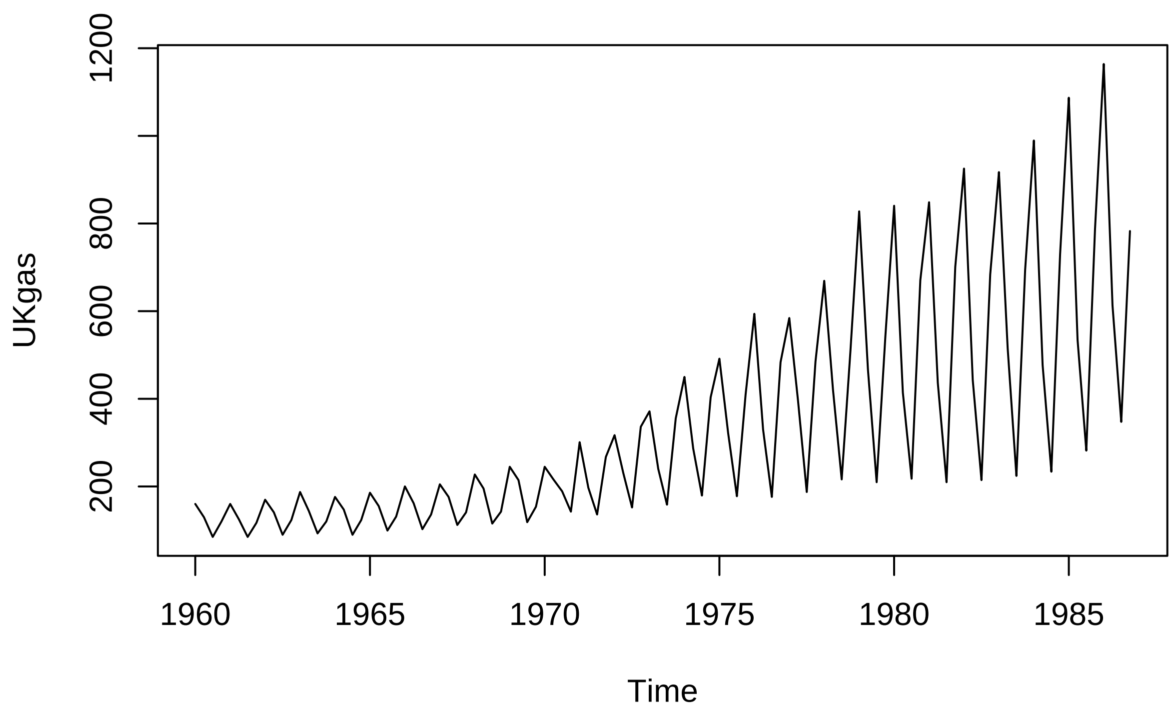
Note de synthèse sur les Series Chronologiques

Hissein Tidei*

2023-05-06

1 Nature des series



1.1 Concepts de Base

Qu'est ce qu'une série chronologique ou Time Series (TS)?

C'est une suite finie des observations effectuer à un intervalle régulier au cours du temps. Dite aussi série temporelles, elle se matérialise économiquement par l'observation des grandeurs telles que: IPC, PNB, PIB, etc.

La notion de processus stationnaire au sens large (SSL) d'une TS :

Un processus est X_t est SSL, au sens faible, ou dite de "second ordre" si et seulement si:

*Twitter

1. $\mathbb{E}(X_t) = \mu$: constant ou ne croit pas avec le temps (t): démunie de tendance;
2. $\mathbb{E}(X_t^2) = Var(X_t) = \sigma^2 \neq \infty$: c'est dire constant et ne tend pas vers l'infinie au fil du temps infinie au fil du temps;
3. $\gamma(k)$, sa fonction d'auto-covariance , est dépendante de l'ampleur du décalage k et non de la position en t, celle-ci étant:
   - $\gamma(k) = Cov(X_t, X_{t+k}) = E\{(X_t - \mathbb{E}(X_t))(X_{t+k} - \mathbb{E}(X_{t+k}))\} \;\; \forall k \in \mathbb{Z}$
   - $\gamma(k=0) = \gamma(0) = \sigma_x^2 = Var(X)$ => confirme le respect de (2.).

La fonction d'auto-corrélation d'un processus X_t SSL

Confondue avec la covariance dans le cadre de la stationnarité des séries chrono, elle écrit comme suit [1]:

$$\rho(k) = cor(X_t, X_{t+k}) = \frac{Cov(X_t, X_{t+k})}{Var(X_t)\,Var(X_{t+k})} \Longrightarrow \frac{\gamma(k)}{\gamma(0)} \tag{1}$$

.

Bruit blanc (BB) ou White noise

a- ***Simple BB***:

$X_t \rightsquigarrow BB\left(0, \sigma^2\right)$ si :

- $\mathbb{E}(X_t) = 0$;
- $Var(X_t) = \sigma^2$;
- $\gamma(k) = Cov(X_t, X_{t+k}) = 0 \quad \forall k \neq O \Longrightarrow$
  - $\gamma(k = o) = \sigma^2$
  - $\gamma(k \neq o) = 0$

NB: Un BB est processus SSL.

b- ***BB Gaussien(BBG)***:

$X_t \rightsquigarrow BBG\left(0, \sigma^2\right)$ *ssi* :

- $X_t : i.i.d \rightsquigarrow N\left(0, \sigma^2\right)$

Attention : si $X_t \rightsquigarrow BB\left(0, \sigma^2\right) \longrightarrow \rho(k) = \frac{\gamma(k)}{\gamma(0)}$

De quoi il s'agit et si X__t n'est pas SSL

Il faut comprendre que si un processus X_t n'est pas SSL, sa non stationnarité decoule du

1.2 Un processsus moyen Mobile d'ordre q: MA(q)

1.2.1 Un MA(q=1):

$MA(1)$ est un processus X_t, tel que :

$$X_t = \mu + a_t + \theta a_{t-1} \tag{2}$$

où:

$$a_t \rightsquigarrow BB\left(0, \sigma^2\right) \; et \; \theta \in \mathbb{R} \tag{3}$$
$$\Longrightarrow \mathbb{E}(a_t) = 0 \; \forall t, \; Var(a_t) = \sigma^2, \; Cov(a_t, a_{t+k}) = \sigma^2 \mu \; \Leftrightarrow \; \{k = 0\} \tag{4}$$

avec comme:

[1] Cf au démonstration

$$\mathbb{E}\left(X_t\right)=\mu \quad (5)$$

$$Var\left(X_t\right)=\left(1+\theta^2\right)\sigma^2 \quad (6)$$

$$\gamma\left(k\right)=Cov\left(X_t,X_{t+k}\right)=\begin{cases}\left(1+\theta^2\right)\sigma^2=Var(X_t) & \text{si } k=O\\ \theta\sigma^2 & \text{si } K=1\\ 0 & \text{si } k\geqslant 2\end{cases} \quad (7)$$

Ainsi, tout processus $MA(1)$ est un SSL. Et la fonction d'auto-corrélation (FAC) pour un $MA(1)$ est telle que:

$$FAC=\rho\left(k\right)=Cov\left(X_t,Xt+k\right)=\begin{cases}1 & \text{si } k=O\\ \frac{\theta}{1+\theta} & \text{si } K=1\\ 0 & \text{si } k\geqslant 2\end{cases} \quad (8)$$

1.2.2 Un MA(q=2):

Pour un tel processus la forme est la suivante:

$$X_t=\mu+a_t+\theta_2 a_{t-1}+\theta_2 a_{t-2} \quad (9)$$

où encore

$$a_t\rightsquigarrow BB\left(0,\sigma^2\right)\ et\ \{\theta_2,\theta_1\}\in\mathbb{R} \quad (10)$$

$$\mathbb{E}\left(X_t\right)=\mu \quad (11)$$

$$Var\left(X_t\right)=\left(1+\Sigma_{j=1}^{2}\theta_j^2\right)\sigma^2\ si\ 1\leqslant k\leqslant 2 \quad (12)$$

$$\gamma\left(k\right)=\begin{cases}\left(\theta_k+\Sigma_{j=1}^{2-k}\theta_j\theta_{k+j}\right)\sigma^2 & \text{si } 1\leqslant k\leqslant 2\\ 0 & \text{si } k>2\end{cases} \quad (13)$$

De même tout processus $MA(2)$ est un SSL et le FAC s'annule lorsque $k>2$.

1.2.3 Un MA(q):

Pour tout $q>2$, une série $X_t $ MA(q) écrit comme suit:

$$X_t=\mu+a_t+\theta_2 a_{t-1}+\theta_2 a_{t-2}+\cdots+\theta_q a_{t-q} \quad (14)$$

$$=\mu+a_t+\Sigma_{j=1}^{Q}\theta_j a_{t-j} \quad (15)$$

où encore

$$a_t\rightsquigarrow BB\left(0,\sigma^2\right)\ et\ \theta_j\in\mathbb{R} \quad (16)$$

et

$$\mathbb{E}\left(X_t\right)=\mu \quad (17)$$

$$Var\left(X_t\right)=\left(1+\Sigma_{j=1}^{q}\theta_j^2\right)\sigma^2 \quad (18)$$

$$\gamma\left(k\right)=\begin{cases}\left(\theta_k+\Sigma_{j=1}^{q-k}\theta_j\theta_{k+j}\right)\sigma^2 & \text{si } 1\leqslant k\leqslant q\\ 0 & \text{si } k>q\end{cases} \quad (19)$$

1.2.4 Sur le Plan pratique:

En générale, la modélisation d'une série chronologique des données réelles, comprise comme obéissant un processus à moyen mobile, pour faire de la prévision requière la compréhension de la structure de ces données, son fonctionnement et sa mode de détermination interne ou exogène.En l'occurence ici, une telle compréhension passe par la détermination de l'ordre q, indiquant le k à partir duquel l'auto-corrélation entre les différentes périodes ou temps de la série univariée s'estompe [2]; en d'autre terme le q détermine les niveaux, en t, de relations d'interdépendances ou de liaisons temporelles au sein de la série. Globalement, ces relations sont basées sur des composantes tendancières (trend) ou stochastiques entre deux ou plusieurs sous-séries d'un échantillon (X_t). Ainsi, se pose la question , en supposant par observation que la série à des composantes tendancière et ou stochastique, de l'ordre q ou le k à partir duquel l'auto-covariances intra-observation de la série ne sont plus significatif?[3] Formellement, modéliser une série revient, signifiant par ailleurs comprendre la structure interne qui génere et lie les observation entre elle[4], à d'abord détecter les liens causaux, mêmes partiels mais significatifs, entre plusieurs niveaux temporels d'une série (X_t); donc pour un $MA(Q)$ le q est déterminer comme suit:

$$\begin{cases} \rho(k) \neq 0 & \text{si } 1 \leqslant k \leqslant q \\ \rho(k) = 0 & \text{si } k \geqslant q+1 \end{cases}$$

En d'autre terme, et en pratique $\rho(.)$ étant l'AC de la série, il faut disposer de la FAC (auto-corrélogramme) qui donne l'AC a différentes périodes de la série pour pouvoir déterminer l'ordre q. Ce, au moyen de l' équation précédente, q obtenu en identifiant la période à partir de laquelle la FAC s'annule ($\rho(k) = 0$); ainsi le k qui en ressort donne les valeurs des différentes q qui lui sont inférieurs et qui pourront modéliser la série en $MA(q)$, l'ordre(q) s'arrête pour les valeurs auxquelles k deviendraient nuls, indiquant une faible liaison temporelles pour les k supérieurs à q.

1.2.5 Determination ou estimation des FAC:

comment avoir les $\gamma(k)$ et $\rho(k)$?

1.2.5.1 Estimation de la fonction d'auto-covariance:

Soit la fonction $\gamma(k)$ d'un processus SSL d'une suite de T observations de X, tel que un :

$X_t : \{X_1, X_2, \ldots, X_T\}$

Et soit la moyenne empirique de cette serie T (population):

$\bar{X}_t = \frac{1}{T}\Sigma_{j=1}^{T} X_j$

L'estimateur ,$\hat{\gamma}$, (pour un échantillon) de la fonction de γ est donnée comme suit :

$$\gamma(k) = E\{(X_t - \mathbb{E}(X_t))(X_{t+k} - \mathbb{E}(X_{t+k}))\} \Longrightarrow \hat{\gamma} = \frac{1}{T}\Sigma_{j=1}^{T-k}\left(X_j - \bar{X_T}\right)\left(X_{j+k} - \bar{X_T}\right), \ \forall\ 0 \leqslant k \leqslant T-1$$

1.2.5.2 Estimation de la fonction d'auto-corrélation:

l'estimateur de ρ ($\hat{\rho}$), désignant la FAC d'un X_t processus SSL, est donnée comme suit:

$$\rho = \frac{\gamma(k)}{\gamma(0)} \Longleftrightarrow \hat{\rho} = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} \Longrightarrow \hat{\rho(k)} = \frac{\Sigma_{j=1}^{T-k}\left(X_j - \bar{X_T}\right)\left(X_{j+k} - \bar{X_T}\right)}{\Sigma_{j=1}^{T}\left(X_j - \bar{X_T}\right)^2}$$

Ainsi nous avons:

- par exemple : $\hat{\rho(1)} = \frac{\Sigma_{j=1}^{T-1}\left(X_1 - \bar{X_T}\right)\left(X_{j+1} - \bar{X_T}\right)}{\Sigma_{j=1}^{T}\left(X_1 - \bar{X_T}\right)^2}$
- Etudiant la correlation entre les ensembles X_1 et X_{j+K}, tel que :
  - $X_1 : (x_1, x_2, \ldots, x_{T-1})$

[2] Comme en (19), pour un $k > q$, la covariance ou, peu importe ici, l'auto-corrélation au sein de la série s'annule, c-à-d qu'il n'y a plus de lien en moyen mobile depassé k temps ou période entre deux plusieurs observations d'une série temporelle.

[3] Il faut garde en tête que les séries sont considérés comme ayant des mémoire, d'où hypothétiquement il peut toujours subsister d'influence d'une observation sur ses prochaines, qui à leur tour sur les autres prochaines même lorsque celle-ci est consideré morte.

[4] Parfois dit la boite noire des données

– $X_{j+1} : (x_2, x_2, \ldots, x_T)$

NB: Sous Eviews, le grahph de $\hat{\rho}$, en fonction de k, dit corrélegramme est donnée par le programme et permet de decider de l'ordre de q, car pour un $MA(q)$, puisque q se trouve être inferieur à toute les k où celui-ci annule le $\hat{\rho}$, soit $\rho(\hat{k}) = 0$, le q qui permetrait modeliser la serie significativement est une valeur inferieure à à la premiere k qui annule le ρ.

1.3 Un processus autoregressif: AR(p)

Un classe des modèles de processus *Stochastique*

1.3.1 Un AR(P=1):