('bottomleft', legend = c('Série MSE', 'Prévisions du modèle'),
       col=c('black', 'red'), lty=1, cex=0.8)
```



Nous nous intéressons donc à l'erreur quadratique moyenne de cette prévision, qui est inférieure à celle obtenue pour le meilleur modèle effectué avec le package **vars**.

```
EQM(MSETest, recon)
```

```
## [1] 1.188978e+14
```