Series Temporelles

Josephson Junior R.

February 23, 2024

Table des matières

Un processus $\{Y_t\}_{t\in\mathbb{Z}}$ est stationnaire si **ses moments d'ordre 1 et 2 sont indépendants par rapport au temps**.

Exemple : Le bruit blanc $\{\epsilon_t\}$

- $E(\varepsilon_t) = 0$
- $cov(\varepsilon_t, \varepsilon_h) = 0 \quad \forall t \neq h$
- $V(\varepsilon_t) = \sigma_\varepsilon^2$

- **Fonction d'autocovariance**

$$\gamma_h = cov(Y_t, Y_{t-h}) = E(Y_t Y_{t-h}) - E(Y_t)E(Y_{t-h})$$

- **Fonction d'autocorrélation**

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{cov(Y_t, Y_{t-h})}{\sqrt{V(Y_t)V(Y_{t-h)}}}$$

Définition

$\{Y_t\}_{t\in\mathbb{Z}}$ est un processus Auto-Régressif(p) s'il s'écrit en fonction **de ses observations antécédentes** c-a-d :

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_p y_{t-p} + \varepsilon_t$$

Les propriétés sont telles que :

- Toujours inversible
- $|\alpha_i| \leq 1 \quad \forall i \in [1, p]$
- $\psi(L) = 1 - \alpha_1 L - \alpha_2 L^2 - ... - \alpha_p L^p$
- $\mathbf{\{Y_t\}_{t\in\mathbb{Z}}}$ **est stationnaire** $\iff |\mathbf{L}| > \mathbf{1}$

Caractéristiques stochastiques

- Espérance

$$E(y_t) = \mu = 0$$

- Variance

$$V(y_t) = \gamma_0 = \alpha_1\gamma_1 + \sigma_\varepsilon^2$$

- Autocovariance d'ordre 1

$$cov(y_t, y_{t-1}) = \gamma_1 = \alpha_1\gamma_0$$

- Autocorrélation d'ordre 1

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\alpha_1\gamma_0}{\gamma_0} = \alpha_1$$

Remarques

- Fonction d'autocovariance

$$\gamma_k = \alpha_1^k \gamma_0 \quad \forall k \geq 0$$

- Fonction d'autocorrélation

$$\rho_k = \alpha_1^k \quad \forall k \geq 0$$

- Condition de stationnarité : $|L| \geq 0$

$$\psi(L) = 1 - \alpha_1 L = 0 \Longrightarrow L = \frac{1}{\alpha_1}$$

$$\{y_t\} \; \textit{est stationnaire} \iff \frac{1}{|\alpha_1|} \geq 0$$

Caractéristiques stochastiques

- Espérance

$$E(y_t) = \mu = 0$$

- Variance

$$\gamma_0 = \alpha_1\gamma_1 + \alpha_2\gamma_2 + \sigma_\varepsilon^2$$

- Autocovariance d'ordre 1

$$\gamma_1 = \alpha_1\gamma_0 + \alpha_2\gamma_1$$

- Autocorrélation d'ordre 1

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\alpha_1\gamma_0 + \alpha_2\gamma_1}{\gamma_0} = \alpha_1 + \alpha_2\rho_1$$

- Autocovariance d'ordre 2

$$\gamma_2 = \alpha_2\gamma_0 + \alpha_1\gamma_1$$

- Autocorrélation d'ordre 2

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{\alpha_2\gamma_0 + \alpha_1\gamma_1}{\gamma_0} = \alpha_2 + \alpha_1\rho_1$$

Les équations de Yule-Walker :
Les équations de Yule-Walker permettent de trouver les coefficients du processus AR(2) à travers :

$$\begin{cases} \rho_1 & = & \alpha_1 + \alpha_2\rho_1 \\ \rho_2 & = & \alpha_2 + \alpha_1\rho_1 \end{cases}$$

On peut genéraliser à l'ordre $k \geq 3$:

$$\begin{cases} \gamma_k & = & \alpha_1\gamma_{k-1} + \alpha_2\gamma_{k-2} \\ \rho_k & = & \alpha_1\rho_{k-1} + \alpha_2\rho_{k-2} \end{cases}$$

Définition

$\{Y_t\}_{t\in\mathbb{Z}}$ est un processus Moving Average(q) s'il s'écrit comme suit :

$$y_t = \varepsilon_t - \Theta_1\varepsilon_{t-1} - \Theta_2\varepsilon_{t-2} - ... - \Theta_q\varepsilon_{t-q}$$

Les propriétés sont telles que :

- Stationnaire
- $|\Theta_i| \leq 1 \quad \forall i \in [1, q]$
- $\phi(L) = 1 - \Theta_1 L - \Theta_2 L^2 - ... - \Theta_q L^q$
- Inversible si $|L| > 1$
- $\rho_k = 0 \ \forall k > q + 1$

Caractéristiques stochastiques

- Espérance

$$E(y_t) = 0$$

- Variance

$$V(y_t) = \gamma_0 = \sigma_\varepsilon^2 \left(1 + \Theta_1^2\right)$$

- Autocovariance d'ordre 1

$$\gamma_1 = -\sigma_\varepsilon^2 \Theta_1$$

- Autocorrélation d'ordre 1

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\Theta_1}{1 + \Theta_1^2}$$

Caractéristiques stochastiques

- Variance

$$\gamma_0 = \sigma_\varepsilon^2(1 + \Theta_1^2 + \Theta_2^2)$$

- Autocovariance d'ordre 1

$$\gamma_1 = \sigma_\varepsilon^2(\Theta_1\Theta_2 - \Theta_1)$$

- Autocorrélation d'ordre 1

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\Theta_1\Theta_2 - \Theta_1}{1 + \Theta_1^2 + \Theta_2^2}$$

- Autocovariance d'ordre 2

$$\gamma_2 = -\sigma_\varepsilon^2\Theta_2$$

- Autocorrélation d'ordre 2

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{-\sigma_\varepsilon^2 \Theta_2}{\sigma_\varepsilon^2(1 + \Theta_1^2 + \Theta_2^2)} = \frac{-\Theta_2}{1 + \Theta_1^2 + \Theta_2^2}$$

Théorème de Wold

Tout processus stationnaire peut être exprimé sous la forme d'un processus **MA(**∞) *comme suit* :

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

Définition

$\{Y_t\}_{t\in\mathbb{Z}}$ est un processus ARMA(p,q) s'il s'écrit comme suit :

$$\psi(L)Y_t = \phi(L)\varepsilon_t$$

Les propriétés sont telles que :

- $\{Y_t\}_{t\in\mathbb{Z}}$ admet une représentation **MA(**∞)

$$Y_t = \frac{\phi(L)}{\psi(L)}\varepsilon_t = \sum_{j=0}^{\infty} h_j\varepsilon_{t-j}; \quad h_0 = 1$$

- $\{Y_t\}_{t\in\mathbb{Z}}$ admet une représentation **AR(**∞)

$$\varepsilon_t = \frac{\psi(L)}{\phi(L)}Y_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j}; \quad \pi_0 = 1$$

Corrélogramme simple

Soit $\{Y_t\}_{t\in\mathbb{Z}}$ un processus ARMA(p,q) :

$$Y_t - \sum_{j=1}^{p} \alpha_j Y_{t-j} = \varepsilon_t - \sum_{j=1}^{q} \Theta_j \varepsilon_{t-j}$$

En multipliant par Y_{t-h} et en appliquant l'espérance de part et d'autre on obtient :

$$\gamma_h - \sum_{j=1}^{p} \alpha_j \gamma_{h-j} = E(\varepsilon_t Y_{t-h}) - \sum_{j=1}^{q} \Theta_j E(Y_{t-h} \varepsilon_{t-j})$$

Or $E(Y_{t-h}\varepsilon_{t-j}) = 0 \; si \; t-h < t-j$. Pour $h > q$ on rétrouve l'équation de récurrence d'ordre comme dans le cas d'un processus AR(p) :

$$\gamma_h - \sum_{j=1}^{p} \alpha_j \gamma_{h-j} = 0$$

Les p équations où $h > q$ sont appelées les **équations de Yule-Walker**.
Les premoères valeures de γ_h sont determinées à partir de :

$$\gamma_h - \sum_{j=1}^{p} \alpha_j \gamma_{t-j} = E(\varepsilon_t Y_{t-h}) - \sum_{j=1}^{q} \Theta_j E(Y_{t-h}\varepsilon_{t-j})$$

Où les termes de droites de l'égalité sont calculés à partir de l'expression MA(∞) *de* Y$_t$.
Donc pour $h > q$ le corrélogramme du processus se ramène à celui d'un processus AR(p).

Carastéristiques stochastiques

Soit le processus $\{Y_t\}$ qui s'ecrit comme suit :

$$Y_t = \alpha_1 Y_{t-1} + \varepsilon_t - \Theta_1 \varepsilon_{t-1}$$

- **Espérance** :

$$E(y_t) = \mu = 0$$

- **Variance** :

$$V(y_t) = \gamma_0 = \sigma_\varepsilon^2 \left[\frac{1 - 2\alpha_1\theta_1 + \theta_1^2}{1 - \alpha_1^2} \right]$$

- **Autocovariance d'ordre 1** :

$$\gamma_1 = \alpha_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2$$

- **Autocorrélation d'ordre 1** :

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \alpha_1 - \frac{\theta_1 \sigma_\varepsilon}{\gamma_0}$$

Une généralisation des coefficients d'autocorrélation nous donne :

$$\rho_{\mathbf{h}} = \alpha_{\mathbf{1}} \rho_{\mathbf{h-1}} \quad \forall \mathbf{h} > \mathbf{1}$$

Les chroniques économiques sont rarement des réalisations de processus aléatoires stationnaires. Parmi les explications de la non-stationnarité d'un processus on peut parler de l'accumulation des chocs ε_t au cours du temps. Pour analyser la non-stationnarité, deux types de processus sont distingués :

- Les processus TS (Trend Stationary) : qui représentent une non-stationnarité de **type déterministe** car determiné par **terme de tendance**.
- Les processus DS (Differency Stationary) : qui représentent une non-stationnarité de **type aléatoire**.

Définition

Un processus TS s'ecrit comme suit :

$$Y_t = \beta_0 + \beta_1 f(t) + \varepsilon_t$$

Le cas le plus simple est $\mathbf{f(t) = t}$ où l'on voit très bien que le processus n'est pas stationnaire car $\mathbf{E(Y_t)}$ dépend du temps.
Pour cette modélisation l'effet produit par un choc à un instant t est **transitoire**.

Remarques

Une bonne manière de stationnariser la série serait la MCO **si l'allure est à tendance linéaire**.

Définition

Un processus DS est un processus que l'on peut rendre stationnaire en appliquant un filtre de différence :

$$\Delta^d Y_t \;\; \textbf{est stationnaire}$$

d est appelé l'odre du filtre aux différences.

Types

- Processus sans dérive : **marche aléatoire**

$$Y_t = Y_{t-1} + \varepsilon_t \Longrightarrow (1-L)Y_t \; est \; stationnaire$$

- Processus avec dérive : $\beta \neq 0$

$$Y_t = \beta + Y_{t-1} + \varepsilon_t \Longrightarrow (1-L)Y_t \; est \; stationnaire$$

Tests de Dickey-Fuller (1979)

Les test DF permettent de mettre en évidence le caractère stationnaire ou non d'une chronique par l'identification d'une tendance **déterministe ou stochastique**.

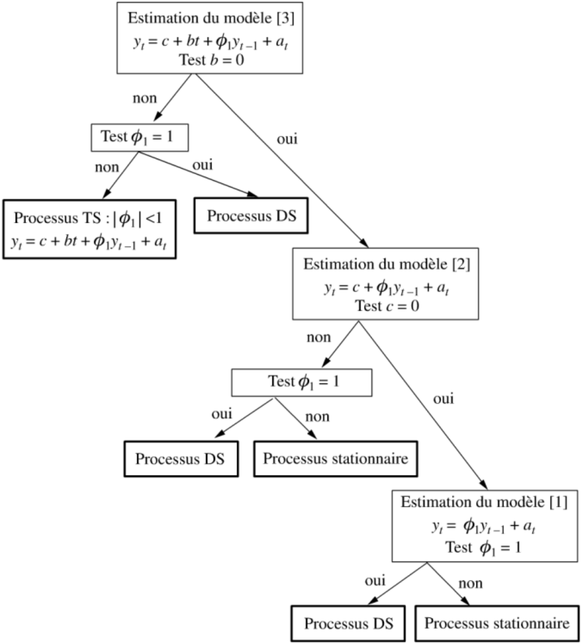
$$\begin{cases} H0 & : & \textbf{il existe une racine unitaire} \\ H1 & : & \textbf{absence de racine unitaire} \end{cases}$$

Ce test se base sur l'estimation et **l'étude séquentielle de 3 modèles de régression** :

1. $y_t = \phi y_{t-1} + c + bt + \varepsilon_t$
2. $y_t = \phi y_{t-1} + c + \varepsilon_t$
3. $y_t = \phi y_{t-1} + \varepsilon_t$

La règle de décision pour le test :

$$\textbf{si } \mathbf{t_{(DF)}} > \mathbf{t^c} : \textbf{On accepte H0}$$


Estimation du modèle [3]
$y_t = c + bt + \phi_1 y_{t-1} + a_t$
Test $b = 0$
non
oui
Test $\phi_1 = 1$
non
oui
Processus TS : $|\phi_1| < 1$
$y_t = c + bt + \phi_1 y_{t-1} + a_t$
Processus DS
Estimation du modèle [2]
$y_t = c + \phi_1 y_{t-1} + a_t$
Test $c = 0$
non
oui
Test $\phi_1 = 1$
oui
non
Processus DS
Processus stationnaire
Estimation du modèle [1]
$y_t = \phi_1 y_{t-1} + a_t$
Test $\phi_1 = 1$
oui
non
Processus DS
Processus stationnaire

Test de Dickey-Fuller Augmenté (1981)

Afin d'améliorer la puissance du test Dickey et Fuller ont proposé une nouvelle version en **introduisant des termes en différence retardés**. Il est fondé sur l'estimation par MCO des 3 modèles suivantes :

1. $\Delta y_t = \rho y_{t-1} + \sum_{j=1}^{p} \alpha_j \Delta y_{t-j} + c + bt + \varepsilon_t$
2. $y_t = \rho y_{t-1} + \sum_{j=1}^{p} \alpha_j \Delta y_{t-j} + c + \varepsilon_t$
3. $y_t = \rho y_{t-1} + \sum_{j=1}^{p} \alpha_j \Delta y_{t-j} + \varepsilon_t$

La règle de décision pour le test :

$$\textbf{si } \mathbf{t_{(ADF)}} > \mathbf{t^c} : \textbf{On accepte H0}$$

Box et Jenkins (1976) ont proposé une demarche générale de prévision pour une série univariée fondée sur la notion du processus **ARIMA(p,d,q)** qui se déroule en **4 étapes**.

Etape 1 : Identification du modèle

Dans cet étape on va déterminer des valeurs possibles **p,d,q** en utilisant le test **ADF** et les correlogrammes simples et partielles.
Après stationnarisation de la série nous pouvons identifier les degrés p,q du modèle di modèle ARMA tel que :

- si le **corrélogramme simple** n'a que ses **q premiers termes** $\neq$ **0** et que le **corrélogramme partiel** diminue lentement alors le processus est de **type MA(q)**
- si le **corrélogramme partiel** n'a que ses **p premiers termes** $\neq$ **0** et que le **corrélogramme simple** diminue lentement alors le processus est de **type AR(p)**

Etape 2 : Estimation

Il s'agit d'estimer les coefficients du modèle $\alpha_{\mathbf{i}}$ et $\theta_{\mathbf{j}}$ avec :

- La méthode des MCO
- La méthode du maximum vraisemblance

Etape 3 : Validation et choix du modèle adéquat

Consiste à verifier la qualité du modèle candidat en utilsant des méthode comme : **Test d'autocorrélation des résidus** ; **Test de nullité des coefficients estimés** et autres critère de séléction du modèle par la qualité d'information tel que AIC ou BIC.

$$AIC = T\ ln\left(\frac{SCR}{T}\right) + 2k \quad ; \quad BIC = T\ ln\left(\frac{SCR}{T}\right) + kln(T)$$

Le modèle à retenir est celui ayant la valeur de AIC ou BIC **le plus faible**.

Etape 4 : Prévision

On remplace les paramètres inconnus par les estimations obtenues du modèle choisi pour pouvoir faire de la prévision.
Cet étape fera l'objet de la matière **Technique de prévision**.