Chapitre 4: Les modèles linéaires

1. Introduction:

- Dans ce chapitre on va voir successivement les modèles linéaires stationnaires: auto-regressifs *(AR)*, de moyennes mobiles *(MA)* et mixtes *(ARMA)* en particulier.
- Finalement, parmi les modèles non stationnaires, on verra les modèles *ARIMA*.

- Dans ce chapitre, la méthode utilisée pour exposer les modèles linéaires est la méthode déductive.
- Ainsi, à partir des hypothèses initiales, on obtiendra les propriétés des différents types de modèles.
- La connaissance de ces propriétés va faciliter la réalisation de l'inférence que l'on verra dans les chapitres suivants.

2. Opérateur retard:

- Soit (Y_t) un processus stochastique. On définit l'opérateur retard "L" tel que:

  $LY_t=Y_{t-1}$

  $L^2Y_t=L[LY_t]=LY_{t-1}=Y_{t-2}$

  En général: $L^kY_t=Y_{t-k}$
- On a aussi, pour c une constante:

  $Lc=c$ et $L(cY_t)=cLY_t=cY_{t-1}$
- $L^kL^s(Y_t)=L^{k+s}(Y_t)=Y_{t-k-s}$

- Si $LY_t=Y_{t-1}$, alors $Y_t=L^{-1}Y_{t-1}$.
- L'opérateur retard identique ou unité est $L^0=I=1$.
- L'application de l'opérateur retard unité ne change pas la période de référence:

  $L^0Y_t=Y_t$.
- L'opérateur L peut être utiliser pour exprimer un modèle de retards. Soit, par exemple, le modèle suivant:

$$Y_t-\Phi_1Y_{t-1}-\Phi_2Y_{t-2}-\ldots-\Phi_pY_{t-p}=e_t$$

- Si on applique l'opérateur L, on obtient

$$[1- \Phi_1 L-\Phi_2 L^2-\ldots-\Phi_p L^p]Y_t=e_t$$

- L'expression entre crochets s'appelle "opérateur polynomial des retards" et on le note $\Phi(L)$. La lettre Φ qui précède (L) indique les coefficients qui interviennent dans l'expression de l'opérateur polynomial des retards.
- Donc, on peut écrire: $\Phi(L)Y_t=e_t$.
- L'expression $\Phi(L)=0$, s'appelle équation polynômiale.

3. Modèles auto-regressifs (AR):

- Dans le chapitre précédent, un modèle auto-regressif d'ordre p, ou un modèle AR(p), a été présenté par:

$$Y_t=\Phi_1 Y_{t-1}+\Phi_2 Y_{t-2}+\ldots+\Phi_p Y_{t-p}+e_t$$

  Où e_t est une variable "bruit blanc".
- En utilisant l'opérateur polynomial des retards "L":

$$\Phi(L)= 1- \Phi_1 L-\Phi_2 L^2-\ldots-\Phi_p L^p$$

- Le modèle peut s'exprimer de forme simplifiée: $\Phi(L)Y_t=e_t$

- Par la suite, on va analyser les caractéristiques des modèles $AR(1)$ et $AR(2)$. Les résultats obtenus seront généralisés à un modèle $AR(p)$.

Modèle AR(1):

Un modèle $AR(1)$ est donné par:

$$Y_t=\Phi_1 Y_{t-1}+e_t$$

Ou, en utilisant l'opérateur retard, par:

$$(1-\Phi_1 L)Y_t=e_t.$$

$$\begin{tikzcd}
Y_{t-2} \arrow[r] & Y_{t-1} \arrow[r] & Y_t \arrow[r] & Y_{t+1} \\
e_{t-2} \arrow[u] & e_{t-1} \arrow[u] & e_t \arrow[u] & e_{t+1} \arrow[u]
\end{tikzcd}$$

- Chaque variable "bruit blanc" influe sur les valeurs de Y_i correspondantes à la même période ou sur des périodes posterieures, mais jamais sur des périodes antérieures.

- Une conséquence importante est:

$$E[e_t Y_{t-\tau}]=0 \qquad \text{pour tout } \tau>0$$

- La racine de l'équation polynômiale $1\text{-}\Phi_1 L=0$ est

$$L = \frac{1}{\Phi_1}$$

- Le processus $AR(1)$ est stationnaire si $|L| = \left|\frac{1}{\Phi_1}\right| > 1$

- Donc $AR(1)$ est stationnaire si $|\Phi_1|<1$.

- Pour analyser le comportement du processus $AR(1)$, on supposera qu'au début le processus commence par la période $-N$.

- Par des substitutions consécutives on aura:

$Y_t=\Phi_1[\Phi_1 Y_{t-2}+e_{t-1}]+e_t=\ldots= \sum_{j=0}^{t+N-1} \Phi_1^j e_{t-j} + \Phi_1^{t+N} Y_{-N}$

- Si on calcul, alors, les epérances à partir de la dernière expression, on aura:

$$E\left[Y_t\right] = \Phi_1^{t+N} Y_{-N}$$

- Où Y_{-N}, qui représente les conditions initiales, a été considérée comme une variable non stochastique. En général, on suppose qu'un processus stochastique commence en une période infiniment petite.
- Si $|\Phi_1|<1$, alors $E\left[Y_t\right] = \Phi_1^{t+N} Y_{-N}$ diminuera en valeur absolue lorsque t augmente.

- Lorsque $|\Phi_1|>1$, alors $E[Y_t] = \Phi_1^{t+N} Y_{-N}$ augmentera en valeur absolue lorsque t augmente à condition que Y_{-N} ne soit pas exactement égale à zero.
- Si $\Phi_1=1$, alors on vérifie que $E(Y_t)= Y_{-N}$
- Si $\Phi_1= -1$, alors on aura une alternance de signes dans la valeur de l'espérance.
- Par conséquent, pour un processus qui commence en un instant fini du temps, si la valeur initiale n'est pas nulle, la moyenne de Y_t reste constante seulement lorsque $\Phi_1=1$.

- Mais, si le processus considéré commence en $-\infty$, alors pour $|\Phi_1|<1$ et pour n'importe quelle valeur initiale, on vérifie que $E[Y_t]=0$.
- Si on inclu une constante dans le modèle $AR(1)$, on aura $Y_t=\delta+\Phi_1 Y_{t-1}+e_t$
- Alors, si le processus commence en $-\infty$ et il est stationnaire, on vérifie que la moyenne du processus sera constante pour toute valeur de t: $E[Y_t]= E[Y_{t-1}]=\mu$

- En prenant les espérances sur $Y_t=\delta+\Phi_1 Y_{t-1}+e_t$ et en tenant compte que $E[Y_t]= E[Y_{t-1}]=\mu$, on obtient que: $\mu = \frac{\delta}{1-\Phi_1}$
- Dans la suite, on supposera, sans perte de généralité, que $\delta=0$.
- Un processus initié en $-N$ aura la variance suivante:

$$\gamma_0 = E\left[Y_t - E(Y_t)\right]^2 = E\left[Y_t - \Phi_1^{t+N} Y_{-N}\right]^2$$

$$\gamma_0 = E\left[\sum_{j=0}^{t+N-1} \Phi_1^j e_{t-j}\right]^2 = \sigma_e^2 \sum_{j=0}^{t+N-1} \Phi_1^{2j}$$

Et par la formule de la somme des termes d'une progression géométrique, on a:

$$\gamma_0 = \sigma_e^2 \frac{1-\Phi_1^{2(t+N-1)}}{1-\Phi_1^2}$$

Finalement, on peut dire que soit dans le cas $\Phi_1=1$, ou dans le cas $|\Phi_1|>1$, la vriance trouvée précédement tend vers l'infini si le processus commence par -∞. Alors, on dit que des processus de ces caractéristiques ont une nature explosive!

- Si le processus est stationnaire, c'est-à-dire, si $|\Phi_1|<1$ et il a commencé en -∞, on vérifie que:

$$\gamma_0 = \frac{\sigma_e^2}{1-\Phi_1^2}$$

• Dans ce qui suit, et sauf mention contraire, on supposera que le processus débute par -∞ et qu'il est stationnaire.

• Pratiquement, même si le processus débute par une valeur finie, sa variance se stabilise rapidement au voisinage de

$$\gamma_0 = \frac{\sigma_e^2}{1-\Phi_1^2}$$

- Si on multiplie les deux membres de $Y_t=\Phi_1 Y_{t-1}+e_t$ par $Y_{t-\tau}$, on vérifie que:

$$\gamma_\tau=E[Y_t Y_{t-\tau}]=\Phi_1 E[Y_{t-1}Y_{t-\tau}]+E[e_t Y_{t-\tau}]$$

- Pour $\tau>0$, et comme dans le diagramme des flèches vu précédement, on vérifie que:

$$\gamma_\tau=\Phi_1\gamma_{\tau-1} \qquad \tau>0$$

- Cette dernière équation $\gamma_\tau-\Phi_1\gamma_{\tau-1}=0$ est une équation aux différences homogène de premier ordre. De même, l'équation $Y_t-\Phi_1 Y_{t-1}=e_t$ est une équation aux différences de premier ordre mais non homogène à cause de la présence du "bruit blanc" e_t.

❖Équations aux différences:

- Soit le modèle suivant:

$$Y_t - \Phi_1 Y_{t-1} - \Phi_2 Y_{t-2} - \ldots - \Phi_p Y_{t-p} = e_t$$

- Si on applique l'opérateur L, on obtient

$$[1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p] Y_t = e_t$$

- Sachant que $\Phi(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p$

on peut écrire: $\Phi(L) Y_t = e_t$

- Cette dernière équation est une équation aux différences non homogène parce que son second membre est non nul, c'est e_t.

- L'ordre de cette équation aux différences est l'ordre de la différence la plus élevée

–ou encore, la variable de plus grand retard- qui apparaît dans l'équation.

- Donc l'équation considérée ici est d'ordre p.
- L'équation sans second membre est:

$$Y_t - \Phi_1 Y_{t-1} - \Phi_2 Y_{t-2} - \ldots - \Phi_p Y_{t-p} = 0$$

ou encore $\Phi(L) Y_t = 0$ c'est l'équation homogène correspondante.

- Résoudre ces équations c'est trouver une expression de Y_t en fonction du temps qui vérifie l'équation.
- Il n'existe pas une solution unique d'une de ces équations.
- Une solution qui englobe toutes les solutions s'appelle la solution complète Y_t^c.
- La solution complète est la somme de deux solutions: la solution générale de l'équation homogène Y_t^h et une solution particulière de l'équation non homogène Y_t^p

$$Y_t^c=Y_t^h+Y_t^p$$

❖Résolution de l'équation homogène:

- La solution générale de l'équation homogène doit être une fonction du temps qui vérifie l'équation

$$Y_t-\Phi_1Y_{t-1}-\Phi_2Y_{t-2}-\ldots-\Phi_pY_{t-p}=0.$$

- Elle dite générale parce qu'elle doit être valide quelques soient les conditions initiales du processus.
- Essayons avec une solution du type: $Y_t=\lambda^t$

- On la substitue dans l'équation, on obtient:

$$\lambda^t-\Phi_1\lambda^{t-1}-\ldots-\Phi_p\lambda^{t-p}=0$$

- On factorise par λ^{t-p}, on a:

$$\lambda^{t-p}[\lambda^p-\Phi_1\lambda^{p-1}-\ldots-\Phi_p]=0$$

- Une solution triviale qui satisfait l'équation antérieure est $\lambda=0$.

  Ou $\lambda^p-\Phi_1\lambda^{p-1}-\ldots-\Phi_p=0$, équation qui admet p racines qu'on note $\lambda_1, \lambda_2,\ldots, \lambda_p$ et on l'appelle équation caractéristique.

- Alternativement, les racines peuvent être obtenu à partir de l'équation polynômiale $\Phi(L)=0$ ou encore $1- \Phi_1L-\Phi_2L^2-\ldots-\Phi_pL^p=0$ qui a p racines notées $L_1,L_2,\ldots,L_p$.
- On peut vérifier que les racines L_i sont exactement les valeurs inverses des racines λ_i.
- En effet, si L_i est une racine, on a:

$$1- \Phi_1L_i-\Phi_2L_i^2-\ldots-\Phi_pL_i^p=0$$

- D'autre part, et si $\lambda_i = \dfrac{1}{L_i}$ on a:

$$\left[\frac{1}{L_i}\right]^p - \Phi_1\left[\frac{1}{L_i}\right]^{p-1} - \ldots - \Phi_p = 0$$

- Si on multiplie les membres de cette équation par L_i^p, on obtient: $1-\Phi_1 L_i - \Phi_2 L_i^2 - \ldots - \Phi_p L_i^p = 0.$
- on a λi est une racine de l'équation caractéristique, alors $Y_t = \lambda_i^t$ est une solution de l'équation homogène aux différences. On l'appelle solution basique.
- Deux théorèmes importants s'utilisent pour obtenir la solution générale:

- Le premier théorème dit que si λ_i^t est une solution et A_i est une constante arbitraire, alors $Y_t = A_i \lambda_i^t$ est aussi une solution.

- En effet, en substituant cette valeur dans l'équation $Y_t - \Phi_1 Y_{t-1} - \Phi_2 Y_{t-2} - \ldots - \Phi_p Y_{t-p} = 0$ et en factorisant par $A_i \lambda_i^{t-p}$, on obtient:

$$A_i \lambda_i^{t-p}[\lambda_i^p - \Phi_1 \lambda_i^{p-1} - \ldots - \Phi_p] = 0$$

puisque l'expression entre crochets est égale à 0.

- Par conséquent, $A_1 \lambda_1^t$, $A_2 \lambda_2^t$,…, $A_m \lambda_m^t$ ou A_1, A_2,…, A_m sont des des constantes arbitraires, sont aussi des solutions de l'équation homogène. C'est à dire, les multiplications d'une solution basique par une constante arbitraire sont des solutions.
- Le deuxième théorème dit que si λ_i^t et λ_j^t sont des solutions de l'équation homogène, alors une combinaison linéaire des deux solutions $A_i \lambda_i^t + A_j \lambda_j^t$ est aussi une solution de l'équation homogène.

- Cette propriété se vérifie en substituant l'expression antérieure dans l'équation homogène:

$$(A_i \lambda_i^t + A_j \lambda_j^t) - \Phi_1 (A_i \lambda_i^{t-1} + A_j \lambda_j^{t-1}) - \ldots - \Phi_p (A_i \lambda_i^{t-p} + A_j \lambda_j^{t-p}) = 0$$

- Puis, on factorise par $A_i \lambda_i^{t-p}$ d'une part et par $A_j \lambda_j^{t-p}$ d'autre part, on obtient:

$$A_i \lambda_i^{t-p}[\lambda_i^p - \Phi_1 \lambda_i^{p-1} - \ldots - \Phi_p] + A_j \lambda_j^{t-p}[\lambda_j^p - \Phi_1 \lambda_j^{p-1} - \ldots - \Phi_p] = 0$$

Puisque les deux expression entre crochets sont nulles.

- Ce deuxième théorème permet d'établir qu'une combinaison linéaire des deux solutions est aussi une solution.
- Précisément, la solution générale de l'équation homogène est composée par les p solutions basiques, multipliées par des constantes arbitraires. C'est à dire:

$$Y_h=Y_t=A_1\lambda_1^{\,t}+A_2\lambda_2^{\,t}+\ldots+A_p\lambda_p^{\,t}$$

qui est la solution générale de l'équation homogène.

- Il est intéressant d'analyser le comportement de Y_t lorsque t augmente indéfiniment.
- D'après l'expression précédente de Y_t, on peut remarquer que lorsque $|\lambda_i|<1,\ \forall i$, alors lorsque $t \to \infty$, il arrive que $Y_t \to 0$ indépendement des valeurs des constantes arbitraires.
- On dit alors que le système est "stable". Pour les processus stochastiques, on utilise plutôt le terme "stationnaire".
- Une condition nécessaire et suffisante de stabilité ou de stationarité est donc $|\lambda_i|<1,\ \forall i$.

- La solution générales est valable pour n'importe quelles valeurs des constantes arbitraires. Mais en pratique, on leur donne des valeurs obtenues à partir des conditions initiales.

- Ainsi, si on suppose que les valeurs Y_0, $Y_1,\dots, Y_{p-1}$ correspondantes aux périodes $t=0, t=1,\dots, t=p-1$ respectivement, on peut considérer le système de $p-1$ équations suivant:

$$\begin{cases} Y_0=A_1+A_2+\dots+A_p \\ Y_1=A_1\lambda_1+A_2\lambda_2+\dots+A_p\lambda_p \\ \quad\vdots \\ Y_{p-1}=A_1\lambda_1^{p-1}+A_2\lambda_2^{p-1}+\dots+A_p\lambda_p^{p-1} \end{cases}$$

- Dans ce système, les inconnus seront A_1, $A_2,\dots,A_p$, alors que $Y_0, Y_1,\dots, Y_{p-1}$ et $\lambda_1, \lambda_2,\dots,\lambda_p$ seront données.