

Processus ARMA stationnaires

Chapitre 2

Économétrie des séries temporelles

Structure du cours

Processus linéaires

Décomposition de Wold

Modèles de séries temporelles univariées : $AR(p)$, $MA(q)$

Équations de Yule Walker

Décomposition de prédiction

Estimation du maximum de vraisemblance

Exemple empirique

1. Processus linéaires

Supposons que $\{y_t\}$ est un processus stochastique :

Opérateur retard

$$\begin{aligned}Ly_t &= y_{t-1} \\ L^j y_t &= y_{t-j} \quad \forall j \in \mathbb{N}\end{aligned}$$

Opérateur différence

$$\begin{aligned}\Delta y_t &= (1 - L) y_t = y_t - y_{t-1} \\ \Delta^j y_t &= (1 - L)^j y_t \quad \forall j \in \mathbb{N}_+ \\ \Delta_s y_t &= (1 - L^s) y_t = y_t - y_{t-s}\end{aligned}$$

Processus linéaires (suite)

Filtre linéaire transforme une série d'entrée $\{x_t\}$ en une série de sortie $\{y_t\}$ en utilisant un polynôme de retard $A(L)$:

$$\begin{aligned}y_t &= A(L)x_t = \left(\sum_{-n}^m a_j L^j \right) x_t = \sum_{j=-n}^m a_j x_{t-j} \\&= a_{-n} x_{t+n} + \cdots + a_0 x_t + \cdots + a_m x_{t-m}\end{aligned}$$

Processus linéaire

$$y_t = A(L)\varepsilon_t = \left(\sum_{-\infty}^{\infty} a_j L^j \right) \varepsilon_t = \sum_{j=-\infty}^{\infty} a_j \varepsilon_{t-j} \quad (1)$$

où $\varepsilon_t \sim \text{WN}(0, \sigma^2)$

Note - pour $|x| < 1$:

$$1 + x + x^2 + \cdots + x^n = \sum_{i=0}^n x^i = \frac{1 - x^{n+1}}{1 - x} \rightarrow \frac{1}{1 - x} = \sum_{i=0}^{\infty} x^i \quad (2)$$

2. Décomposition de Wold

La **décomposition de Wold** - tout processus stationnaire de covariance à moyenne nulle $\{y_t\}$ peut être représenté sous la forme :

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \kappa_t$$

où $\psi_0 = 1$ et $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. ε_t est un bruit blanc ; il représente l'erreur commise dans la prévision de y_t sur la base d'une fonction linéaire de son passé $Y_{t-1} = y_{t-j}$ pour $j \geq 1$:

$$\varepsilon_t \equiv y_t - \widehat{E}(y_t | Y_{t-1})$$

Notez que $\text{corr}(\kappa_t, \varepsilon_{t-j}) = 0, \forall j$, mais :

$$\kappa_t = \widehat{E}(\kappa_t | Y_{t-1})$$

\widehat{E} indique une projection linéaire sur un vecteur de variables aléatoires Y_t .

Approche de Box-Jenkins

Approximation du polynôme à retard infini avec le rapport de deux polynômes d'ordre fini $\alpha(L)$ et $\beta(L)$:

$$\Psi(L) = \sum_{j=0}^{\infty} \psi_j L^j \simeq \frac{\beta(L)}{\alpha(L)} = \frac{1 + \beta_1 L + \dots + \beta_q L^q}{1 - \alpha_1 L - \dots - \alpha_p L^p}$$

Types de modèles de séries temporelles

Type	Modèle	p	q
AR(p)	$\alpha(L)y_t = \varepsilon_t$	$p > 0$	$q = 0$
MA(q)	$y_t = \beta(L)\varepsilon_t$	$p = 0$	$q > 0$
ARMA(p, q)	$\alpha(L)y_t = \beta(L)\varepsilon_t$	$p > 0$	$q > 0$

3. Processus AR(1)

Le processus AR(1) satisfait l'équation différentielle :

$$y_t = \nu + \alpha y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (3)$$

En utilisant l'opérateur de retard :

$$(1 - \alpha L)y_t = \nu + \varepsilon_t. \quad (4)$$

Lorsque $|\alpha| < 1$:

$$(1 - \alpha L)^{-1} = \lim_{j \rightarrow \infty} (1 + \alpha L + \alpha^2 L^2 + \dots + \alpha^j L^j) \quad (5)$$

Par conséquent :

$$y_t = (1 + \alpha + \alpha^2 + \dots) \nu + (\varepsilon_t + \alpha \varepsilon_{t-1} + \alpha^2 \varepsilon_{t-2} + \dots) \quad (6)$$

En prenant les espérances avec $|\alpha| < 1$:

$$E(y_t) = \frac{\nu}{1 - \alpha} = \mu \quad (7)$$

Analyse de stabilité basée sur une équation différentielle linéaire non homogène :

$$(y_t - \mu) = \alpha(y_{t-1} - \mu) + \varepsilon_t \quad (8)$$

Par conséquent :

$$E(y_t - \mu)^2 = \alpha^2 E(y_{t-1} - \mu)^2 + E(\varepsilon_t^2) + 2\alpha E[(y_{t-1} - \mu)\varepsilon_t] \quad (9)$$

Sous la condition $|\alpha| < 1$:

$$E(y_t - \mu)^2 = E(y_{t-1} - \mu)^2 = \gamma(0) \quad (10)$$

de sorte que :

$$V(y_t) = \gamma(0) = \frac{\sigma^2}{1 - \alpha^2} \quad (11)$$

Fonction d'autocovariance :

$$E[(y_t - \mu)(y_{t-1} - \mu)] = \alpha E(y_{t-1} - \mu)^2 + E[(y_{t-1} - \mu)\varepsilon_t] \quad (12)$$

Par conséquent :

$$\gamma(1) = \alpha\gamma(0). \quad (13)$$

Aussi :

$$E[(y_t - \mu)(y_{t-2} - \mu)] = \alpha E[(y_{t-1} - \mu)(y_{t-2} - \mu)] + E[(y_{t-2} - \mu)\varepsilon_t] \quad (14)$$

pour obtenir :

$$\gamma(2) = \alpha\gamma(1). \quad (15)$$

En général :

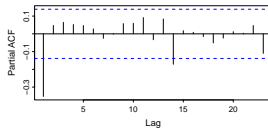
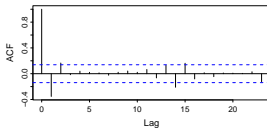
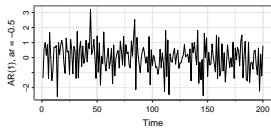
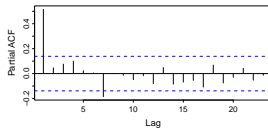
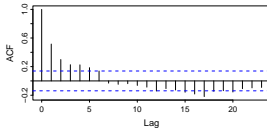
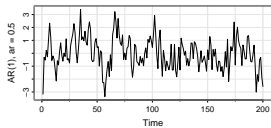
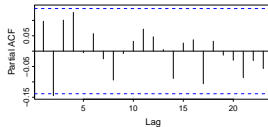
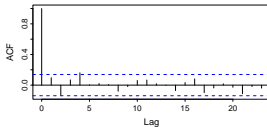
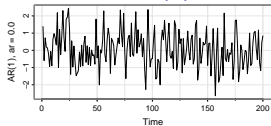
$$\gamma(h) = \alpha\gamma(h-1) = \alpha^h\gamma(0) \quad (16)$$

pour $h \neq 0$.

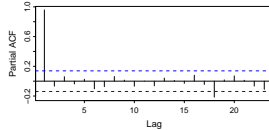
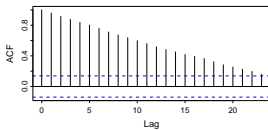
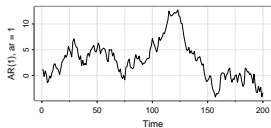
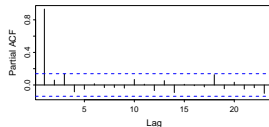
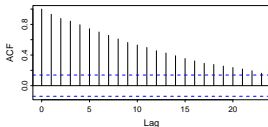
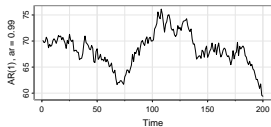
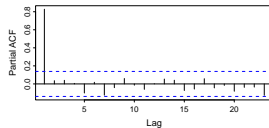
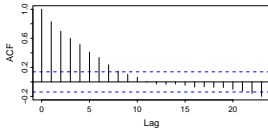
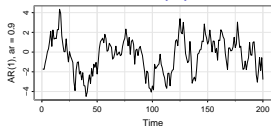
Coefficient d'autocorrélation :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\alpha^h \gamma(0)}{\gamma(0)} = \alpha^h \quad (17)$$

Processus AR(1) stationnaires



Processus AR(1) persistents et non stationnaires



4. Processus AR(2)

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (18)$$

En supposant la stationnarité :

$$E(y_t) = \frac{\nu}{1 - \alpha_1 - \alpha_2} \quad (19)$$

Aussi :

$$V(y_t) = \gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \sigma^2 \quad (20)$$

Et :

$$\gamma(1) = \alpha_1 \gamma(0) + \alpha_2 \gamma(1)$$

$$\gamma(2) = \alpha_1 \gamma(1) + \alpha_2 \gamma(0)$$

En résolvant pour $\gamma(0)$:

$$\gamma(0) = \frac{(1 - \alpha_2)\sigma^2}{(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \quad (21)$$

Conditions pour la stationnarité :

$$\alpha_2 + \alpha_1 < 1; \quad \alpha_2 - \alpha_1 < 1; \quad |\alpha_2| < 1 \quad (22)$$

Les racines complexes apparaissent si :

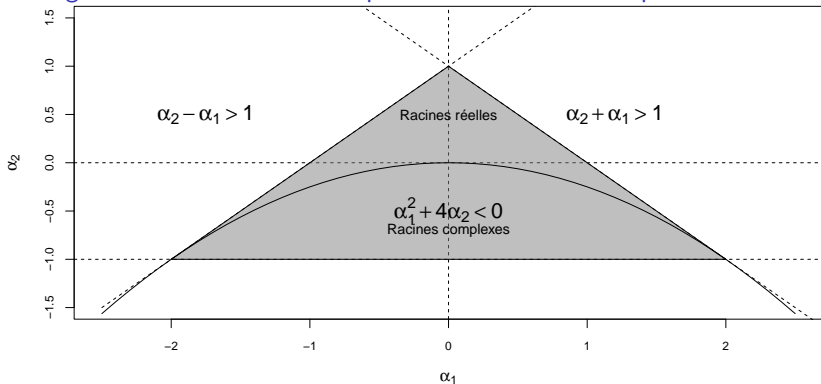
$$\alpha_1^2 + 4\alpha_2 < 0 \quad (23)$$

Équations de Yule-Walker :

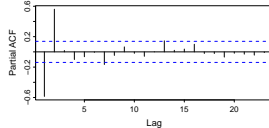
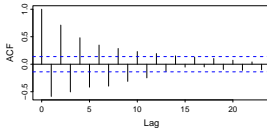
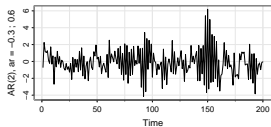
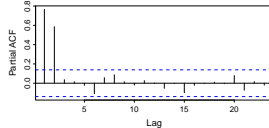
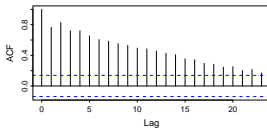
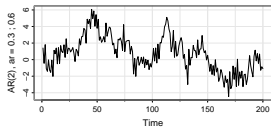
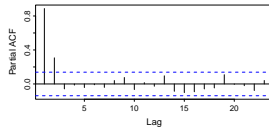
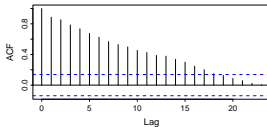
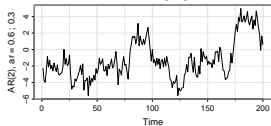
$$\rho(1) = \alpha_1 + \alpha_2 \rho(1)$$

$$\rho(2) = \alpha_1 \rho(1) + \alpha_2$$

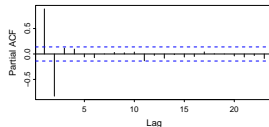
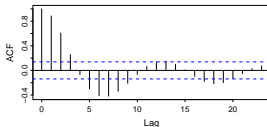
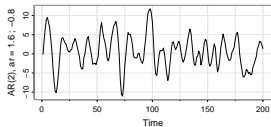
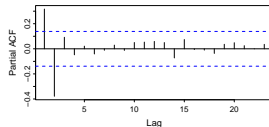
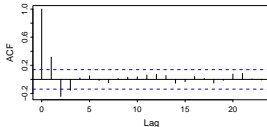
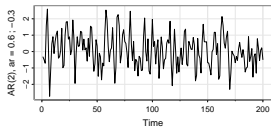
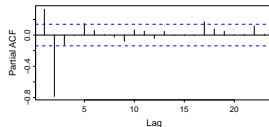
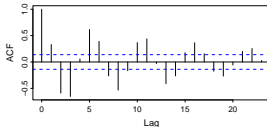
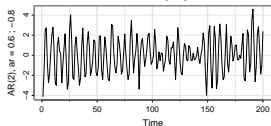
Triangle de stationnarité avec séparation des racines complexes et réelles



Processus AR(2) avec racines réelles



Processus AR(2) avec racines complexes



5. Processus AR(p)

$$y_t = \nu + \sum_{j=1}^p \alpha_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (1) \quad (24)$$

Stabilité : $\alpha(z) = 0 \Rightarrow |z| > 1$ garantit la stationnarité et la représentation MA(∞) :

$$y_t = \alpha(1)^{-1} \nu + \alpha(L)^{-1} \varepsilon_t$$

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

où $\mu = \frac{\nu}{\alpha(1)}$ et $\Psi(L) = \alpha(L)^{-1}$ avec $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

En prenant les espérances de l'équation (1) :

$$E(y_t) = \mu = \frac{\nu}{(1 - \alpha_1 - \dots - \alpha_p)} \quad (25)$$

Fonction d'autocovariance :

$$\begin{aligned}\gamma(h) &= E(y_t y_{t-h}) = E[(\alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t) y_{t-h}] \\ &= \alpha_1 E(y_{t-1} y_{t-h}) + \cdots + \alpha_p E(y_{t-p} y_{t-h}) + E(\varepsilon_t y_{t-h}) \\ &= \alpha_1 \gamma(h-1) + \cdots + \alpha_p \gamma(h-p)\end{aligned}$$

Équations de Yule-Walker :

$$\rho(1) = \alpha_1 + \alpha_2 \rho(1) + \cdots + \alpha_p \rho(p-1)$$

$$\rho(2) = \alpha_1 \rho(1) + \alpha_2 + \cdots + \alpha_p \rho(p-2)$$

$$\vdots$$

$$\rho(p) = \alpha_1 \rho(p-1) + \alpha_2 \rho(p-2) + \cdots + \alpha_p$$

pour ρ_1, \dots, ρ_p

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \cdots + \alpha_p \rho(k-p)$$

pour $k > p$

6. Processus de moyenne mobile

$$y_t = \mu + \beta(L)\varepsilon_t = \mu + \sum_{i=1}^q \beta_i \varepsilon_{t-i} + \varepsilon_t \quad (26)$$

où $\varepsilon_t \sim \text{WN}(0, \sigma^2)$ et $\beta(L) = 1 + \beta_1 L + \dots + \beta_q L^q$, $\beta_q \neq 0$.

Le processus $\text{MA}(q)$ est stationnaire pour tout $(\beta_1, \beta_2, \dots, \beta_q)$.

Inversibilité : $\beta(z) = 0 \Rightarrow |z| > 1$ garantit la représentation $\text{AR}(\infty)$:

$$\begin{aligned} \beta(L)^{-1} y_t &= \beta(1)^{-1} \mu + \varepsilon_t \\ y_t &= \mu + \sum_{j=1}^{\infty} \phi_j (y_{t-j} - \mu) + \varepsilon_t \end{aligned}$$

où $\phi(L) = 1 - \sum_{j=1}^{\infty} \phi_j L^j = 1 - \phi_1 L - \phi_2 L^2 - \dots = \beta(L)^{-1}$.

Fonction d'autocovariance :

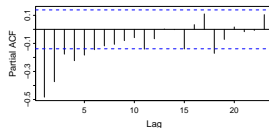
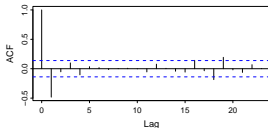
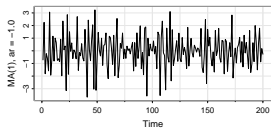
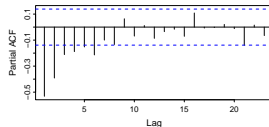
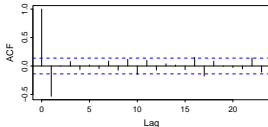
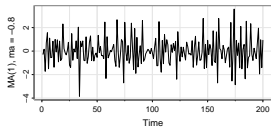
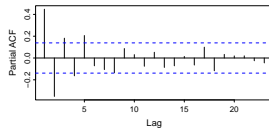
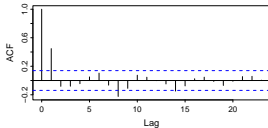
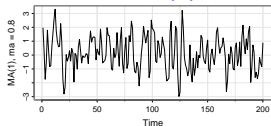
$$\gamma(0) = \left(\sum_{i=0}^q \beta_i^2 \right) \sigma^2$$

$$\gamma(k) = \left(\sum_{i=0}^{q-k} \beta_i \beta_{i+k} \right) \sigma^2 \quad \text{pour } k = 1, 2, \dots, q$$

$$\gamma(k) = 0 \quad \text{pour } k > q$$

Les processus MA sont stationnaires et ergodiques.

Processus MA(1)



7. Processus ARMA(p, q)

$$\alpha(L)y_t = \nu + \beta(L)\varepsilon_t$$

$$y_t = \nu + \sum_{j=1}^p \alpha_j y_{t-j} + \varepsilon_t + \sum_{i=1}^q \beta_i \varepsilon_{t-i}$$

où $\varepsilon_t \sim \text{WN}(0, \sigma^2)$;

$$\alpha(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p, \alpha_p \neq 0 ;$$

$$\beta(L) = 1 + \beta_1 L + \dots + \beta_q L^q, \beta_q \neq 0.$$

1. Stabilité

$\alpha(z) = 0 \Rightarrow |z| > 1$ garantit la stationnarité et la représentation MA(∞) :

$$y_t = \alpha(1)^{-1} \nu + \alpha(L)^{-1} \beta(L) \varepsilon_t$$

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

2. Inversibilité

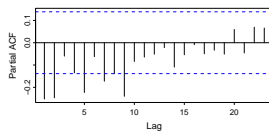
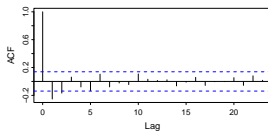
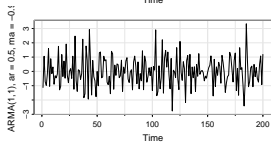
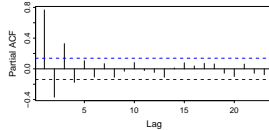
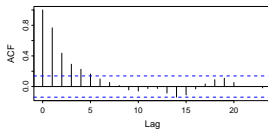
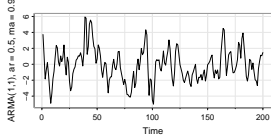
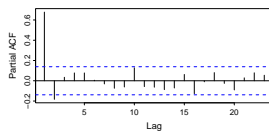
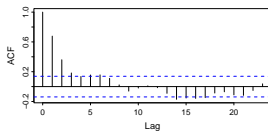
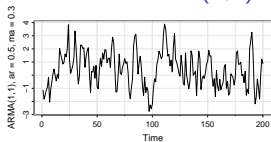
$\beta(z) = 0 \Rightarrow |z| > 1$ permet la représentation $AR(\infty)$:

$$\begin{aligned}\beta(L)^{-1}\alpha(L)(y_t - \mu) &= \varepsilon_t \\ y_t &= \mu + \sum_{j=1}^{\infty} \phi_j(y_{t-j} - \mu) + \varepsilon_t\end{aligned}$$

3. Pas de racines communes dans $\alpha(L)$ et $\beta(L)$

$$\begin{aligned}\alpha(L) &= \prod_{j=1}^p (1 - \lambda_j L) \\ \beta(L) &= \prod_{i=1}^q (1 - \mu_i L) \\ \Rightarrow \lambda_j &\neq \mu_i \quad \forall i, j\end{aligned}$$

Processus ARMA(1,1)



8. Formulation statistique du modèle AR(1)

Indépendance conditionnelle : $(y_t | y_0, \dots, y_{t-1}) \stackrel{d}{=} (y_t | y_{t-1})$;

Distribution conditionnelle : $(y_t | y_{t-1}) \stackrel{d}{=} \mathcal{N}(\nu + \alpha y_{t-1}, \sigma^2)$; $t \geq 1$;

Espace des paramètres : $\nu, \alpha, \sigma^2 \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$.

L'observation initiale y_0 n'est pas modélisée – conditionnelle à y_0 .

Le régresseur est la variable dépendante retardée.

$$y_t = \nu + \alpha y_{t-1} + \varepsilon_t \quad \text{pour } t = 1, \dots, T.$$

où $\varepsilon_t \sim \text{NID}(0, \sigma^2)$ sont des innovations.

Dans la pratique (modélisation - données)

y_t sont les rendements de l'action Apple sur trois années autour de la période covid.

```
AAPL <- getSymbols(Symbols = 'AAPL', src = 'yahoo',  
                  from='2018-12-01', to='2020-12-01', auto.assign = FALSE)  
AAPL$Date = as.Date.character(AAPL$Date, format = "%Y/%m/%d") #changement  
# du format dans la dataframe  
AAPL.date <- AAPL$Date #création de la variable date (non utilisée ici)  
AAPL.p <- AAPL$AAPL.Adjusted #création de la variable prix  
AAPL.r = na.omit(diff(log(AAPL.p))*100) #création de la variable rendements  
# stationnaires  
summary(cbind(AAPL.p[-1], AAPL.r))
```

##	Index	AAPL.Adjusted	AAPL.Adjusted.1
##	Min. :2018-12-04	Min. : 33.95	Min. : -13.7708
##	1st Qu.:2019-06-06	1st Qu.: 48.02	1st Qu.: -0.7839
##	Median :2019-12-03	Median : 62.88	Median : 0.1964
##	Mean :2019-12-03	Mean : 68.10	Mean : 0.1936
##	3rd Qu.:2020-06-03	3rd Qu.: 79.20	3rd Qu.: 1.3388
##	Max. :2020-11-30	Max. :130.98	Max. : 11.3157

Dans la pratique (modélisation - stationarité, méthode *a la* {tidyverse})

Test KPSS sur les prix - H_0 : données stationnaires

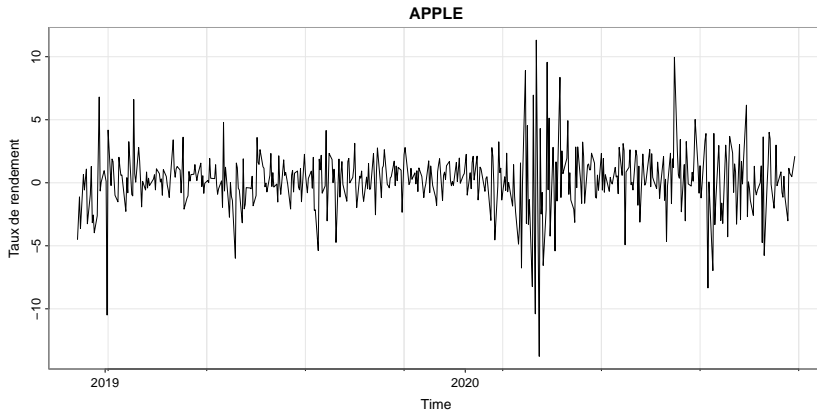
```
## $test_stat
## [1] 7.460073
##
## $critical_values
##           10pct  5pct 2.5pct  1pct
## critical values 0.347 0.463 0.574 0.739
```

Test KPSS sur les rendements - H_0 : données stationnaires

```
## $test_stat
## [1] 0.1184969
##
## $critical_values
##           10pct  5pct 2.5pct  1pct
## critical values 0.347 0.463 0.574 0.739
```

Dans la pratique (modélisation - graphique)

```
tsplot(index(AAPL.r), AAPL.r, main = "APPLE",  
       ylab = "Taux de rendement")
```



9. Vraisemblance autorégressive

Densité conditionnelle :

$$\begin{aligned}f(y_T, y_{T-1}, \dots, y_1 | y_0) &= f(y_T | y_{T-1}, \dots, y_1, y_0) \times f(y_{T-1}, \dots, y_1 | y_0) \\&= f(y_T | y_{T-1}, \dots, y_1, y_0) \times f(y_{T-1} | y_{T-2}, \dots, y_1, y_0) \times f(y_{T-2}, \dots, y_1 | y_0) \\&= \dots\end{aligned}$$

$$= \prod_{t=1}^T f(y_t | y_{t-1}, \dots, y_1, y_0)$$

Cette formule est généralement valable - décomposition de prédiction.

En utilisant la propriété de Markov :

$$f_{\nu, \alpha, \sigma^2}(y_T, \dots, y_1 | y_0) = \prod_{t=1}^T f_{\nu, \alpha, \sigma^2}(y_t | y_{t-1})$$

Vraisemblance autorégressive (suite)

Étant donné l'hypothèse de normalité conditionnelle :

$$\begin{aligned} f_{\nu, \alpha, \sigma^2}(y_T, \dots, y_1 | y_0) &= \prod_{t=1}^T (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y_t - \nu - \alpha y_{t-1})^2\right) \\ &= (2\pi\sigma^2)^{-T/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \nu - \alpha y_{t-1})^2\right) \end{aligned}$$

Et la fonction de vraisemblance est :

$$L(y_1, \dots, y_T | y_0; \nu, \alpha, \sigma^2) = (2\pi\sigma^2)^{-T/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \nu - \alpha y_{t-1})^2\right)$$

La vraisemblance exacte inclut la condition initiale. La vraisemblance conditionnelle est conditionnée par y_0 .

MLE

L'estimation par maximum de vraisemblance (MLE) de (ν, α) est obtenue en minimisant la somme des carrés des résidus :

$$\arg \max_{(\nu, \alpha)} l(\nu, \alpha) = \arg \min_{(\nu, \alpha)} \sum_{t=1}^T \varepsilon_t^2(\nu, \alpha)$$

La MLE de $(\tilde{\nu}, \tilde{\alpha})$ est équivalente à l'estimation par les moindres carrés ordinaires (OLS) de $(\hat{\nu}, \hat{\alpha})$.

$(\tilde{\nu}, \tilde{\alpha})$ sont des estimateurs convergents si y_t est stationnaire et $\sqrt{T}(\tilde{\delta} - \delta)$ est asymptotiquement normalement distribué, où $\delta = (\nu, \alpha)'$.

Les MLE basées sur la vraisemblance exacte et la vraisemblance conditionnelle sont asymptotiquement équivalentes.

L'estimation par la méthode des moments (en utilisant les équations de Yule-Walker) est également équivalente pour α .

Dans la pratique (estimation - {stats})

On suppose une distribution conditionnelle gaussienne des ε_t .

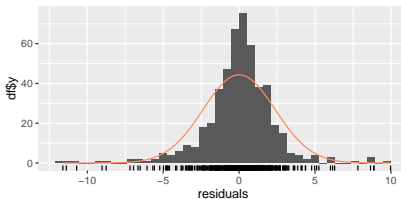
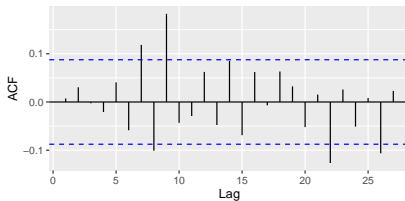
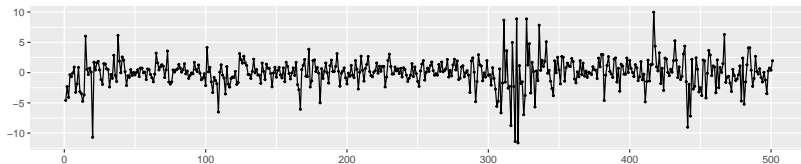
```
AAPL.estim = arima(AAPL.r, order = c(1, 0, 0))  
coeftest(AAPL.estim)
```

```
##  
## z test of coefficients:  
##  
##           Estimate Std. Error z value Pr(>|z|)  
## ar1         -0.211687   0.043821 -4.8307 1.361e-06 ***  
## intercept    0.194581   0.087407  2.2262   0.026 *  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Dans la pratique (estimation - {forecast})

```
fit = Arima(AAPL.r, order = c(1, 0, 0))  
checkresiduals(fit)
```

Residuals from ARIMA(1,0,0) with non-zero mean



```
##  
## Ljung-Box test  
##  
## data: Residuals from ARIMA(1,0,0) with non-zero mean  
## Q* = 33.733, df = 9, p-value = 9.947e-05  
##  
## Model df: 1. Total lags used: 10
```

Dans la pratique (estimation - autoarima{forecast})

```
fit = auto.arima(AAPL.r)
fit
```

```
## Series: AAPL.r
## ARIMA(4,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ar2      ar3      ar4      ma1      mean
##      -1.0701  -0.1448  0.0191  -0.0724  0.8901  0.1948
## s.e.    0.0859   0.0676  0.0662   0.0504  0.0748  0.0870
##
## sigma^2 = 5.517:  log likelihood = -1135.81
## AIC=2285.62   AICc=2285.85   BIC=2315.14
```

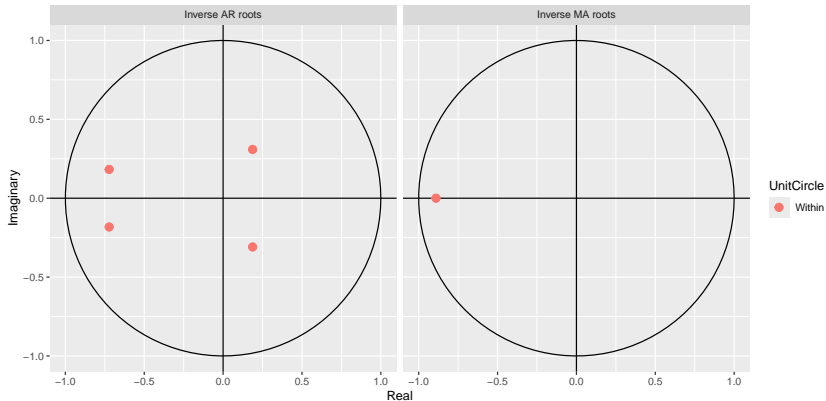
Dans la pratique (estimation - autoarima{forecast}, pvalues)

```
coeftest(fit)
```

```
##
## z test of coefficients:
##
##           Estimate Std. Error  z value Pr(>|z|)
## ar1         -1.070092   0.085930 -12.4531 < 2e-16 ***
## ar2         -0.144838   0.067607  -2.1423  0.03217 *
## ar3          0.019084   0.066192   0.2883  0.77311
## ar4         -0.072442   0.050358  -1.4385  0.15028
## ma1          0.890074   0.074820  11.8962 < 2e-16 ***
## intercept    0.194787   0.086957   2.2400  0.02509 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

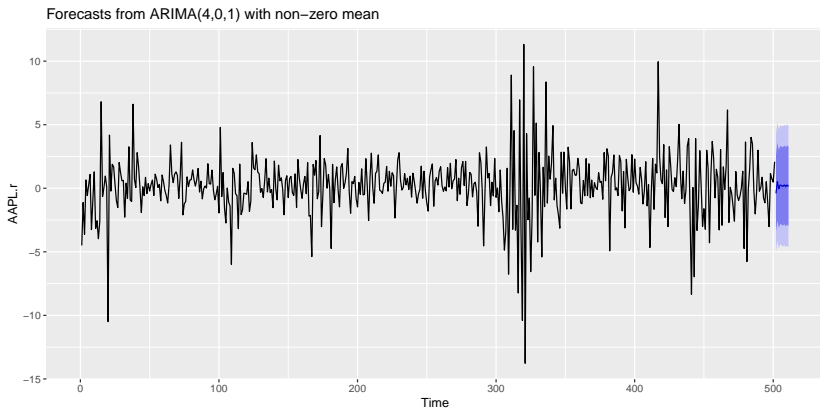
Dans la pratique (estimation - autoarima{forecast}, racines unitaires)

```
autoplot(fit)
```



Dans la pratique (forecast - {forecast})

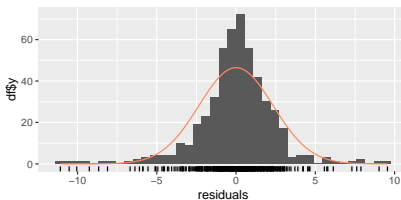
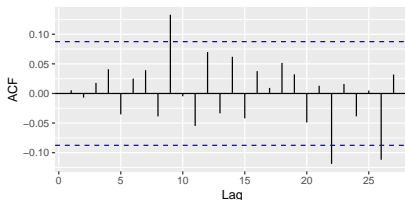
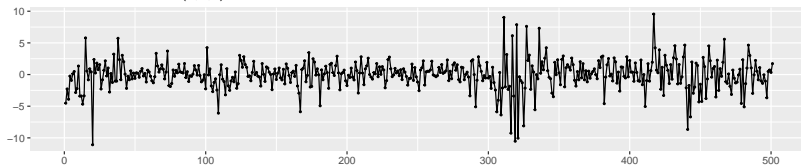
```
autoplot(forecast(fit))
```



Dans la pratique (residuals check - autoarima{forecast})

```
checkresiduals(fit)
```

Residuals from ARIMA(4,0,1) with non-zero mean



```
##  
## Ljung-Box test  
##  
## data: Residuals from ARIMA(4,0,1) with non-zero mean  
## Q* = 12.629, df = 5, p-value = 0.02712  
##  
## Model df: 5. Total lags used: 10
```

AIC et BIC

Pour effectuer une sélection de modèles avec des critères d'information.

AIC est la méthode de sélection de modèles la plus connue et utilisée.

AIC Akaike (1973) :

$$AIC = -2 \times \text{LogLik} + 2 \times p$$

BIC Schwarz (1978) :

$$BIC = -2 \times \text{LogLik} + \log(n) \times p$$

avec n le nombre d'observations dans l'échantillon étudié et p le nombre de paramètres.

Plus petit est AIC/BIC, meilleur est le modèle.

```
AIC(AAPL.estim) #AR(1)
```

```
## [1] 2292.332
```

```
AIC(fit) #ARMA(4,1)
```

```
## [1] 2285.623
```

10. Analyse de la mauvaise spécification

Les hypothèses faites lors de la construction de modèles économétriques.

Exemple – Modèle AR(1) :

$$y_t = \nu + \alpha y_{t-1} + \varepsilon_t$$

Le modèle statistique conditionnel de y_1, y_2, \dots, y_T étant donné y_0 est défini par les hypothèses suivantes :

1. **Indépendance** : $\varepsilon_1, \dots, \varepsilon_T$ sont indépendants ;
2. **Normalité conditionnelle** : $\varepsilon_t \stackrel{D}{=} N(0, \sigma^2)$;
3. **Espace des paramètres** : $\nu, \alpha, \sigma^2 \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$.

Il est important de tester si ces hypothèses sont valides.

Normalité

Tester si l'asymétrie et l'aplatissement correspondent à une distribution normale.

Soit $x_t \sim D(\mu, \sigma^2)$ tel que :

$$y_t = \frac{x_t - \mu}{\sigma} \sim D(0, 1)$$

Définir :

$$\kappa_3 = E(y_t^3) \quad (\text{asymétrie})$$

$$\kappa_4 = E(y_t^4) - 3 \quad (\text{excès d'aplatissement})$$

Le test de **Jarque-Bera** - $H_0 : \kappa_3 = \kappa_4 = 0$ pour la normalité.

$$JB = \frac{T-k}{6} \left(\kappa_3^2 + \frac{(\kappa_4)^2}{4} \right) \sim \chi_2^2 \text{ et } k \text{ le nombre de variables explicatives.}$$

Dans la pratique

```
AAPL.resid = AAPL.estim$residuals  
jarque.bera.test(AAPL.resid)
```

```
##  
## Jarque Bera Test  
##  
## data: AAPL.resid  
## X-squared = 527.32, df = 2, p-value < 2.2e-16
```

Hétéroscédasticité

Hétéroscédasticité - variance non constante

Test de White (1980) : La variance $V(\varepsilon_t|Y_{t-1})$ varie-t-elle avec Y_{t-1} ?

1. Obtenir les résidus $\hat{\varepsilon}_t$ du modèle original.
2. Obtenir R_{het}^2 à partir de la régression auxiliaire :

$$\hat{\varepsilon}_t^2 = \beta_1 + \beta_2 y_{t-1} + \beta_3 y_{t-1}^2 + \eta_t$$

3. Tester $\beta_2 = \beta_3 = 0$ en utilisant :

$$TR_{het}^2 \approx F_{het} = \frac{R_{het}^2/m}{(1 - R_{het}^2)/(T - k - m - 1)}$$

où k est le nombre de régresseurs et m est le nombre de termes quadratiques et de produits croisés.

Notez que le $TR_{het}^2 \sim \chi_2^2$

Dans la pratique (Test de White - hétéroscédasticité)

```
# Étape 1 : Extraire les résidus
residuals <- resid(AAPL.estim)
# Étape 2 : Construire les termes quadratiques et croisés
residuals_squared <- residuals^2
x1 = lag(AAPL.r)
x2 = lag(AAPL.r)^2
auxiliary_data <- data.frame(x1, x2)
# Supprimer les valeurs NA introduites par le décalage
auxiliary_data <- na.omit(auxiliary_data)
# Étape 3 : Régression auxiliaire pour le test de White
auxiliary_model <- lm(residuals_squared ~ x1 + x2, data = auxiliary_data)
# Étape 4 : Test de significativité globale (statistique de White)
white_stat <- summary(auxiliary_model)$r.squared * nrow(auxiliary_data) # Stat
p_value <- 1 - pchisq(white_stat, df = 2) # Degré de liberté = nb de variables
# Résultat : H0 - hétéroscédasticité
cat("Statistique de White :", white_stat, "\n")
```

```
## Statistique de White : 22.61797
```

```
cat("P-valeur :", p_value, "\n")
```

```
## P-valeur : 1.226225e-05
```

Forme fonctionnelle

La forme fonctionnelle log-linéaire est souvent supposée.

Test RESET de Ramsey (1969) : Erreur de spécification de la régression.

Les polynômes de la variable prédite \hat{y}_t aident-ils à expliquer y_t ? Sous l'hypothèse nulle, les variables de la forme fonctionnelle correcte sont irrelevantes.

1. Obtenir les variables prédites \hat{y}_t du modèle original.
2. Obtenir la corrélation partielle R_{reset}^2 de y_t et \hat{y}_t^2 étant donné y_{t-1} à partir de la régression auxiliaire :

$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 \hat{y}_t^2 + \eta_t$$

3. Tester $\beta_3 = 0$ en utilisant :

$$TR_{reset}^2 \approx F_{reset} = \frac{R_{reset}^2 / m}{(1 - R_{reset}^2) / (T - k - m - 1)}$$

où k est le nombre de régresseurs et m est le nombre de restrictions.

Dans la pratique (Test de Ramsey - erreur de spécification)

Devoir #1 - A faire (de préférence sous sa forme TR^2)

Erreurs autocorrélées

L'indépendance des innovations dans un AR(1) implique $H_0 : \alpha_2 = 0$ dans :

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$$

où $\varepsilon_t \sim \text{NID}(0, \sigma^2)$.

Test du rapport de vraisemblance pour $\alpha_2 = 0$:

$$LR = -T \log(1 - r_{02.1}^2)$$

Cela peut être appliqué à la régression auxiliaire – obtenir R_{ar}^2 à partir de :

$$\widehat{\varepsilon}_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 \widehat{\varepsilon}_{t-1} + \eta_t$$

Tester $\beta_2 = \beta_3 = 0$ en utilisant :

$$TR_{ar}^2 \approx F_{ar} = \frac{R_{ar}^2/m}{(1 - R_{ar}^2)/(T - k - m - 1)}$$

où k est le nombre de régresseurs et m est le nombre de restrictions.

Test d'absence d'autocorrélation

Box-Pierce et Ljung-Box

H_0 : absence d'autocorrélation

H_1 : autocorrélation

Idée principale : On estime la fonction d'autocorrélation

$$H_0 : \rho(1) = \rho(2) = \dots = \rho(p) = 0$$

La statistique du Ljung-Box est donnée par :

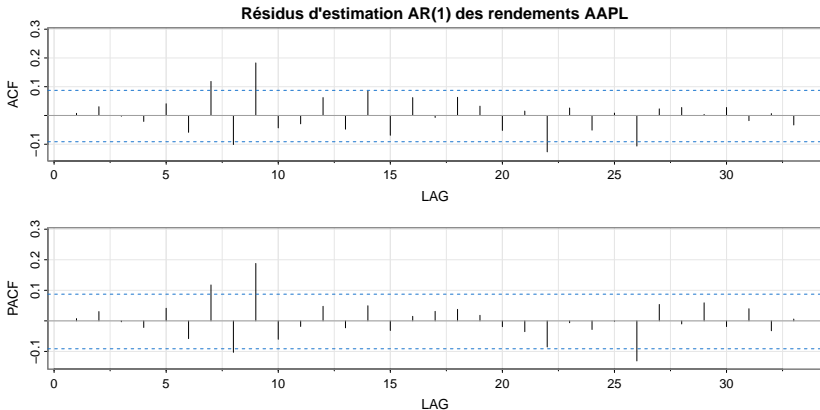
$$Q(m) = T(T+2) \sum_{h=1}^m \frac{\hat{\rho}(h)^2}{T-h}$$

$$\text{avec } \hat{\rho}(h) = \frac{\sum_{t=h+1}^T (\varepsilon_t - \bar{\varepsilon})(\varepsilon_{t-h} - \bar{\varepsilon})}{(\varepsilon_t - \bar{\varepsilon})^2}$$

On rejette H_0 si $Q(m) > q_\alpha$ avec q_α le $100(1 - \alpha)$ quantile d'une χ_{m-p-q}^2

Dans la pratique (ACF et PACF)

```
AAPL_ACF_PACF=acf2(resid(AAPL.estim),  
                    main = "Résidus d'estimation AR(1) des rendements AAPL")
```



Dans la pratique (Test de Ljung-Box - absence d'autocorrélation)

```
Box.test(resid(AAPL.estim), lag = 1, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 0.02769, df = 1, p-value = 0.8678
```

```
Box.test(resid(AAPL.estim), lag = 6, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 3.3109, df = 6, p-value = 0.7689
```

```
Box.test(resid(AAPL.estim), lag = 12, type = "Ljung")
```

```
##  
## Box-Ljung test  
##  
## data: resid(AAPL.estim)  
## X-squared = 36.169, df = 12, p-value = 0.0003043
```

Dans la pratique (Test de Box-Pierce - absence d'autocorrélation)

```
Box.test(resid(AAPL.estim), lag = 1, type = "Box")
```

```
##  
## Box-Pierce test  
##  
## data: resid(AAPL.estim)  
## X-squared = 0.027525, df = 1, p-value = 0.8682
```

```
Box.test(resid(AAPL.estim), lag = 6, type = "Box")
```

```
##  
## Box-Pierce test  
##  
## data: resid(AAPL.estim)  
## X-squared = 3.2648, df = 6, p-value = 0.775
```

```
Box.test(resid(AAPL.estim), lag = 12, type = "Box")
```

```
##  
## Box-Pierce test  
##  
## data: resid(AAPL.estim)  
## X-squared = 35.428, df = 12, p-value = 0.0004002
```

Erreurs ARCH

Engle (1982) : Hétéroscédasticité conditionnelle autorégressive
ARCH - les variances des innovations sont une fonction du temps :

$$V(\varepsilon_t | \varepsilon_{t-1}) = \beta_1 + \beta_2 \varepsilon_{t-1}^2$$

1. Obtenir les résidus $\hat{\varepsilon}_t$ du modèle original.
2. Obtenir R_{arch}^2 à partir de la régression auxiliaire :

$$\hat{\varepsilon}_t^2 = \beta_1 + \beta_2 \hat{\varepsilon}_{t-1}^2 + \eta_t$$

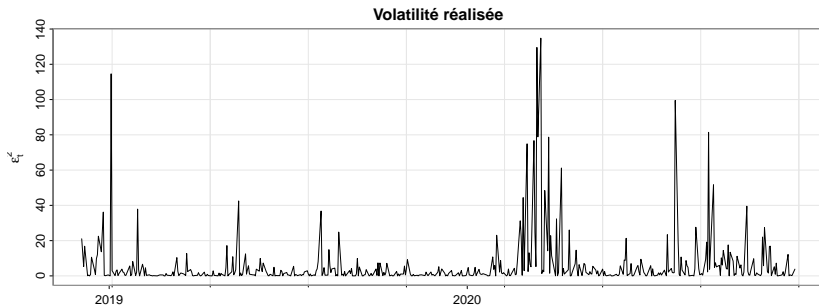
3. Tester $H_0 : \beta_2 = 0$ en utilisant :

$$TR_{arch}^2 \approx F_{arch} = \frac{R_{arch}^2 / m}{(1 - R_{arch}^2) / (T - k - m - 1)}$$

où k est le nombre de régresseurs et m est le nombre de restrictions.

Dans la pratique (Test d'Engle - ARCH)

```
# L'un des proxies de la volatilité et l'élévation de la série au carré :  
tsplot(index(AAPL.r), resid(AAPL.estim)^2, main = "Volatilité réalisée",  
        ylab = expression(epsilon[t]^2), xlab = '')
```



```
# Le test statistique, quant à lui, existe dans le package {FinTS}  
ArchTest(resid(AAPL.estim), lags = 1)
```

```
##  
## ARCH LM-test; Null hypothesis: no ARCH effects  
##  
## data: resid(AAPL.estim)  
## Chi-squared = 20.499, df = 1, p-value = 5.966e-06
```

11. Application

Taux de chômage au USA (1948-2024)

Le modèle est donné par :

$$x_t = \nu + \alpha x_{t-1} + \varepsilon_t$$

Les résidus $\{\varepsilon_t\}$ sont obtenus à partir de $\varepsilon_t = x_t - \nu - \alpha x_{t-1}$.

Les résultats sont les suivants :

$$x_t = 5.5867 + 0.9704x_{t-1} + \varepsilon_t$$

(0.0000) (0.0000)

(pvalues)

Les données

```
USunrate = fredr(series_id = "UNRATE",  
                 observation_start = as.Date("1948-01-01"),  
                 observation_end = as.Date("2024-10-01"),  
                 frequency = "m")  
x = USunrate$value  
xDate = USunrate$date
```

```
# Tester la stationarité  
adf_test = adf.test(x)  
print(adf_test)
```

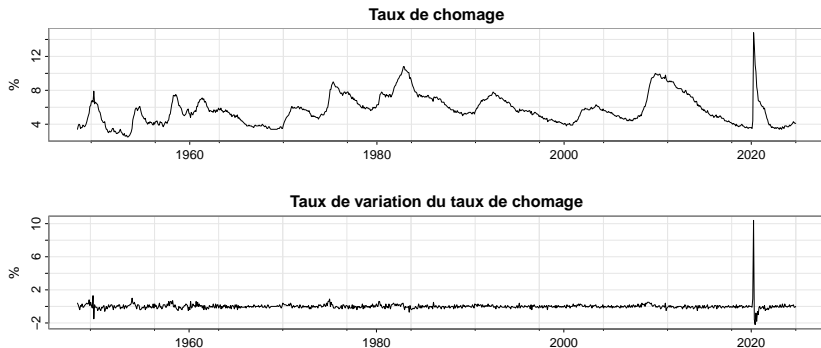
```
##  
## Augmented Dickey-Fuller Test  
##  
## data: x  
## Dickey-Fuller = -3.9542, Lag order = 9, p-value = 0.01129  
## alternative hypothesis: stationary
```

```
#kpss.test(x)
```

```
# Si la série n'est pas stationnaire, appliquer une différenciation  
# x = diff(x)
```

Graph de la série

```
par(mfrow=c(2,1))
tsplot(USunrate$date,USunrate$value,
main="Taux de chômage", ylab = "%", xlab='')
# Si on avait dû différencier la série :
tsplot(xDate[-1],diff(USunrate$value),
      main= "Taux de variation du taux de chômage", ylab = "%", xlab='')
```



Estimation d'un AR(1)

```
fit = arima(x, order=c(1,0,0))
coefs = coef(fit)
print(coefs)
```

```
##          ar1 intercept
## 0.9704322 5.5865874
```

```
# Significativité statistique
std_errors = sqrt(diag(fit$var.coef))
p_values = 2 * (1 - pnorm(abs(coefs / std_errors)))
print(p_values)
```

```
##          ar1 intercept
##          0          0
```

```
# Ou simplement
```

```
#coeftest(fit)
```

```
#Résidus et serie ajustée
```

```
xres = residuals(fit)
```

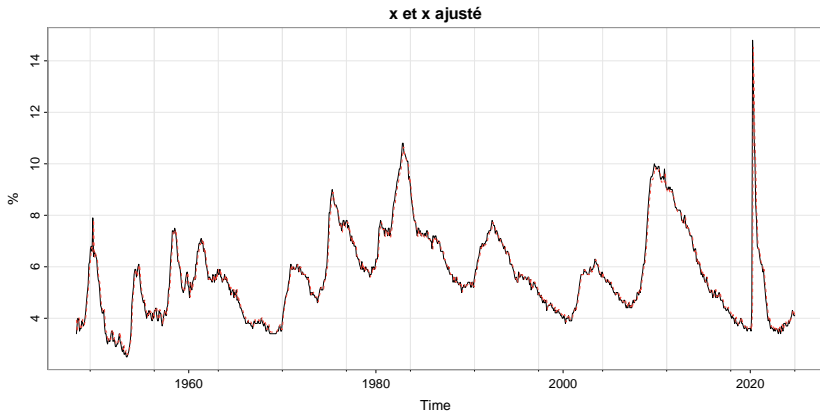
```
xfit = x - xres
```

```
# print(head(xres))
```

```
# print(head(xfit))
```

Graph de y et \hat{y}

```
par(mfrow=c(1,1))  
tsplot(xDate, x, main="x et x ajusté", ylab = "%")  
points(xDate, xfit, type = "l", col = 2, lty = 2)
```



Tests de mauvaises spécification

```
Box.test(xres, lag=2, type="Ljung-Box")#H0: absence d'autocorrelation
```

```
##  
## Box-Ljung test  
##  
## data: xres  
## X-squared = 2.851, df = 2, p-value = 0.2404
```

```
shapiro.test(xres)#H0: normalité
```

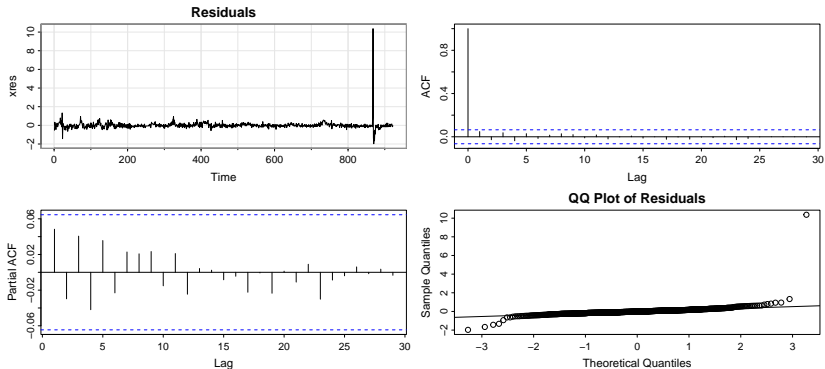
```
##  
## Shapiro-Wilk normality test  
##  
## data: xres  
## W = 0.38724, p-value < 2.2e-16
```

```
bptest(xres ~ xfit)#H0: homoscedasticité
```

```
##  
## studentized Breusch-Pagan test  
##  
## data: xres ~ xfit  
## BP = 0.11118, df = 1, p-value = 0.7388
```

Graph des résidus

```
par(mfrow = c(2, 2))  
tsplot(xres, main = "Residuals")  
acf(xres, main = "ACF of Residuals")  
pacf(xres, main = "PACF of Residuals")  
qqnorm(xres, main = "QQ Plot of Residuals")  
qqline(xres)
```



Conclusion

Le modèle est non congruent - échoue au test de normalité.

Un modèle $AR(p)$ ($p > 1$) pourrait être essayé.

L'explication macroéconomique suggère que des covariables sont nécessaires.

Pas de preuves statistiques de non-stationnarité, mais forte persistance :

$$\alpha = 0.9704^{***}$$

Par conséquent, examiner les données comme non stationnaires ensuite.

Puis considérer les modèles dynamiques multivariés.