

Chapitre II

Exercice sur le PIB des Etats-Unis

Importation des données

On importe les données trimestrielles du PIB des Etats-Unis du 01/01/1947 au 23/01/01 (305 observations). Les données sont importées du site db.nomics. Les données proviennent du “Bureau of Labor Statistics”. On leur importe en utilisant leur identifiant ids.

```
library(data.table)
library(rdbnomics)

df<-rdb(ids = "BEA/NIPA-T10106/A191RX-Q")
#class(df)
str(df)

## Classes 'data.table' and 'data.frame': 305 obs. of 19 variables:
## $ @frequency : chr "quarterly" "quarterly" "quarterly" "quarterly" ...
## $ concept : chr "gross-domestic-product" "gross-domestic-product" "gross-domestic-product" ...
## $ Concept : chr "Gross domestic product" "Gross domestic product" "Gross domestic product" ...
## $ dataset_code : chr "NIPA-T10106" "NIPA-T10106" "NIPA-T10106" "NIPA-T10106" ...
## $ dataset_name : chr "Table 1.1.6. Real Gross Domestic Product, Chained Dollars - LastRevised: J" ...
## $ FREQ : chr "Q" "Q" "Q" "Q" ...
## $ Frequency : chr "Quarterly" "Quarterly" "Quarterly" "Quarterly" ...
## $ indexed_at : POSIXct, format: "2023-06-30 07:44:41" "2023-06-30 07:44:41" ...
## $ metric : chr "millions-of-chained-dollars" "millions-of-chained-dollars" "millions-of-chained-dollars" ...
## $ Metric : chr "Millions of chained Dollars" "Millions of chained Dollars" "Millions of chained Dollars" ...
## $ original_period: chr "1947-Q1" "1947-Q2" "1947-Q3" "1947-Q4" ...
## $ original_value : chr "2034450" "2029024" "2024834" "2056508" ...
## $ period : Date, format: "1947-01-01" "1947-04-01" ...
## $ provider_code : chr "BEA" "BEA" "BEA" "BEA" ...
## $ series_code : chr "A191RX-Q" "A191RX-Q" "A191RX-Q" "A191RX-Q" ...
## $ series_name : chr "Gross domestic product (line 1) - Quarterly" "Gross domestic product (line 1) - Quarterly" ...
## $ unit : chr "level" "level" "level" "level" ...
## $ Unit : chr "Level" "Level" "Level" "Level" ...
## $ value : num 2034450 2029024 2024834 2056508 2087442 ...
## - attr(*, ".internal.selfref")=<externalptr>
```

On crée un dataframe contenant uniquement les dates (colonne period) à partir du premier trimestre de l'année 1990 et la série du PIB (colonne value).

```
df_Y_us<-df[df$period>="1990-01-01",c("period","value")]
colnames(df_Y_us)<-c('Date','PIB')

head(df_Y_us)
```

```
##      Date      PIB
## 1: 1990-01-01 9364259
## 2: 1990-04-01 9398243
## 3: 1990-07-01 9404494
```

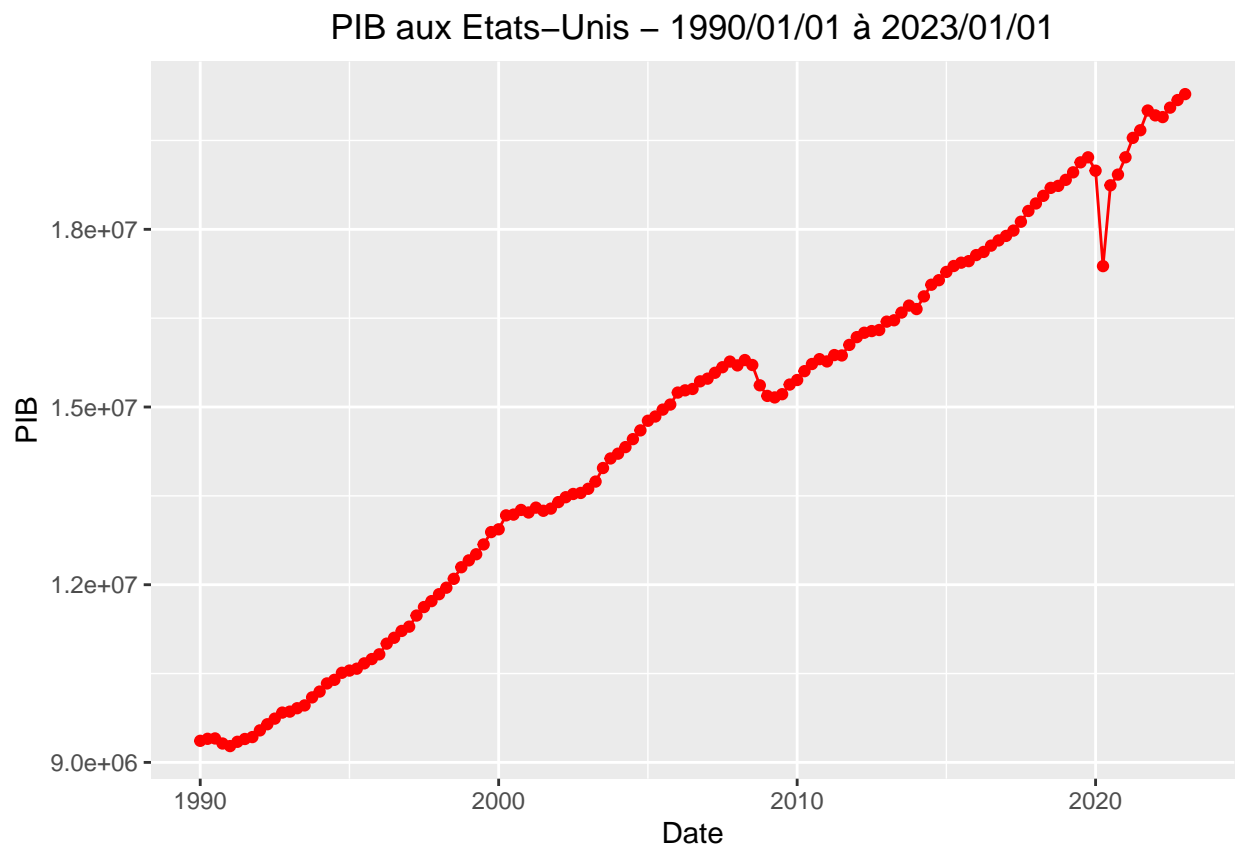
```
## 4: 1990-10-01 9318876
## 5: 1991-01-01 9275276
## 6: 1991-04-01 9347597
```

```
tail(df_Y_us)
```

```
##      Date      PIB
## 1: 2021-10-01 20006181
## 2: 2022-01-01 19924088
## 3: 2022-04-01 19895271
## 4: 2022-07-01 20054663
## 5: 2022-10-01 20182491
## 6: 2023-01-01 20282760
```

```
library(ggplot2)
```

```
p_us<-ggplot(data=df_Y_us,aes(x=Date,y=PIB))+geom_point(color='red')+geom_line(color='red')+xlab('Date')
p_us
```

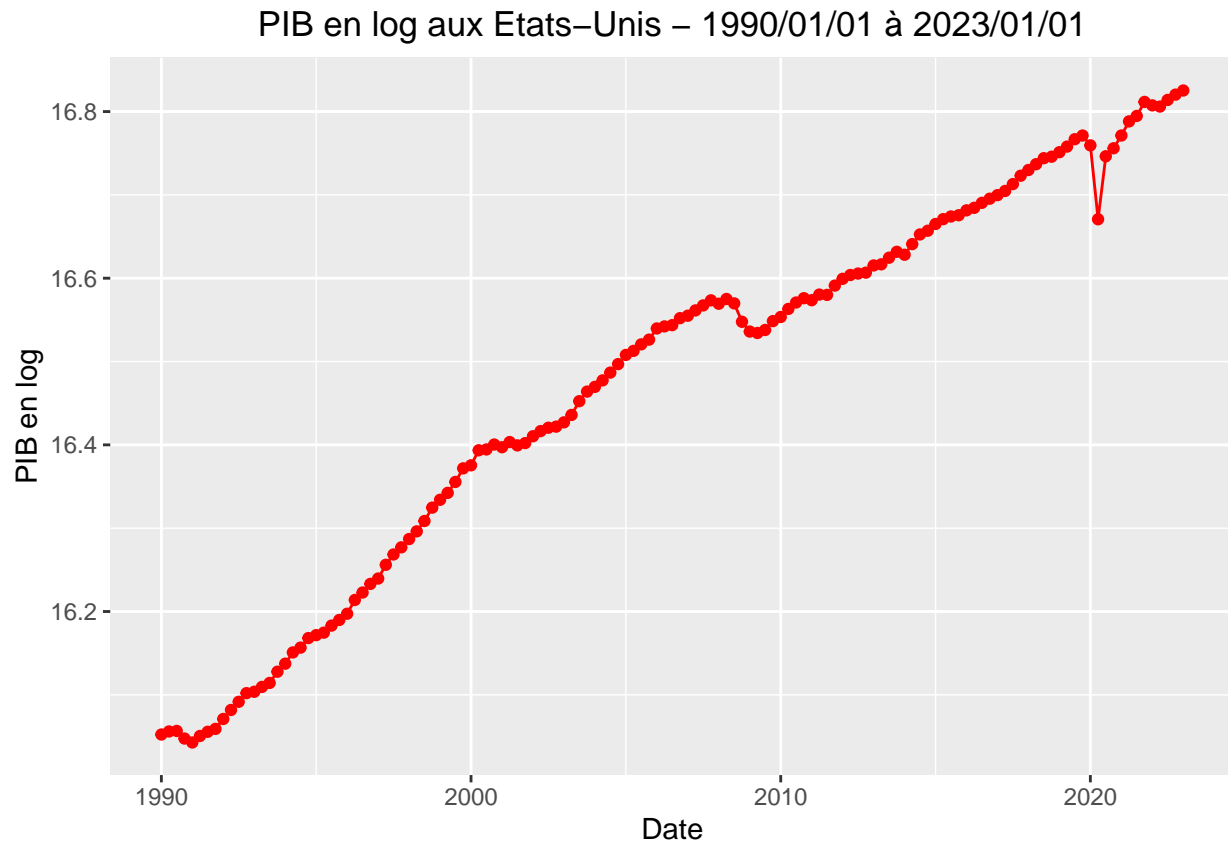


Le graphique montre que le PIB augmente de façon tendancielle au cours de la période, avec cependant une baisse de niveau en 2009, due à la grande récession et un creux très prononcé au moment de déclenchement de l'épidémie de COVID. On va exclure dans un premier temps la période du COVID de l'échantillon et l'inclure par la suite pour voir son impact sur les résultats de tests de racine unitaire et de stationnarité.

Visualisation du PIB en logarithme

On applique la fonction logarithme au PIB. Les tests de racine unitaire et de stationnarité seront appliqués au PIB en logarithme.

```
df_Y_us$logPIB<-log(df_Y_us$PIB)
library(ggplot2)
p_us<-ggplot(data=df_Y_us,aes(x=Date,y=logPIB))+geom_point(color='red')+geom_line(color='red')+xlab('Date')
p_us
```



Tests sans la période Covid

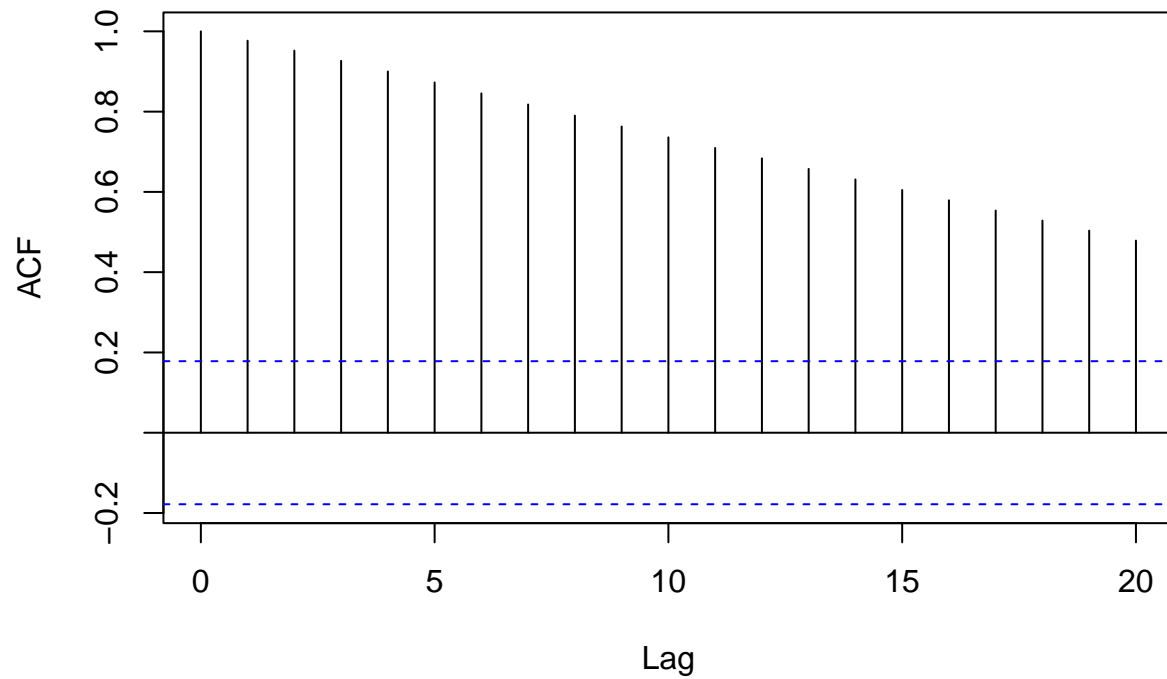
On exclut les observations postérieures au premier semestre de l'année 2020. Les résultats des tests de racine unitaire et de stationnarité ne sont modifiés. On peut dire aussi qu'ils sont robustes à cette modification de la période de test.

```
df_Y_us_bc=df_Y_us[df_Y_us$Date<"2020-04-01",]
```

Autocorrélogramme du taux de chômage

```
acf_us_rate<-acf(x=df_Y_us_bc$logPIB,main='Autocorrélogramme du PIB (en log) des USA 1990-Q1 à 2023-Q1')
```

Autocorrélogramme du PIB (en log) des USA 1990–Q1 à 2023–Q1

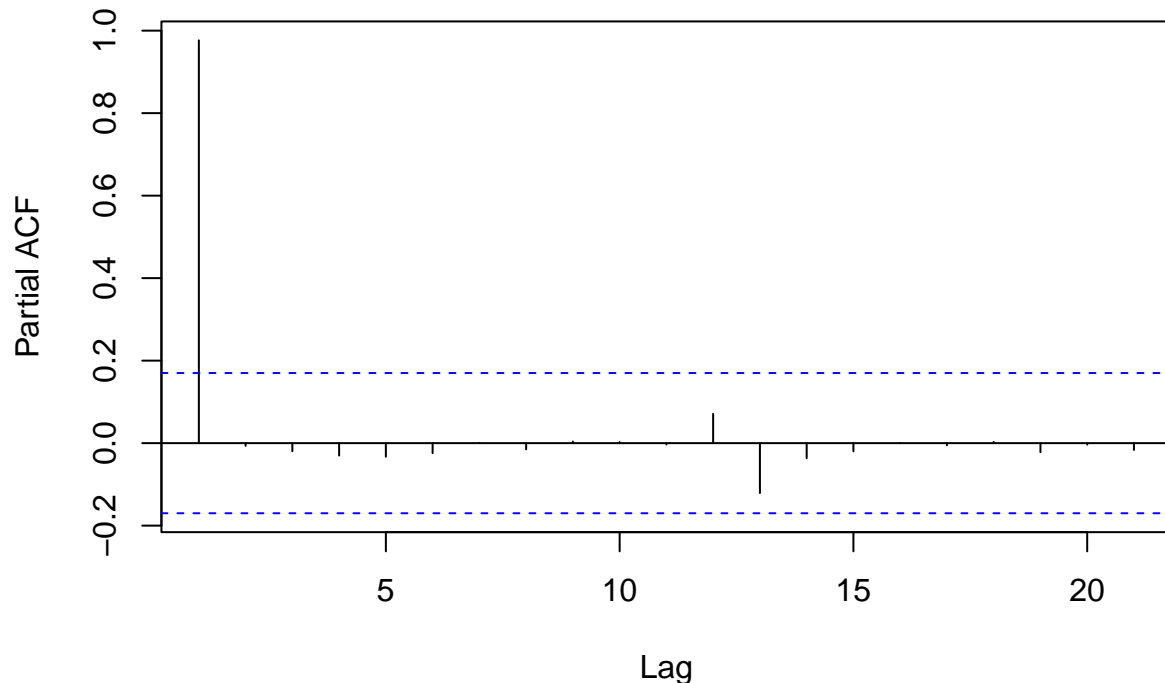


```
#acf_us_rate$acf
```

Les autocorrélations sont toutes significatives et proches de 1 pour les premiers décalages. Elles décroissent relativement lentement.

```
pacf(x=df_Y_us$logPIB,main='Autocorrélations partielles du PIB des USA, 1990-Q1 à 2023-Q1')
```

Autocorrélations partielles du PIB des USA, 1990–Q1 à 2023–Q1



Aucune autocorrélation partielle n'est significativement différente de 0.

Test de Dickey-Fuller augmenté

Test ADF avec constante et tendance déterministe

On applique le test ADF avec une tendance déterministe. Le nombre de retards maximum est fixé à 8 et le nombre de retards optimal est déterminé par la minimisation du critère AIC.

```
library(urca)
adf1<-ur.df(y = df_Y_us$logPIB,type=c("trend"), lag=8, selectlags = c('AIC'))
summary(adf1)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.096751 -0.002460  0.000902  0.003590  0.048920
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.9052059  0.4289022   2.111  0.0369 *
## z.lag.1      -0.0555629  0.0266440  -2.085  0.0392 *
## tt           0.0002647  0.0001539   1.720  0.0880 .
## z.diff.lag   -0.1745390  0.0883950  -1.975  0.0506 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01177 on 120 degrees of freedom
## Multiple R-squared:  0.08485,    Adjusted R-squared:  0.06197
## F-statistic: 3.709 on 3 and 120 DF,  p-value: 0.01356
##
##
## Value of test-statistic is: -2.0854 14.5209 3.5916
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

La statistique du test ADF est égale à $t_{ADF} = -2.085$. Les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire.

Le t-stat de la tendance déterministe est égal à $t_b = 1.720$. Le seuil de rejet qui figurent la table de t_b est égal à 3.14 pour un risque de première espèce de 5% et 100 observations. on a donc $|t_b| = 1.720 < 3.14$: on ne peut pas rejeter l'hypothèse nulle que la tendance déterministe n'est pas significative.

Test ADF avec constante

```
adf2<-ur.df(y = df_Y_us$logPIB,type=c("drift"), lag=8, selectlags = c('AIC'))
summary(adf2)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.095440 -0.002564  0.000972  0.003821  0.053087
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.182096  0.085818   2.122  0.0359 *
## z.lag.1      -0.010597  0.005197  -2.039  0.0436 *
```

```
## z.diff.lag -0.192696 0.088470 -2.178 0.0313 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01186 on 121 degrees of freedom
## Multiple R-squared: 0.06228, Adjusted R-squared: 0.04678
## F-statistic: 4.018 on 2 and 121 DF, p-value: 0.02043
##
##
## Value of test-statistic is: -2.0388 19.9783
##
## Critical values for test statistics:
##      1pct 5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1 6.52 4.63 3.81
```

a statistique du test ADF est égale à $t_{ADF} = -2.039$. Les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire.

Le t-stat de la constante est égal à $t_{\hat{c}} = 2.122$. Le seuil de rejet de la table de $t_{\hat{c}}$ est égal à 2.86 pour un risque de première espèce de 5% et $T=100$ observations. on a donc $|t_{\hat{c}}| = 2.122 < 2.86$: on ne peut pas rejeter l'hypothèse nulle que la contante n'est pas significative.

Test ADF sans constante ni tendance déterministe

```
adf3<-ur.df(y = df_Y_us$logPIB,type=c("none"), lag=8, selectlags = c('AIC'))
summary(adf3)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.097903 -0.002073  0.000540  0.004264  0.053246
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      4.307e-04  7.336e-05  5.871 3.83e-08 ***
## z.diff.lag -1.713e-01  8.915e-02 -1.921  0.057 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01203 on 122 degrees of freedom
## Multiple R-squared: 0.2235, Adjusted R-squared: 0.2108
## F-statistic: 17.56 on 2 and 122 DF, p-value: 1.982e-07
##
```

```
##
## Value of test-statistic is: 5.8707
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

La statistique du test ADF est égale à $t_{ADF} = 5.871$. Les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire. On conclut que la série n'est pas stationnaire.

Conclusion

Les résultats des tests ADF nous amènent à conclure que le PIB (en log) des Etats-Unis est une série DS sans tendance ni constante. On dit aussi qu'elle est un processus DS sans dérive (drift).

Test de stationnarité de KPSS

On applique le test de stationnarité de KPSS. On présente les résultats du test avec une tendance déterministe. La représentation graphique du PIB montre que la série pourrait être stationnaire autour d'une tendance déterministe.

```
kpss_mu<-ur.kpss(y=df_Y_us$logPIB,type="tau",lags="short")
summary(kpss_mu)
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: tau with 4 lags.
##
## Value of test-statistic is: 0.5664
##
## Critical value for a significance level of:
##      10pct  5pct 2.5pct  1pct
## critical values 0.119 0.146  0.176 0.216
```

La statistique du test KPSS avec une constante est égale à $LM_{KPSS} = 0.5664$. Elle est supérieure au seuil de rejet à 5% :

$$0.5664 > 0.146$$

(et aussi à celle à 1%). On rejette l'hypothèse nulle de stationnarité du taux du PIB autour d'une tendance déterministe.

```
kpss_tau<-ur.kpss(y=df_Y_us$logPIB,type="mu",lags="short")
summary(kpss_tau)
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: mu with 4 lags.
##
## Value of test-statistic is: 2.659
```



```
##
## Critical value for a significance level of:
##          10pct  5pct 2.5pct  1pct
## critical values 0.347 0.463  0.574 0.739
```

Les résultats des tests de stationarité autour d'une constante montre que l'on rejette également l'hypothèse de stationnarité.

Conclusion générale sur les test de racine unitaire et de stationnarité

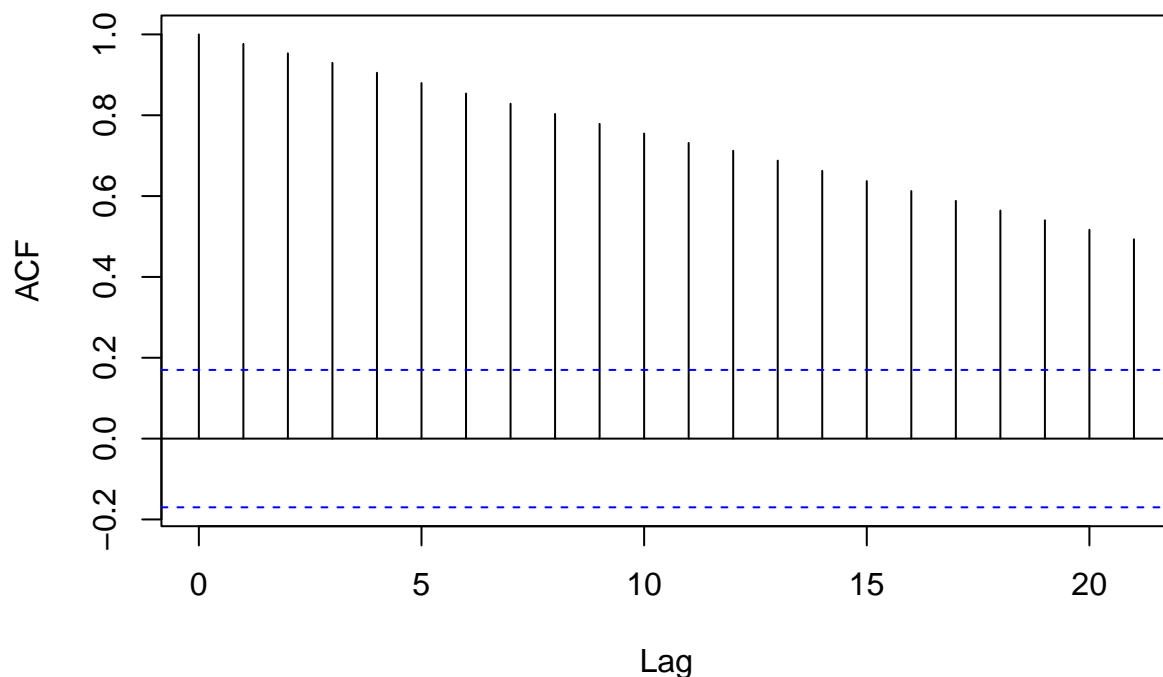
Les tests ADF et KPSS rejettent tous la stationnarité du PIB des Etats-Unis. On étudie l'impact de la chute brutale et temporaire du PIB lors du début de l'épidémie de COVID sur les résultats des tests de racine unitaire et de stationnarité.

Test sur la période 1990:Q1 à 2023:Q1

Autocorrélogramme du taux de chômage

```
acf_us_rate<-acf(x=df_Y_us$logPIB,main='Autocorrélogramme du PIB des Etats-Unis - 1990:Q1 à 2023:Q1')
```

Autocorrélogramme du PIB des Etats-Unis – 1990:Q1 à 2023:Q1



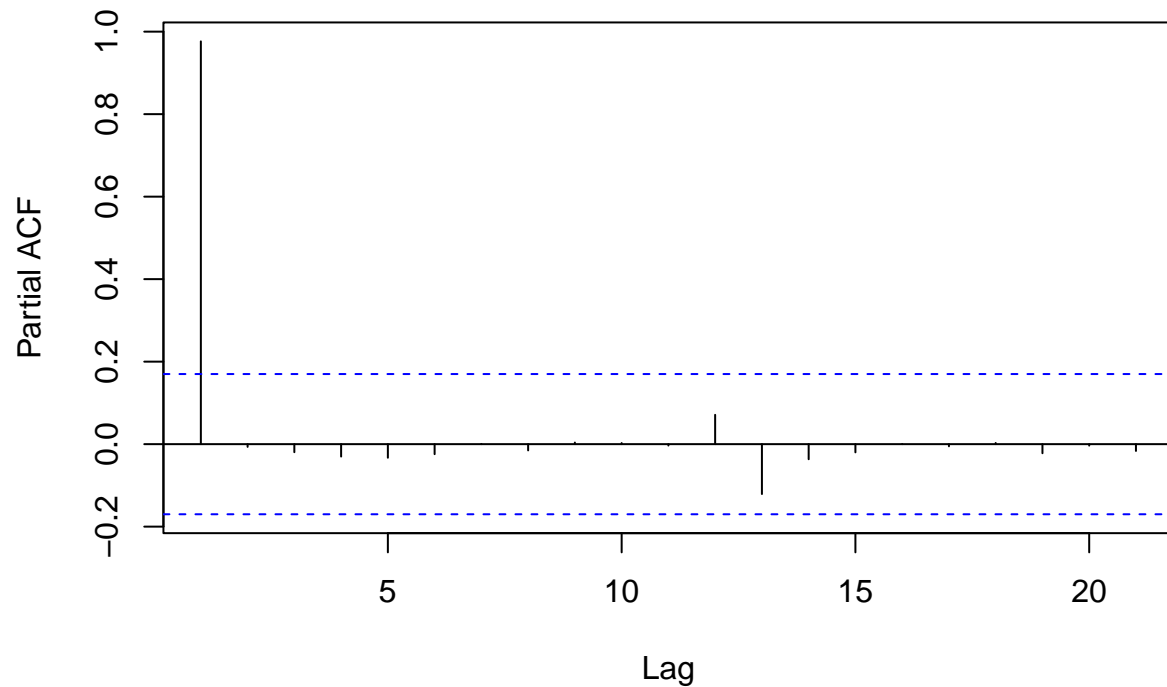
```
acf_us_rate$acf
```

```
## , , 1
##
##      [,1]
## [1,] 1.0000000
```

```
## [2,] 0.9764676
## [3,] 0.9531802
## [4,] 0.9295363
## [5,] 0.9050636
## [6,] 0.8796625
## [7,] 0.8537588
## [8,] 0.8285111
## [9,] 0.8031527
## [10,] 0.7786043
## [11,] 0.7548138
## [12,] 0.7315116
## [13,] 0.7121717
## [14,] 0.6877868
## [15,] 0.6625408
## [16,] 0.6371693
## [17,] 0.6124874
## [18,] 0.5880739
## [19,] 0.5642230
## [20,] 0.5401761
## [21,] 0.5165653
## [22,] 0.4930008
```

```
pacf(x=df_Y_us$logPIB,main='Autocorrélogramme partiel du PIB des Etats-Unis - 1990:Q1 à 2023:Q1')
```

Autocorrélogramme partiel du PIB des Etats-Unis – 1990:Q1 à 2023:Q1



Test de Dickey-Fuller augmenté

On applique le test ADF avec une tendance déterministe. Le nombre de retards maximum est fixé à 8 et le nombre de retards optimal est déterminé par la minimisation du critère AIC.

```
library(urca)
adf1<-ur.df(y = df_Y_us$logPIB,type=c("trend"), lag=8, selectlags = c('AIC'))
summary(adf1)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.096751 -0.002460  0.000902  0.003590  0.048920
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.9052059  0.4289022   2.111  0.0369 *
## z.lag.1      -0.0555629  0.0266440  -2.085  0.0392 *
## tt           0.0002647  0.0001539   1.720  0.0880 .
## z.diff.lag   -0.1745390  0.0883950  -1.975  0.0506 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01177 on 120 degrees of freedom
## Multiple R-squared:  0.08485,    Adjusted R-squared:  0.06197
## F-statistic: 3.709 on 3 and 120 DF,  p-value: 0.01356
##
##
## Value of test-statistic is: -2.0854 14.5209 3.5916
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

La statistique du test ADF est égale à $t_{ADF} = -2.085$. Les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire.

Le t-stat de la tendance déterministe est égal à $t_{\hat{\delta}} = 1.720$. Le seuil de rejet qui figurent la table de $t_{\hat{\delta}}$ est égal à 3.14 pour un risque de première espèce de 5%. on a donc $|t_{\hat{\delta}}| = 1.720 < 3.14$: on ne peut pas rejeter l'hypothèse nulle que la tendance déterministe n'est pas significative.

```
adf2<-ur.df(y = df_Y_us$logPIB,type=c("drift"), lag=8, selectlags = c('AIC'))
summary(adf2)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.095440 -0.002564  0.000972  0.003821  0.053087
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.182096   0.085818   2.122   0.0359 *
## z.lag.1      -0.010597   0.005197  -2.039   0.0436 *
## z.diff.lag   -0.192696   0.088470  -2.178   0.0313 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01186 on 121 degrees of freedom
## Multiple R-squared:  0.06228,    Adjusted R-squared:  0.04678
## F-statistic: 4.018 on 2 and 121 DF,  p-value: 0.02043
##
##
## Value of test-statistic is: -2.0388 19.9783
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

a statistique du test ADF est égale à $t_{ADF} = -2.0388$ les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire.

Le t-stat de la constante est égal à $t_{\hat{c}} = 2.122$. Le seuil de rejet de la table de $t_{\hat{c}}$ est égal à 2.86 pour un risque de première espèce de 5%. on a donc $|t_{\hat{c}}| = 2.122 < 2.86$: on ne peut pas rejeter l'hypothèse nulle que la contante n'est pas significative.

```
adf3<-ur.df(y = df_Y_us$logPIB,type=c("none"), lag=8, selectlags = c('AIC'))
summary(adf3)
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.097903 -0.002073  0.000540  0.004264  0.053246
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1      4.307e-04  7.336e-05   5.871 3.83e-08 ***
## z.diff.lag -1.713e-01  8.915e-02  -1.921   0.057 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01203 on 122 degrees of freedom
## Multiple R-squared:  0.2235, Adjusted R-squared:  0.2108
## F-statistic: 17.56 on 2 and 122 DF,  p-value: 1.982e-07
##
##
## Value of test-statistic is: 5.8707
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

La statistique du test ADF est égale à $t_{ADF} = 5.870$. Les seuils de rejet figurent sur la ligne tau3. La statistique de test t_{ADF} est supérieure aux seuils de rejets pour les risques de première espèce de 1%, 5% et 10%. On ne peut pas rejeter l'hypothèse nulle de racine unitaire. On conclut que la série n'est pas stationnaire.

Test de stationnarité de KPSS

On applique le test de stationnarité de KPSS. On présente les résultats du test avec une constante. La représentation graphique de la série et les conclusions du test ADF montrent que la taux de chômage ne contient pas de tendance déterministe. On ne rejette pas l'hypothèse nulle de stationnarité autour d'une constante.

```
kpss_tau<-ur.kpss(y=df_Y_us$logPIB,type="tau",lags="short")
summary(kpss_tau)
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: tau with 4 lags.
##
## Value of test-statistic is: 0.5664
##
## Critical value for a significance level of:
##      10pct  5pct 2.5pct  1pct
## critical values 0.119 0.146  0.176 0.216
```

On rejette à nouveau l'hypothèse de stationnarité autour d'une tendance déterministe.

```
kpss_mu<-ur.kpss(y=df_Y_us$logPIB,type="mu",lags="short")
summary(kpss_mu)
```

```
##
```

```
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: mu with 4 lags.
##
## Value of test-statistic is: 2.659
##
## Critical value for a significance level of:
##          10pct  5pct  2.5pct  1pct
## critical values 0.347 0.463  0.574 0.739
```

On rejette l'hypothèse de stationnarité autour d'une constante.

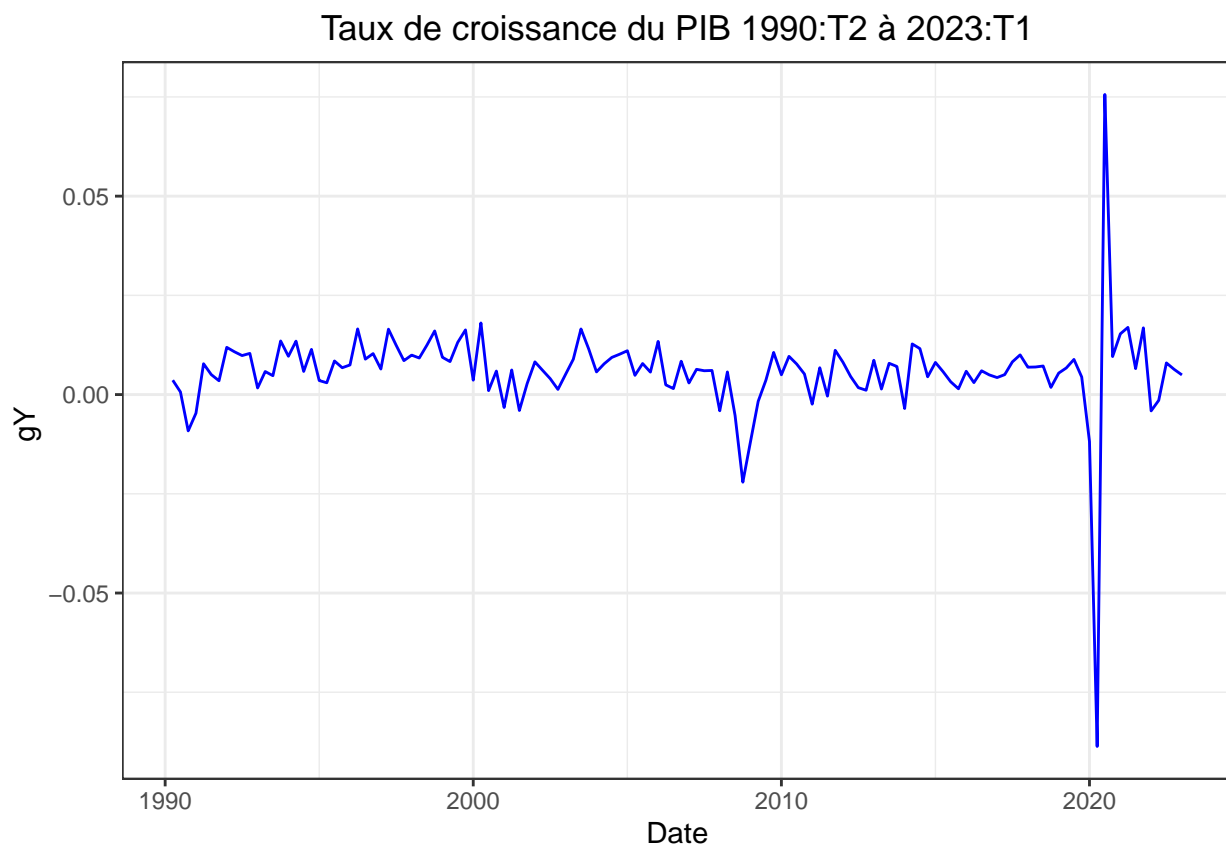
Etude du PIB en différence première

La différence première du PIB en log correspond à son taux de croissance (voir l'exercice sur le CAC40 pour une explication).

```
df_gY_us<-data.frame(df_Y_us$Date[-1],diff(df_Y_us$logPIB))
colnames(df_gY_us)<-c('Date','gY')
```

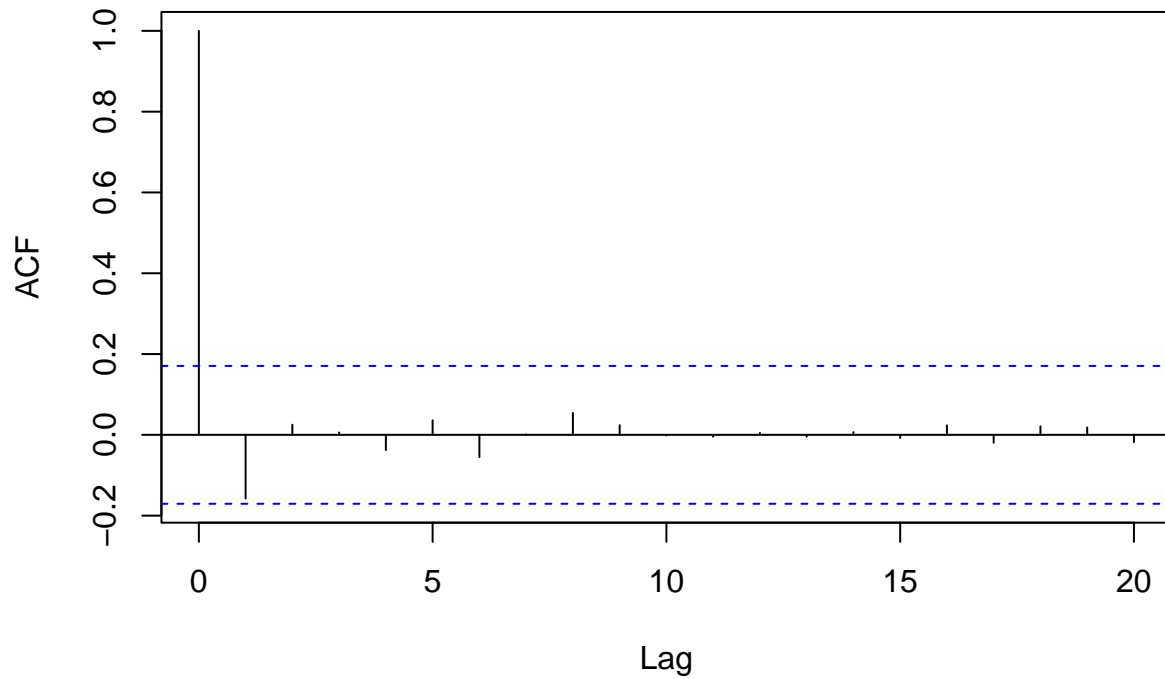
Représentation graphique du taux de croissance du PIB des Etats-Unis

```
p<-ggplot(data=df_gY_us,aes(x=Date,y=gY))+geom_line(colour='blue')+ggtitle('Taux de croissance du PIB 1990:T2 à 2023:T1')
p
```



```
acf(x = df_gY_us$gY, lag.max = 20, main="Autocorrélogramme du taux de croissance du PIB, 1990:T2 à 2023:T")
```

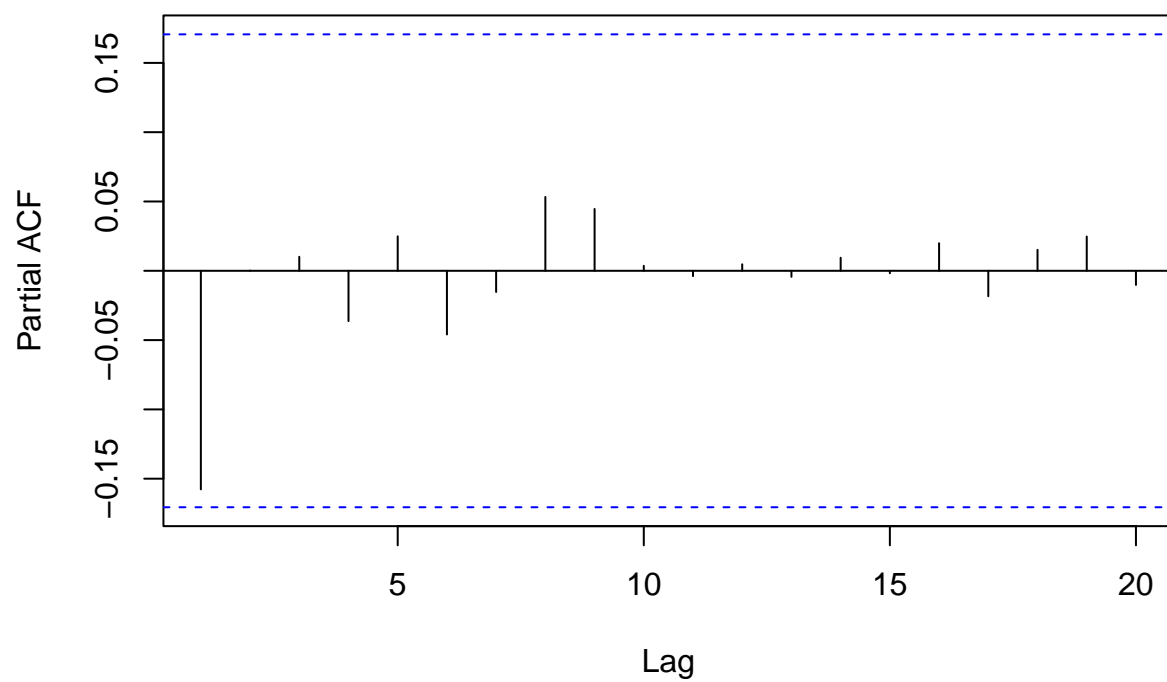
Autocorrélogramme du taux de croissance du PIB, 1990:T2 à 2023:T



Aucune des autocorrélations estimées n'est significativement différente de 0.

```
pacf(x = df_gY_us$gY, lag.max = 20, main='Autocorrelations partielles du PIB des Etats-Unis, 1990:T2 à 2023:T')
```

Autocorrelations partielles du PIB des Etats-Unis, 1990:T2 à 2023:T



Aucune des autocorrélations partielles n'est significativement différente de 0.