

Chapitre 5

Modèles pour les séries chronologiques non stationnaires

Ghislain Léveillé
École d'actuariat, Université Laval

5.1 Contenu du chapitre

Nous avons déjà vu que n'importe quel modèle de séries chronologiques avec moyenne non constante ne pouvait être stationnaire. Qui plus est, dans plusieurs applications, l'hypothèse d'une tendance déterministe ne peut tenir indéfiniment. Cependant, plusieurs processus non stationnaires peuvent être « reconvertis » en processus stationnaires, de sorte à pouvoir utiliser les propriétés qui leurs sont associées.

Ainsi, dans ce chapitre, nous introduisons d'abord l'opérateur de différenciation qui nous permettra d'induire la stationnarité sur certains processus non stationnaires. Cela nous amènera par la suite à définir les processus autorégressifs à moyenne mobile intégrés (ARIMA), dont nous examinerons les propriétés. Enfin, d'autres transformations vers la stationnarité seront considérées, telles les changements de pourcentage, le logarithme et les puissances (ou transformées Box-Cox).

5.2 Stationnarité par opération de différenciation

Afin d'introduire le sujet, considérer le processus autorégressif AR(1) suivant :

$$Y_t = 3Y_{t-1} + e_t \quad ,$$

où $\{e_t; t \in \mathbb{Z}\}$ est un bruit blanc, $E[e_t] = 0$ et e_t indépendant de Y_{t-1} .

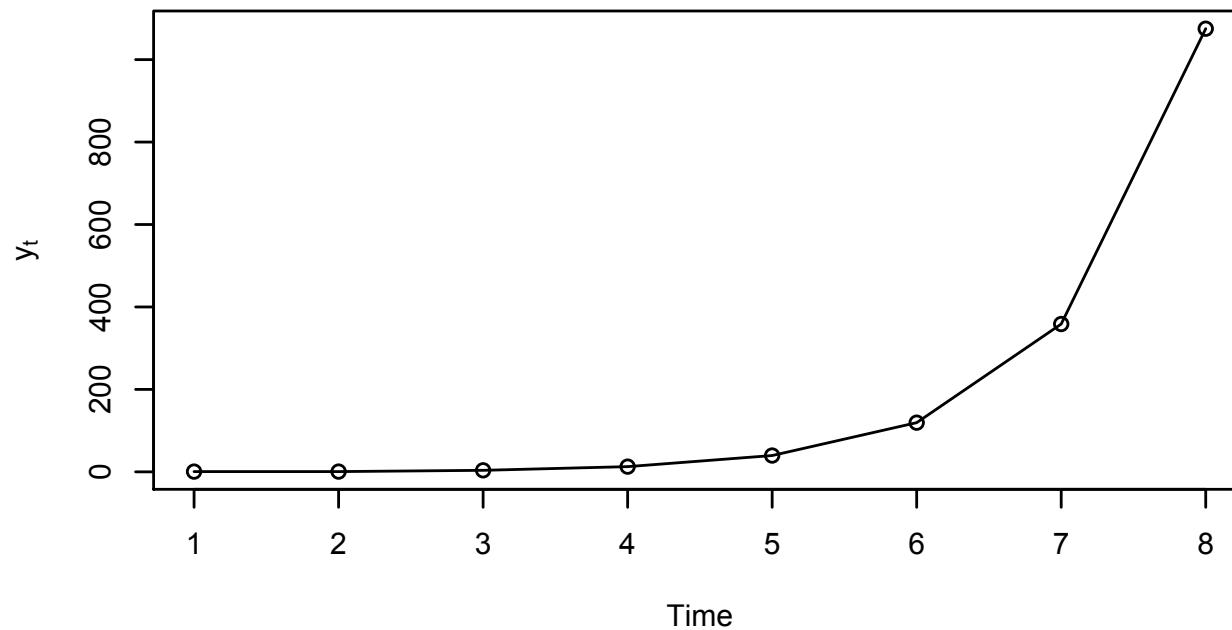
Nous savons que ce processus n'est pas stationnaire puisque le paramètre ϕ ($\phi = 3$) est supérieur à 1. Ceci est confirmé par le fait que ce processus « explose » avec le temps, puisque

$$Y_t = e_t + 3e_{t-1} + 3^2 e_{t-2} + \dots + 3^{t-1} e_1 + 3^t Y_0 \quad ,$$

impliquant une croissance exponentielle des facteurs associés à Y_0 et à e_1 .

Illustrons le processus précédent.

```
> data(explode.s)  
> win.graph(width=4.875, height=3, pointsize=8)  
> plot(explode.s, ylab=expression(y[t]), type='o')
```



Ce comportement « explosif » est aussi confirmé par les fonctions de variance et d'autocovariance. En supposant $Y_0 = 0$, nous avons

$$- \text{Var}[Y_t] = E(Y_t^2) = (1 + 9 + 9^2 + \dots + 9^{t-1})\sigma_e^2 = \frac{9^t - 1}{8}\sigma_e^2$$

$$- \text{Cov}(Y_t, Y_{t-k}) = E[Y_t Y_{t-k}] = E\left[3^k (1 + 9 + \dots + 9^{t-k-1})\sigma_e^2\right] = \frac{3^k (9^{t-k} - 1)}{8}\sigma_e^2$$

Nous notons que, pour t « grand » et k « relativement petit » ($k = 1$ ou 2 par exemple), nous avons

$$- \text{Corr}(Y_t, Y_{t-k}) = 3^k \frac{9^{t-k} - 1}{9^t - 1} \approx 3^{-k}$$

ce qui indique qu'une certaine corrélation (positive) persiste avec le temps.

Considérons maintenant le processus autorégressif AR(1) suivant :

$$Y_t = Y_{t-1} + e_t \quad ,$$

avec les mêmes conditions que précédemment.

Ce processus est non stationnaire puisque $\phi = 1$. De plus, nous avons

- $Y_t = e_t + e_{t-1} + \dots + e_1 + Y_0$
- En supposant $Y_0 = 0$, nous avons

$$Var[Y_t] = t\sigma_e^2 \quad , \quad Cov(Y_t, Y_{t-k}) = (t-k)\sigma_e^2 \quad \Rightarrow \quad Corr(Y_t, Y_{t-k}) = \frac{t-k}{t}$$

Notons toutefois qu'en appliquant l'opérateur de différenciation ∇ (première différence avec écart 1) défini par

$$\nabla Y_t = Y_t - Y_{t-1}$$

sur le processus précédent, nous obtenons

$$\nabla Y_t = e_t ,$$

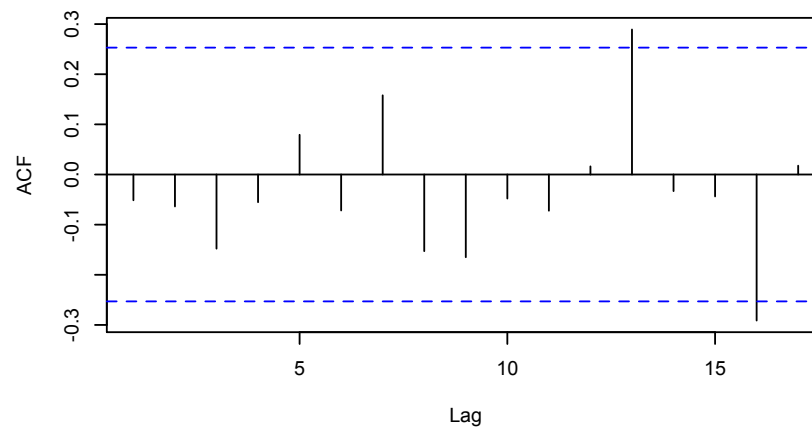
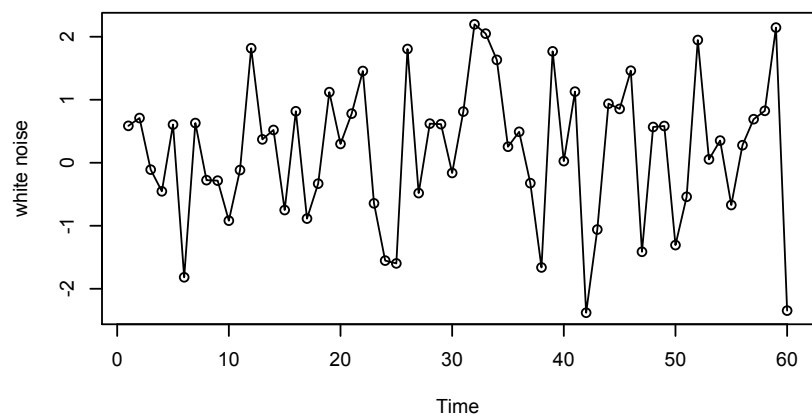
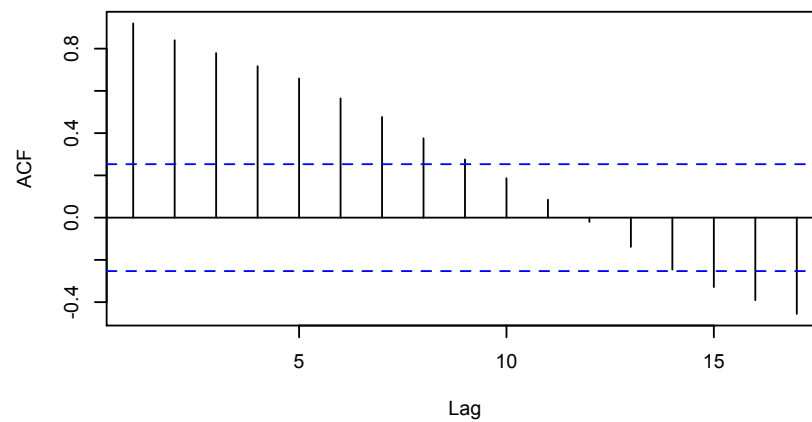
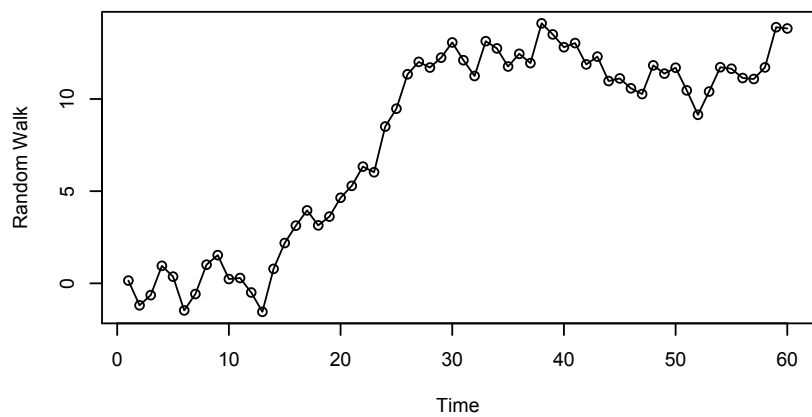
qui lui est un processus stationnaire (e_t générant un bruit blanc).

Remarque : Si nous appliquons cet opérateur au processus AR(1) défini par $Y_t = 3Y_{t-1} + e_t$, nous n'arriverons pas à « provoquer » la stationnarité puisque

$$\nabla Y_t = 2Y_{t-1} + e_t .$$

Illustrons les opérations précédentes pour notre processus non stationnaire $Y_t = Y_{t-1} + e_t$, où $e_t \sim N(0,1)$.

```
> n=60
> set.seed(12345)
> white.noise.norm=ts(rnorm(n),freq=1,start=1)
> sim.random.walk=ts(cumsum(rnorm(n)),freq=1,start=1)
> model7=lm(white.noise.norm~time(white.noise.norm))
> model8=lm(sim.random.walk~time(sim.random.walk))
> win.graph(width=4.875,height=3,pointsize=8)
> plot(white.noise.norm,type='o',ylab='white noise')
> plot(sim.random.walk,type='o',ylab='Random Walk')
> acf(rstudent(model7),main='')
> acf(rstudent(model8),main='')
```

La méthode de différenciation appliquée à la marche aléatoire précédente peut être étendue à d'autres séries chronologiques dont la première différence sera un processus stationnaire (et pas nécessairement un bruit blanc).

- Comme premier exemple, considérer le processus (non stationnaire)

$$Y_t = \beta_0 + \beta_1 t + X_t ,$$

où X_t génère un processus (faiblement) stationnaire.

Sous cette hypothèse, la première différence donnera

$$\nabla Y_t = \beta_1 + \nabla X_t ,$$

qui est un processus (faiblement) stationnaire puisque ...

$$\begin{aligned} - E[\nabla Y_t] &= \beta_1 + E[X_t] - E[X_{t-1}] = \beta_1 \\ - Cov(\nabla Y_t, \nabla Y_{t-k}) &= Cov(X_t, X_{t-k}) - Cov(X_t, X_{t-k-1}) \\ &\quad - Cov(X_{t-1}, X_{t-k}) + Cov(X_{t-1}, X_{t-k-1}) \end{aligned}$$

... la dernière identité ne dépendant que de l'écart k , par hypothèse de (faible) stationnarité du processus généré par X_t .

- Comme deuxième exemple, considérer le processus

$$Y_t = M_t + e_t ,$$

où $M_t = M_{t-1} + \varepsilon_t$, e_t et ε_t générant des bruits blanc indépendants, et $E[e_t] = E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = \sigma_\varepsilon^2$, $Var[e_t] = \sigma_e^2$.

Sous cette hypothèse, nous obtenons

$$\nabla Y_t = \nabla M_t + \nabla e_t = \varepsilon_t + e_t - e_{t-1} \quad ,$$

qui est un processus (faiblement) stationnaire puisque

$$- E[\nabla Y_t] = 0$$

$$- Var[\nabla Y_t] = \sigma_\varepsilon^2 + 2\sigma_e^2$$

$$- \rho_1 = \frac{Cov(\nabla Y_t, \nabla Y_{t-1})}{Var(\nabla Y_t)} = \frac{-\sigma_e^2}{\sigma_\varepsilon^2 + 2\sigma_e^2} \quad , \quad \rho_k = 0 \quad , \quad k = 2, 3, \dots$$

Qu'advient-il si la première différence ne nous donne pas un processus (faiblement) stationnaire? Il nous faudra alors considérer une deuxième différence ...

- Comme premier exemple, considérer le processus

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + X_t ,$$

où X_t génère un processus (faiblement) stationnaire.

Sous cette hypothèse, la première différence donnera

$$\nabla Y_t = \beta_1 - \beta_2 + 2\beta_2 t + \nabla X_t ,$$

ce qui ne correspond à pas un processus (faiblement) stationnaire, ayant

$$E[\nabla Y_t] = \beta_1 - \beta_2 + 2\beta_2 t .$$

Nous pouvons cependant observer qu'une deuxième différence amènera la stationnarité (faible) voulue. Ainsi ...

$$\begin{aligned}\nabla^2 Y_t &= \nabla Y_t - \nabla Y_{t-1} \\ &= 2\beta_2 + \nabla^2 X_t \\ &= 2\beta_2 + X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

... nous donne maintenant un processus (faiblement) stationnaire puisque

$$- E[\nabla^2 Y_t] = 2\beta_2$$

$$\begin{aligned}- Cov(\nabla^2 Y_t, \nabla^2 Y_{t-k}) &= Cov(X_t, X_{t-k}) - 2Cov(X_t, X_{t-k-1}) + Cov(X_t, X_{t-k-2}) \\ &\quad - 2Cov(X_{t-1}, X_{t-k}) + 4Cov(X_{t-1}, X_{t-k-1}) - 2Cov(X_{t-1}, X_{t-k-2}) \\ &\quad + Cov(X_{t-2}, X_{t-k}) - 2Cov(X_{t-2}, X_{t-k-1}) + Cov(X_{t-2}, X_{t-k-2})\end{aligned}$$

dernière expression qui ne dépend que de k.

- Comme deuxième exemple, considérer le processus

$$Y_t = M_t + e_t \quad ,$$

où $M_t = M_{t-1} + W_t$, $W_t = W_{t-1} + \varepsilon_t$, e_t et ε_t générant des bruits blanc indépendants, et $E[e_t] = E[\varepsilon_t] = 0$.

Sous cette hypothèse, nous obtenons

$$\nabla Y_t = \nabla M_t + \nabla e_t \Rightarrow \nabla^2 Y_t = \nabla W_t + \nabla^2 e_t = \varepsilon_t + e_t - 2e_{t-1} + e_{t-2} \quad ,$$

un processus (faiblement) stationnaire dont la fonction d'autocorrélation est

$$\rho_1 = \frac{Cov(\nabla^2 Y_t, \nabla^2 Y_{t-1})}{Var[\nabla^2 Y_t]} = \frac{-4\sigma_e^2}{\sigma_\varepsilon^2 + 6\sigma_e^2} \quad , \quad \rho_2 = \frac{Cov(\nabla^2 Y_t, \nabla^2 Y_{t-2})}{Var[\nabla^2 Y_t]} = \frac{\sigma_e^2}{\sigma_\varepsilon^2 + 6\sigma_e^2} \quad ,$$

$$\rho_k = 0 \quad , \quad k = 3, 4, \dots$$

Remarque : Nous avons introduit précédemment l'opérateur de différenciation d'ordre 2. Dans la littérature sur les séries chronologiques, cet opérateur est lié à un autre opérateur, appelé l'opérateur de « retard » (backshift operator), noté par B et défini par l'identité suivante :

$$BY_t = Y_{t-1} .$$

Nous avons ainsi

$$\nabla Y_t = Y_t - Y_{t-1} = (1 - B)Y_t \Rightarrow \nabla^k Y_t = (1 - B)^k Y_t , k = 1, 2, \dots$$

Pour le processus $AR(p)$, réécrit avec l'opérateur de retard,

$$(1 - \phi_1 B - \dots - \phi_p B^p)Y_t = e_t ,$$

nous retrouvons le polynôme caractéristique du processus $AR(p)$, que nous avons noté $\phi(B)$.

Pour le processus MA(q), réécrit avec l'opérateur de retard,

$$Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) e_t ,$$

nous obtenons le polynôme caractéristique du processus MA(q), que nous noterons $\theta(B)$.

Pour le processus ARMA(p,q), réécrit avec l'opérateur de retard,

$$\phi(B)Y_t = \theta(B)e_t ,$$

nous obtenons une identité fort intéressante, en termes de polynômes caractéristiques des processus AR(p) et MA(q), que nous utiliserons plus loin.

5.3. Processus ARIMA

5.3.1 Définition

Un processus $\{Y_t; t \in \mathbb{Z}\}$ est un processus autorégressif à moyenne mobile intégré (ARIMA) si la d -ième différence $W_t = \nabla^d Y_t$ est un processus (faiblement) stationnaire ARMA. Si le processus $\{W_t; t \in \mathbb{Z}\}$ est un processus (faiblement) stationnaire ARMA(p, q), nous dirons que $\{Y_t; t \in \mathbb{Z}\}$ est un processus ARIMA(p, d, q).

Si le processus $\{W_t; t \in \mathbb{Z}\}$ ne contient aucune partie autorégressive, nous dirons qu'il s'agit d'un processus à moyenne mobile intégré, noté IMA(d, q). Si ce même processus ne contient aucune partie à moyenne mobile, nous dirons qu'il s'agit d'un processus autorégressif intégré, noté ARI(p, d).

Illustrons cette définition en supposant que $\{Y_t; t \in \mathbb{Z}\}$ est un processus ARIMA(p,1,q). Ayant $W_t = Y_t - Y_{t-1}$, nous avons alors

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad .$$

Ainsi,

$$Y_t - Y_{t-1} = \phi_1 (Y_{t-1} - Y_{t-2}) + \dots + \phi_p (Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad ,$$

que nous pouvons réécrire sous la forme suivante

$$Y_t = (1 + \phi_1) Y_{t-1} + (\phi_2 - \phi_1) Y_{t-2} + \dots + (\phi_p - \phi_{p-1}) Y_{t-p} - \phi_p Y_{t-p-1} \\ + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad ,$$

ce qui correspond à une ARMA(p+1,q) non stationnaire.

Remarques : (1) Le processus généré par Y_t n'est pas stationnaire puisque son polynôme caractéristique

$$1 - (1 + \phi_1)x - (\phi_2 - \phi_1)x^2 - \dots - (\phi_p - \phi_{p-1})x^p + \phi_p x^{p+1} = (1-x)(1 - \phi_1 x - \dots - \phi_p x^p)$$

a une racine égale à 1 ...

(2) Pour le processus ARIMA(p,d,q), réécrit avec l'opérateur de retard,

$$\phi(B)(1-B)^d Y_t = \theta(B)e_t ,$$

nous obtenons une identité qui nous indique que $\{Y_t; t \in \mathbb{Z}\}$ sera une ARMA(p+d,q) non stationnaire, où $x = 1$ est une racine de multiplicité d du polynôme $\phi(x)(1-x)^d$.

5.3.2 Le modèle IMA(1,1)

Considérons le processus IMA(1,1)

$$Y_t = Y_{t-1} + e_t - \theta e_{t-1} .$$

Alors nous pouvons réécrire ce processus comme suit

$$(1 - B)Y_t = (1 - \theta B)e_t \Leftrightarrow W_t = (1 - \theta B)e_t .$$

Le processus généré par Y_t n'est pas stationnaire puisque la valeur absolue de la racine du polynôme $\phi(x) = 1 - x$ est égale 1. Quant au processus W_t , une MA(1), il sera inversible ssi la valeur absolue de la racine du polynôme $\theta(x) = 1 - \theta x$ est supérieure à 1, i.e. $|\theta| < 1$.

Remarque :

Pour avoir une idée comment se comportent $Var[Y_t]$ et $Corr(Y_t, Y_{t-k})$, nous aimerions pouvoir réécrire ce processus comme une somme infinie pondérée (convergente) des e_t , tel que nous l'avons fait au chapitre 4. Cela n'est malheureusement pas possible puisque le processus généré par Y_t n'est pas stationnaire.

Nous pouvons cependant supposer que ce processus a été observé à partir d'un temps $t = -m < 0$, et que $Y_t = 0$ pour $t < -m$. Ainsi en utilisant le fait que $W_t = Y_t - Y_{t-1}$, nous obtenons

$$Y_t = \sum_{j=-m}^t (Y_j - Y_{j-1}) = \sum_{j=-m}^t W_j \quad ,$$

et notre processus IMA(1,1) peut alors se réécrire en terme des valeurs du bruit blanc

$$\begin{aligned}
 Y_t &= \sum_{j=-m}^{t-1} W_j + e_t - \theta e_{t-1} \\
 &= \sum_{j=-m}^{t-1} (e_j - \theta e_{j-1}) + e_t - \theta e_{t-1} \\
 &= e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \dots + (1-\theta)e_{-m} - \theta e_{-m-1} .
 \end{aligned}$$

Ainsi, même si la constante m est inconnue, nous avons

$$- E[Y_t] = 0$$

$$- Var[Y_t] = E[Y_t^2] = \left[1 + \theta^2 + (1-\theta)^2 (t+m) \right] \sigma_e^2$$

$$\begin{aligned}
- \text{Corr}(Y_t, Y_{t-k}) &= \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{\text{Var}[Y_t] \text{Var}[Y_{t-k}]}} \\
&= \frac{1 - \theta + \theta^2 + (1 - \theta)^2 (t + m - k)}{\sqrt{\left[1 - \theta + \theta^2 + (1 - \theta)^2 (t + m)\right] \left[1 - \theta + \theta^2 + (1 - \theta)^2 (t + m - k)\right]}} \\
&= \sqrt{\frac{1 - \theta + \theta^2 + (1 - \theta)^2 (t + m - k)}{1 - \theta + \theta^2 + (1 - \theta)^2 (t + m)}}
\end{aligned}$$

La variance est donc une fonction linéaire du temps, et la corrélation est presque 1 pour t (ou m) « grand » et k « petit ».

5.3.3 Le modèle IMA(2,2)

Considérons le processus IMA(2,2)

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} .$$

Alors nous pouvons réécrire ce processus comme suit

$$(1-B)^2 Y_t = (1 - \theta_1 B - \theta_2 B^2) e_t \Leftrightarrow W_t = (1 - \theta_1 B - \theta_2 B^2) e_t .$$

Le processus généré par Y_t n'est pas stationnaire puisque la valeur absolue de la racine double du polynôme $\phi(x) = (1-x)^2$ est égale 1. Quant au processus W_t , une MA(2), il sera inversible ssi les valeurs (normes) absolues des racines du polynôme $\theta(x) = 1 - \theta_1 x - \theta_2 x^2$ sont supérieures à 1.

Remarque :

Pour avoir une idée comment se comportent $Var[Y_t]$ et $Corr(Y_t, Y_{t-k})$, nous aimerions pouvoir réécrire ce processus comme une somme infinie pondérée (convergente). Cela n'est toujours pas possible puisque $\{Y_t; t \in \mathbb{Z}\}$ n'est pas un processus stationnaire.

Nous supposons donc que ce processus a été observé à partir d'un temps $t = -m < 0$, et que $Y_t = 0$ pour $t < -m$.

En utilisant le fait que $W_t = \nabla Y_t - \nabla Y_{t-1}$, nous obtenons ainsi

$$\nabla Y_j = Y_j - Y_{j-1} = \sum_{i=-m}^j (\nabla Y_i - \nabla Y_{i-1}) = \sum_{i=-m}^j W_i ,$$

et donc

$$Y_t = \sum_{j=-m}^t \nabla Y_j = \sum_{j=-m}^t \sum_{i=-m}^j W_i .$$

Ainsi notre processus IMA(2,2) peut alors se réécrire en terme de valeurs du bruit blanc (à vérifier...)

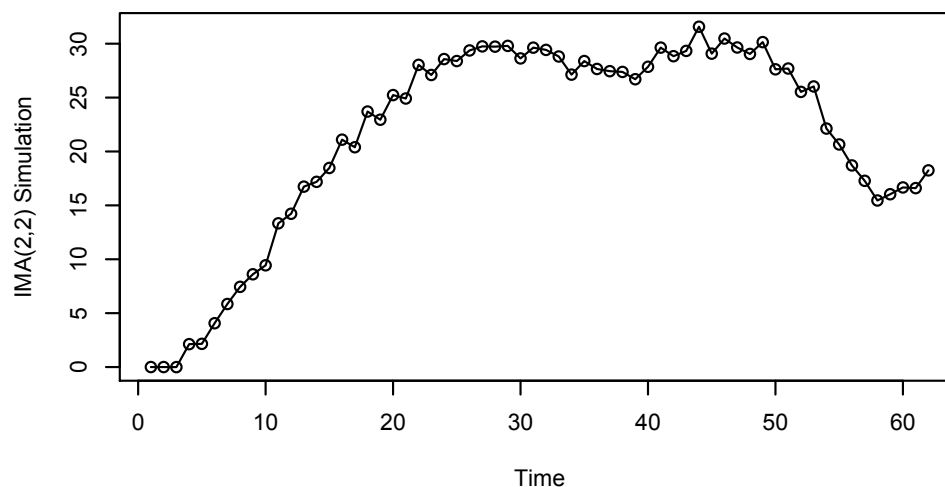
$$Y_t = e_t + \sum_{j=1}^{t+m} \left[1 + \theta_2 + (1 - \theta_1 - \theta_2)j \right] e_{t-j} \\ - \left[(t+m+1)\theta_1 + (t+m)\theta_2 \right] e_{-m-1} - (t+m+1)\theta_2 e_{-m-2} .$$

Les calculs pour la variance et la corrélation sont un peu plus lourds qu'en 5.3.2, et restent donc à faire. Les conclusions sont cependant les mêmes, i.e. que la variance est une fonction linéaire du temps et la corrélation est presque 1 pour t (ou m) « grand » et k « petit ».

Examinons graphiquement le comportement de cette IMA(2,2) pour $\theta_1 = 1$ et $\theta_2 = -0.6$, ainsi que ceux des deux premières différences de ce processus.

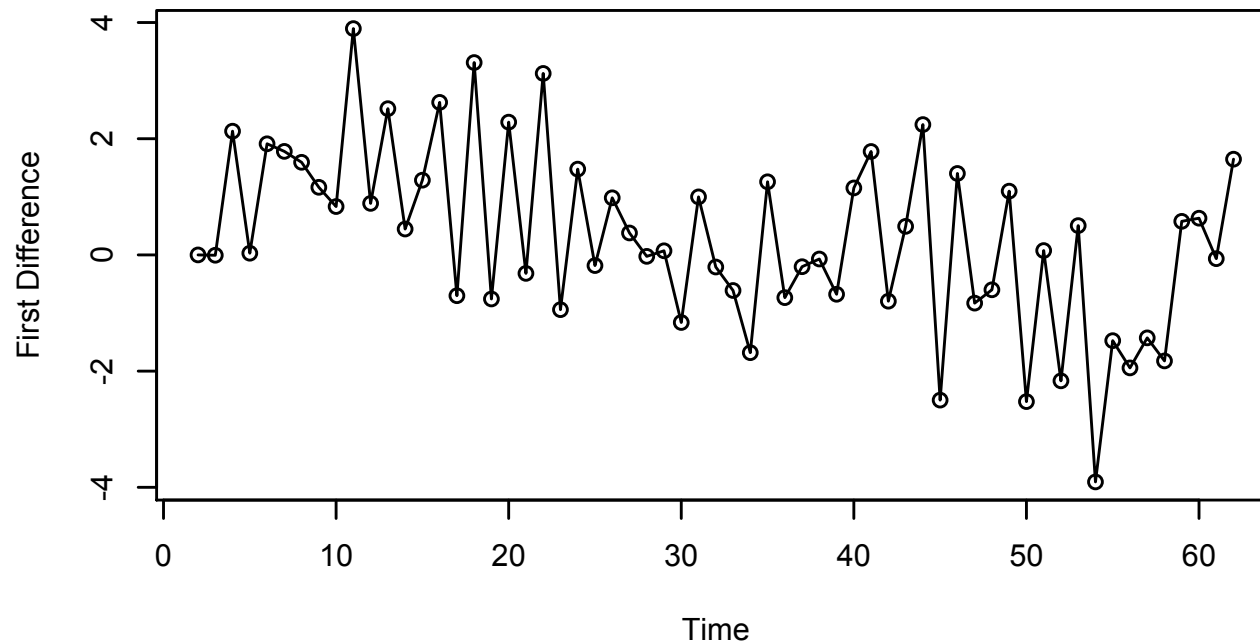
D'abord, une simulation de cette IMA(2,2) ...

```
> data(ima22.s)
> win.graph(width=4.875, height=3, pointsize=8)
> plot(ima22.s, ylab="IMA(2,2) Simulation", type='o')
```



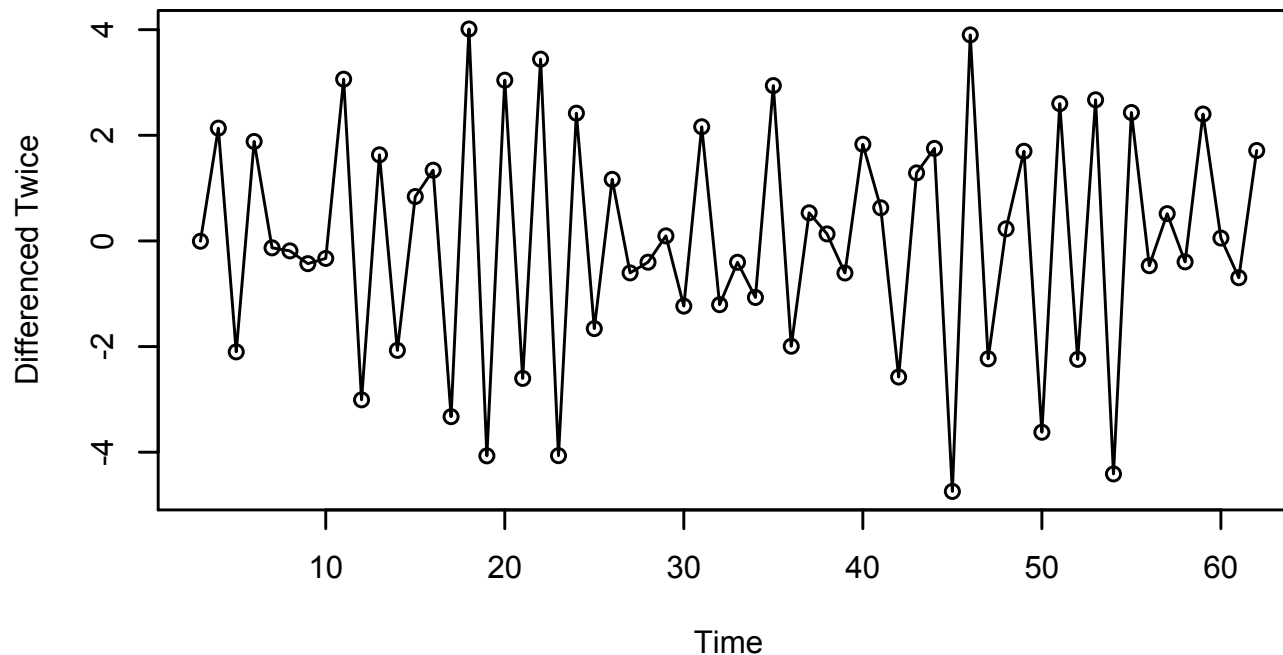
Puis, un graphique de la première différence (... une IMA(1,2))

```
> plot(diff(ima22.s), ylab='First Difference', type='o')
```



Enfin, un graphique de la deuxième différence (... une MA(2) faiblement stationnaire, voir la section 4.3.2)

```
> plot(diff(ima22.s, difference=2), ylab='Differenced Twice',  
      type='o')
```



5.3.4 Le modèle ARI(1,1)

Considérons le processus ARI(1,1)

$$Y_t = (1 + \phi)Y_{t-1} - \phi Y_{t-2} + e_t, \quad |\phi| < 1$$

Alors nous pouvons réécrire ce processus comme suit

$$\begin{aligned} [1 - (1 + \phi)B + \phi B^2]Y_t = e_t &\Leftrightarrow (1 - B)Y_t = \phi(1 - B)Y_{t-1} + e_t \\ &\Leftrightarrow W_t = \phi W_{t-1} + e_t \end{aligned}$$

Le processus généré par Y_t n'est pas stationnaire puisque $\phi_1 + \phi_2 = (1 + \phi) + (-\phi) = 1$. Quant au processus W_t , un AR(1), il est inversible puisque $|\phi| < 1$.

Remarque :

Pour les mêmes raisons qu'en 5.3.2 et 5.3.3, nous aimerions trouver les coefficients (poids ψ) du développement de Y_t en fonction des valeurs du bruit blanc.

Pour ce faire, nous développons une méthode qui se généralisera à n'importe quel modèle ARIMA(p,d,q). Nous nous inspirons en fait de la remarque (2) en section 5.3.1.

Les poids ψ peuvent ainsi être obtenus de l'identité

$$\left(1 - \phi_1 x - \dots - \phi_p x^p\right) (1 - x)^d \left(1 + \psi_1 x + \psi_2 x^2 + \dots\right) = 1 - \theta_1 x - \dots - \theta_q x^q$$

en égalant les coefficients des puissances de x .

Pour notre processus $ARI(1,1)$, nous avons

$$(1 - \phi x)(1 - x)(1 + \psi_1 x + \psi_2 x^2 + \dots) = 1 \quad ,$$

ce qui nous donne les coefficients ψ_k (« autant que nécessaire »)

$$\psi_1 = 1 + \phi \quad , \quad \psi_k = (1 + \phi)\psi_{k-1} - \phi\psi_{k-2} \quad , \quad k = 2, 3, 4, \dots$$

$$\Rightarrow \psi_k = \frac{1 - \phi^{k+1}}{1 - \phi} \quad , \quad k = 1, 2, \dots$$

En effet, cela suit directement l'identité suivante

$$(1 + \psi_1 x + \psi_2 x^2 + \dots) = \frac{1}{(1 - \phi x)(1 - x)} = \sum_{i=0}^{\infty} (\phi x)^i \sum_{j=0}^{\infty} x^j = \sum_{i=0}^{\infty} \phi^i \sum_{k=i}^{\infty} x^k \quad .$$

5.4 Termes constants dans les modèles ARIMA

Jusqu'à maintenant, nous avons supposé que le modèle stationnaire ARMA(p,q), en particulier le processus stationnaire $\{W_t = \nabla^d Y_t, t \in \mathbb{Z}\}$, avait une moyenne nulle. Cette hypothèse n'est pas nécessaire puisque si $E[W_t] = \mu$, nous pouvons toujours écrire

$$W_t - \mu = \phi_1(W_{t-1} - \mu) + \dots + \phi_p(W_{t-p} - \mu) + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

ou encore ce modèle équivalent

$$W_t = \theta_0 + \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} ,$$

où

$$\mu = \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p} \Leftrightarrow \theta_0 = \mu(1 - \phi_1 - \dots - \phi_p) .$$

Maintenant, est-ce que d'ajouter une constante non nulle au processus définissant W_t a une incidence sur le processus ARIMA(p,d,q) si nous voulons réécrire Y_t en fonction des e_t ? La réponse est qu'il peut effectivement y avoir une incidence sur ce développement.

Reprenons l'exemple du processus IMA(1,1) en section 5.3.2, en y ajoutant une constante non nulle θ_0 . Nous avons alors

$$Y_t = \theta_0 + Y_{t-1} + e_t - \theta e_{t-1} \quad .$$

Par récursion, en supposant $Y_t = 0$ pour $t < -m \leq 0$, nous obtenons

$$Y_t = (t - m + 1)\theta_0 + e_t + (1 - \theta)e_{t-1} + (1 - \theta)e_{t-2} + \dots + (1 - \theta)e_{-m} - \theta e_{-m-1}.$$

Nous avons donc ajouté une tendance déterministe linéaire de pente θ_0 .

Une représentation équivalente du processus $\{Y_t; t \in \mathbb{Z}\}$ serait donc

$$Y_t = Y'_t + \beta_0 + \beta_1 t \quad , \quad \{Y'_t; t \in \mathbb{Z}\} \text{ est une IMA}(1,1) \quad ,$$

avec $E[\nabla Y'_t] = 0 \quad (\Rightarrow \quad E[\nabla Y_t] = \beta_1) \quad .$

Remarque : (À vérifier ...)

Plus généralement, pour une ARIMA(p,d,q) tel $E[\nabla^d Y_t] \neq 0$, une représentation équivalente peut aussi être construite, où

$$Y_t = Y'_t + \mu_t \quad , \quad \{Y'_t; t \in \mathbb{N}\} \text{ est une ARIMA}(p,d,q) \quad ,$$

avec $E[\nabla^d Y'_t] = 0 \quad (\Rightarrow \quad E[\nabla^d Y_t] = d! \beta_d \quad)$, et μ_t est un polynôme de degré d en t.

5.5 Autres transformations

5.5.1 Pourcentage de changement et logarithme

Pour certaines séries chronologiques non stationnaires, par exemple celles dont l'écart-type augmente proportionnellement par rapport à la valeur moyenne de la série, il pourrait s'avérer intéressant d'utiliser d'abord une transformation logarithmique.

Ainsi supposons que $Y_t > 0$ pour tout t et que

$$E[Y_t] = \mu_t \quad , \quad \sqrt{Var(Y_t)} = \mu_t \sigma \quad .$$

En approximant le logarithme de Y_t/μ_t par le premier terme de son développement de Taylor (i.e. qu'on suppose que les autres termes de ce développement sont négligeables) ...

... nous aurons

$$\ln\left[\frac{Y_t}{\mu_t}\right] = \ln\left[1 + \frac{Y_t - \mu_t}{\mu_t}\right] \approx \frac{Y_t - \mu_t}{\mu_t} ,$$

ce qui entraîne que

$$\ln(Y_t) \approx \ln \mu_t + \frac{Y_t - \mu_t}{\mu_t} .$$

Nous obtenons une nouvelle série (générée par $\ln(Y_t)$) tel que

$$E\left[\ln(Y_t)\right] \approx \ln(\mu_t) , \quad Var(\ln Y_t) \approx \sigma^2 ,$$

dont nous avons, à tout le moins, « stabilisé » la variance. Il restera ensuite à examiner s'il y a stationnarité de ce nouveau processus, sinon à appliquer les techniques de différenciation.

Considérons une autre situation où Y_t tend à changer d'une manière relativement stable (quasi stationnaire) d'une période à l'autre, i.e.

$$Y_t = (1 + X_t) Y_{t-1} \quad ,$$

où $100X_t$ est le pourcentage de changement de Y_{t-1} à Y_t , et X_t génère un processus (faiblement) stationnaire.

Alors, nous aurons

$$\ln(Y_t) - \ln(Y_{t-1}) = \ln(1 + X_t) \quad .$$

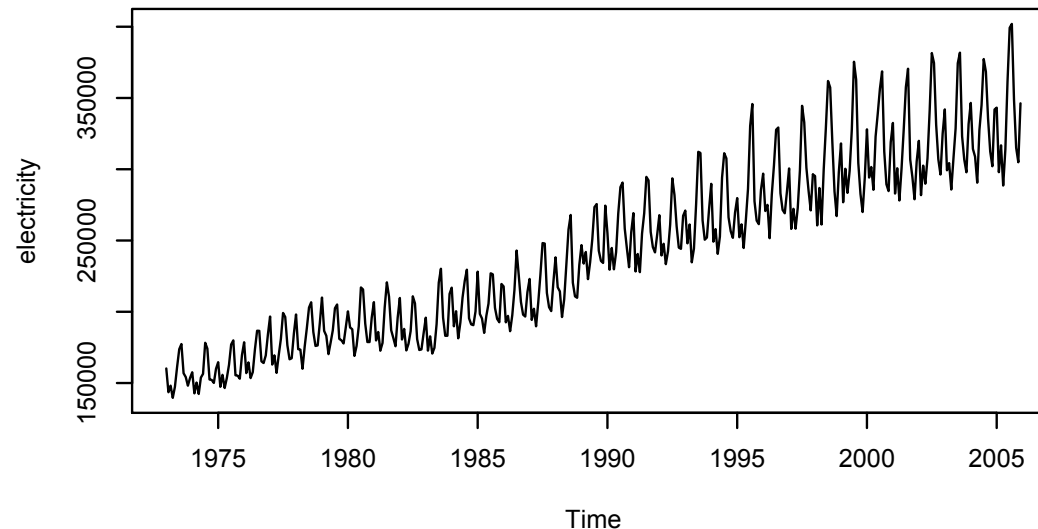
Si $|X_t| < 0.2$, nous pouvons approximer $\ln(1 + X_t)$ par X_t , pour ainsi modéliser $\nabla \ln(Y_t)$ par un processus stationnaire, i.e.

$$\nabla \ln(Y_t) \approx X_t \quad .$$

Illustrons quelque peu l'usage de cette transformation logarithmique à partir de la base de données « Electricity ».

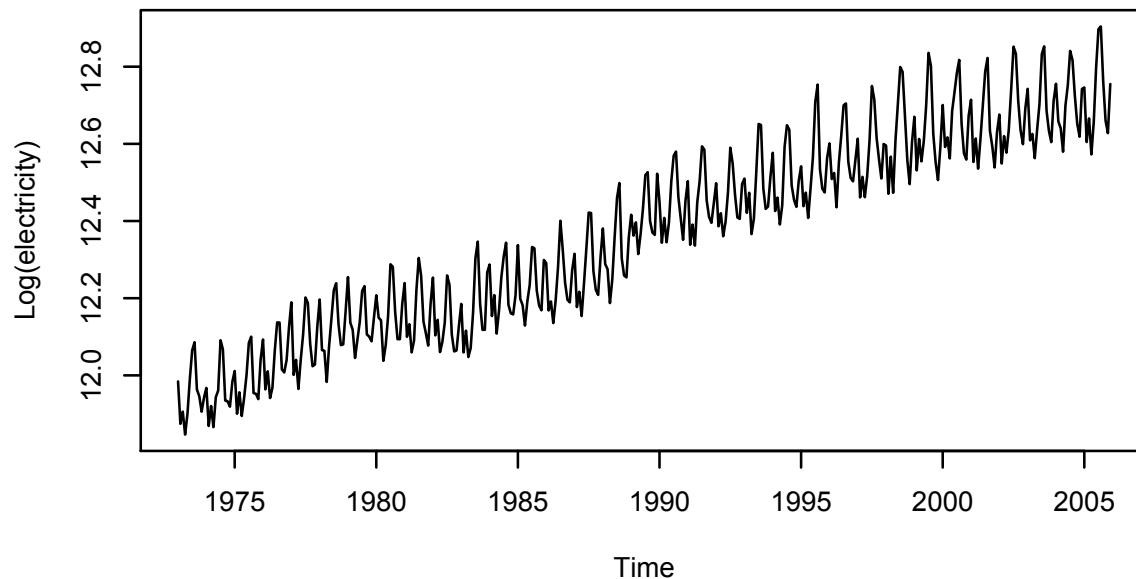
Montrons d'abord un graphique de cette série d'observations.

```
> data(electricity)
> win.graph(width=4.875, height=3, pointsize=8)
> plot(electricity)
```



Dressons maintenant le graphique du logarithme de ces valeurs.

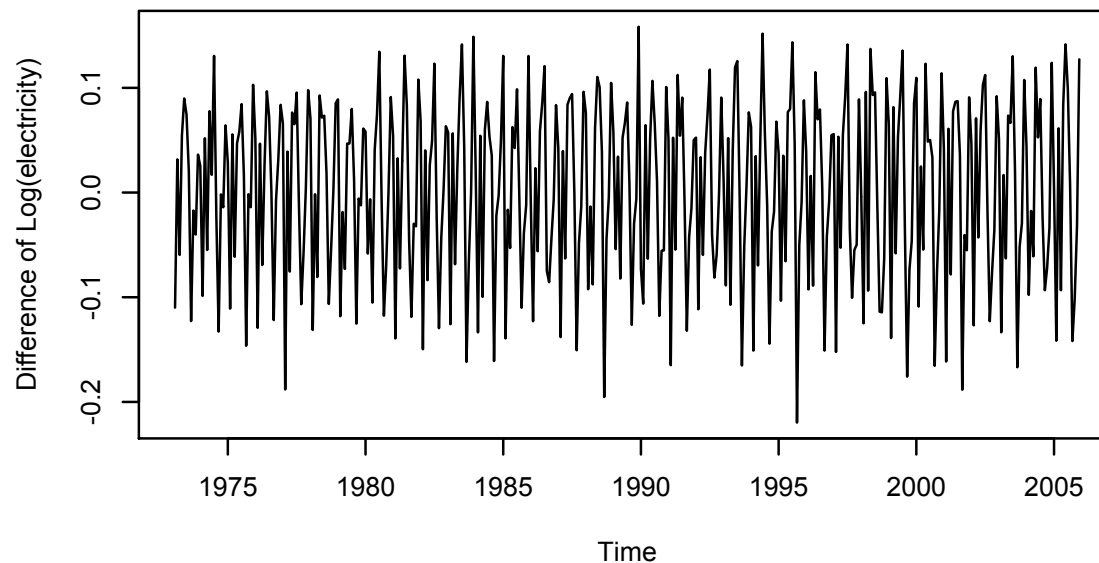
```
> plot(log(electricity), ylab='Log(electricity)')
```



Nous notons que les variations de ce graphique sont maintenant plus uniformes des plus petites aux plus grandes valeurs de $\ln(Y_t)$.

Montrons enfin un graphique de la première différence des logarithmes de notre série d'observations.

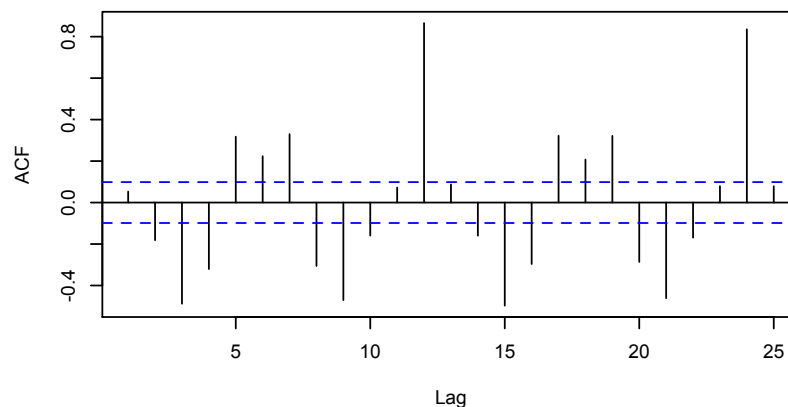
```
> plot(diff(log(electricity)), ylab='Difference of Log(electricity)')
```



La moyenne semble constante (près de 0) et la variance s'est stabilisée. Cependant, à partir de ce graphique, est-ce que cette différence des logarithmes est stationnaire? Il semble que non!

En effet, en regardant de plus près le graphique précédent, nous pourrions aussi conclure que la moyenne semble varier peu mais de manière cyclique. Nous renforçons cette observation dans le graphique suivant en montrant qu'il y a une structure de corrélation caractéristique de ce comportement cyclique.

```
> diff.log.elec=ts(diff(log(electricity)),freq=1,start=1)
> model9=lm(diff.log.elec~time(diff.log.elec))
> win.graph(width=4.875, height=3,pointsize=8)
> acf(rstudent(model9),main='')
```



5.5.2 Transformée de Box-Cox

Une famille de transformations présentant une grande flexibilité, toujours avec objectif de stabiliser la variance de la série chronologique initiale, est la famille des fonctions de puissance

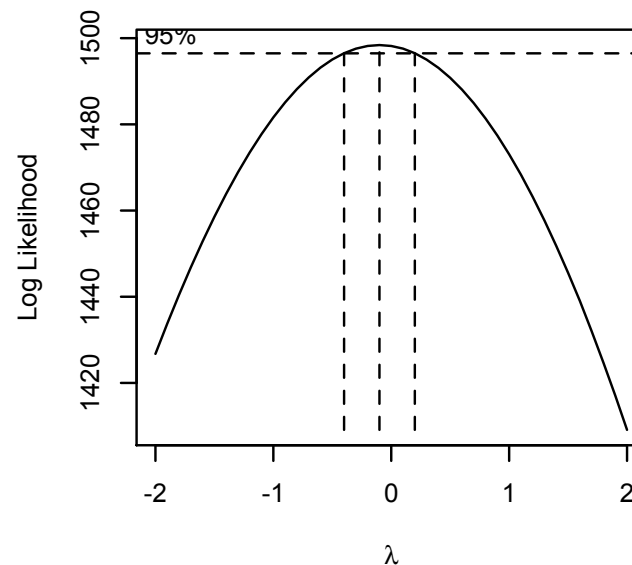
$$g(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} , & \lambda \neq 0 \\ \ln(x) , & \lambda = 0 \end{cases}$$

Cette forme s'applique à des données positives uniquement, quitte à modifier les données pour les rendre positives, et est choisie de sorte à ce que

$$\lim_{\lambda \rightarrow 0} \frac{x^\lambda - 1}{\lambda} = \ln(x) \quad .$$

La valeur du paramètre λ qui réalisera le mieux cet objectif sera celle qui maximisera la fonction de log-vraisemblance de nos données (sous l'hypothèse de la normalité). À titre d'exemple, faisons ce calcul avec la fonction R, `BoxCox.ar`, pour la base de données « Electricity » .

```
> win.graph(width=3, height=3, pointsize=8)  
> BoxCox.ar(electricity)
```



Nous constatons que $\lambda = -0.1$ semble être l'estimé MLE (maximum likelihood estimator) situé au centre de l'intervalle de confiance à 95%, soit $]-0.4, 0.2[$.

Donc, la transformée de Box-Cox ...

$$g(x) = 10(x^{-0.1} - 1)$$

... constitue un bon choix, mais comme $\lambda = 0$ est dans notre intervalle de confiance, la transformée logarithmique constitue aussi un choix qui n'est pas déraisonnable (comme exhibé dans les graphiques précédents).

* En conclusion, peu importe la transformée Box-Cox utilisée pour « stabiliser » la variance, l'étude de la stationnarité reste à faire.