14/09/23

SÉRIES TEMPORELLES

TD + TP
Logiciel R

→ examen écrit
  partie théorique
→ mémoire (= projet)
  partie application (on aura une BDD)
→ groupe de 3 ou 4 personnes
→ faire des prédictions
→ proposer une démarche cohérente de projection

Chapitre 5 important pour le mémoire surtout pour le début

Chapitre 1 Introduction aux séries temporelles définitions et exemples

dates mensuelles pas équiréparties car pas le m nb de jours par mois, il faut rapporter la suite de var aux nb de jours ouvrés pour avoir des données équiréparties NB var = variables ici

On n'est plus dans le cadre i i d

ex processus de Markov = processus à mémoire réduite
→ d'ordre 1 : on a besoin que de la dernière valeur connue pour prédire la loi de la prochaine

ex processus stochastique (cours de M2 1er semestre) ≠ avec ce cours c'est qu'on est en temps C° alors qu'ici on sera en temps discret

ex processus ponctuels → les dates ne sont pas forcément équiréparties et sont elles mêmes aléatoires
ici on fera que avec des dates déterministes

Σ temp = variable aléatoire
→ série temporelle
caractérisée par sa distribution (fonct de répartition,
{ si elle est ≥ 0 ⇒ tranformée de Laplace
{ si elle ∈ ℝ ⇒ fonction caractéristique)

variables aléatoires

ici on sera avec des $\overbrace{v.a}$ multivariée ex (X, Y)
sa loi : fonction de répartition, densité, fonct caractéristique
$\forall x, y$ $\quad \mathbb{P}(X \leq x, Y \leq y)$ $\quad f_{(X,Y)}(x,y)$

Ici on aura un vecteur de v.a infini et dénombrable

$(\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}^{\mathbb{Z}}$

$(X_t)_{t \in \mathbb{Z}}$

Comment caractériser la loi d'un vecteur de taille ∞ ?
car $\overset{\infty}{\cap}$ d'évènements $= \begin{cases} \Omega \\ \text{ou} \\ \emptyset \end{cases}$

donc fonction de répartition donnera toujours 0

On va passer par le thm de Kolmogorov
$\forall n \in \mathbb{N}^*$, $\forall t_1 \leq \ldots \leq t_n \in \mathbb{Z}^n$
on doit connaître la loi du vecteur $(X_{t_1}, \ldots, X_{t_n})$ pour connaître la loi du vecteur de taille ∞. On doit donc caractériser toutes les lois des vecteurs de v.a extraits du vecteur ∞.

$X_t(\omega) = x_t \leftarrow$ réalisation de la série temporelle
= trajectoire
Pour une suite de réalisations indicées, sa rep graphique sera une ligne brisée 6°
important pour le mémoire

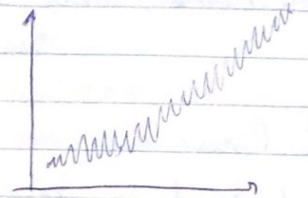
2 grandes natures
- de niveau ou d'état ou de stock
- de flux = variation du système

pas la même modélisation (stationnaire vs non stationnaire)

Ici on a besoin que d'1 seule réalisation pour avoir notre Σ temporelle
NB : on a plus l'hyp d' $\underbrace{\perp\!\!\!\perp}_{}$ = indépendance

Exemples

* Evolution concentrat° CO_2 Hawai

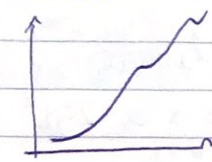


Comment dire ce qui va se passer en 2030 ?
∃ structure ⇝ croissance
comment est la croissance ?
quadratique, cubique, linéaire ?

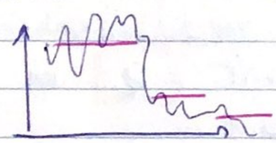
On assume ce qui se passera dans le futur en se basant sur le passé : on a envie de continuer la trend croissante

pente de la droite + saisonnalité = déterminer le futur

* Evolution PIB France (va d'état)



* Variat° du PIB (va du flux)

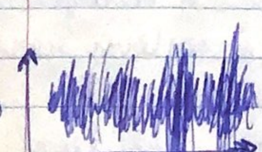


pas constant au cours du temps
composante de trend constante par morceaux "—"
↳ comment déterminer ces constantes ?

* Rentabilité log du CAC 40

rentabilité prix ≥ 0 $\quad \dfrac{P_t - P_{t-1}}{P_{t-1}} = \dfrac{P_t}{P_{t-1}} - 1 \in [-1, +\infty[$

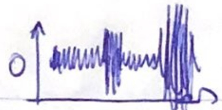
rentabilité log $\ln\left(\dfrac{P_t}{P_{t-1}}\right)$ $\quad$ m̂ ordre de grandeur quand petit
distribution + symétrique avec les gains et les pertes



→ pas de trend $\quad$ → pas saisonnière

{ niveau moyen nul
{ rentabilité moyenne nulle

variance rentabilité : aléatoire

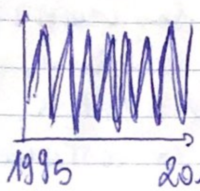
S&P 500 

variabilité augmente de la m̂ façon que CAC 40

on a une corréation à une date fixée probable entre les 2 marchés

↳ spatiale

* Températures moyennes journalières à Paris



forte saisonnalité (chaud en été, froid hiver)

trend ? évolution moyenne ?

difficile à lire pour interpréter

il faudrait effectuer un filtrage

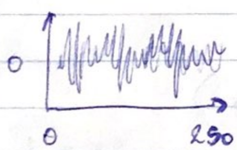
Sous R

rt (250, df = 10)

250 → nb de v.a

df = 10 → nb de dég de libertés

rt → random Student



Quelle valeur donner à la 251^e observation

$\boxed{251^e \text{ obs ?}}$ ici iid

donc ne dépend pas des 250^e observations précédentes

On va lui donner comme valeur la moyenne de la loi Student $\boxed{= 0}$ car symétrique

De plus, IE minimise la perte quadratique

1.2 Problèmes statistiques

Description écriture de la Σ temp + identification de la trend et de la saisonnalité + aléatoire

Prévision on connaît jusqu'à l'instant t, que va t'il se passer en t+1, t+2, ... ?

Pour ça il faut choisir un bon modèle (à définir)

Filtrage Σ temp pas assez satisfaisante (erreurs d'enregistrements, valeurs abherrantes) Il y aura forcément des "bouts" de la Σ à enlever car ils ne seront pas bons.

Traitement des données manquantes reconstruire valeurs manquantes à partir de leur passé et leur futur observé

Detection de rupture

Le niveau trend, l'intensité de la saisonnalité sont déterministes mais il peut y avoir des cassures et dans ce cas il faut les réajuster (pas étudié dans ce cours)

Causalité va multivariée : l'influence d'une composante sur une autre

1.3 Modelisation de la Σ temporelle

Plusieurs remarques ∃ plusieurs approches

⇒ ∃ plusieurs modèles (il peut avoir un modèle meilleur que les autres, à définir avec un critère (prédict° à 1 mois, 2 mois, ...))

Plusieurs composantes sous jacentes peuvent être déterministes mais il y aura toujours une composante aléatoire qui ne dépend pas du passé

À partir d'un nb d'obs finies on doit estimer une fonction ∞.

⇒ il faudra imposer des contraintes

Schéma décomposition additive en 3 termes

$$X_t = m_t + S_t + Y_t$$

m_t : trend (determ)

S_t : saisonnalité (determ)

Y_t : aléatoire mais aura une composante stable dans le temps

il faut lui donner une structure (à Y_t) pour modéliser sa loi

⇒ stable dans le temps

Rmq On suppose que cette décomposition existe, il est fort probable qu'il faille transformer la Σ temporelle pour qu'elle rentre dans notre cadre

Le choix de la transformation dépend de notre expérience

$$Y_t = h(Y_{t-1}, \dots, Y_{t-p}, \varepsilon_t, \dots, \varepsilon_{t-q})$$

modelisation de Y

Y_t est une fonction du passé

on utilise une information passée limitée

ε → c'est l'innovation
→ apporte une information qui n'appartient pas au passé
→ imprévisible
→ i i d

On va supposer que
Y_t est une CL des va passées et les innovations

1.4 Processus stationnaires

stationnaire fort
⇒ la loi est la $\underbrace{\hat{m}}_{= \text{même}}$ au cours du temps
⇒ invariante par translation en loi

Dans la définition h entier relatif, $h \in \mathbb{Z}$

Definition stationnaire $\begin{cases} \text{faible} \\ \text{au second ordre} \end{cases}$
↳ impose des conditions sur les moments d'ordre 1 et 2

Cela suppose que ces moments existent
(ex pour la loi de Cauchy ils n'existent pas)
↳ $\mathbb{E}[X_t^2] < +\infty \quad \forall t$

$\mu_X(t) = \mathbb{E}[X_t]$ variance $\mathbb{V}ar[X_r] = \gamma_X(r, r)$

covariance $\gamma_X(r, s) = \mathrm{Cov}(X_r, X_s)$

i) $t \mapsto \mu_X(t)$ ⫫ de t, $\forall t \in \mathbb{Z}$ $\quad \mu_X(t) = \mu$

ii) $t \mapsto \gamma_X(t, t+h)$ ⫫ de t, $\forall h$ $\quad \gamma_X(t, t+h) = \gamma_X(h)$

h est l'écart temporel entre les 2 v.a
[constante au cours du temps]

exemple :

processus stationnaire fort = v a i i d

faible = v.a iid si moments existent

fort ⇒ faible qd $\exists\ \mathbb{E}[X]$, $\mathbb{E}[X^2]$

faible : v a $\perp\!\!\!\perp$

Contre exemple
$X_{2k} \sim \mathcal{N}(0,1)$
$X_{2k+1} \sim \Gamma(0,1)$

$\mathbb{E}$ est nulle
covariance nulle car $\perp\!\!\!\perp$
⇒ faible

mais pas fort

Definition pour X_t stationnaire faible

$\forall h \in \mathbb{Z}$

autocovariance de (X_t) $\gamma_X(h) = \mathrm{Cov}(X_t, X_{t+h})$

autocorrélation de (X_t) $\rho_X(h) = \dfrac{\gamma_X(h)}{\gamma_X(0)}$

$= \mathrm{Cor}(X_t, X_{t+h})$

Propriétés

Rappel cov bilinéaire symétrique

$\gamma_X(-h) = \mathrm{cov}(X_t, X_{t-h})$

X_t stationnaire faible donc $\perp\!\!\!\perp$ de t

par symétrie de Cov

$= \mathrm{cov}(X_{t-h}, X_t) = \gamma_X(t-h, t) = \gamma_X(t-h+h, t+h)$

$= \mathrm{cov}(X_t, X_{t+h}) = \gamma_X(h)$

$\gamma_X(0) \geq 0$ car Var ≥ 0

$|\gamma_X(h)| = |\mathbb{E}[(X_t - \mathbb{E}[X_t])(X_{t+h} - \mathbb{E}[X_{t+h}])]|$

$\leq \left(\mathbb{E}[(X_t - \mathbb{E}[X_t])^2]\,\mathbb{E}[(X_{t+h} - \mathbb{E}[X_{t+h}])^2]\right)^{1/2}$

Hölder Cauchy

$\leq (\gamma_X(0)\gamma_X(0))^{1/2}$

$\leq \gamma_X(0)$

$$\left| \frac{\gamma_X(h)}{\gamma_X(0)} \right| = |\rho_X(h)| \leq 1$$

$\rho_X(h) = 1 \iff \exists$ relat° affine ↗

$= -1 \iff \exists$ relat° affine ↘

{ qd $\rho_X(h) = 0$ $\cancel{\Rightarrow}$ ⫫ } mais { ⫫ $\Rightarrow \rho_X(h) = 0$ }

Contre-exemple $Z \sim \mathcal{N}(0, 1)$

$X = Z$

$Y = Z^2$ pas ⫫

mais

$$\mathrm{cov}(X, Y) = \underbrace{\mathbb{E}[XY]}_{=0} - \underbrace{\mathbb{E}[X]}_{=0} \underbrace{\mathbb{E}[Y]}_{=1} = 0$$

car impair

$\rho_X(0) = 1$ car relat° linéaire parfaite entre moi et moi-m̂

$$X = \begin{pmatrix} X_{t_1} \\ | \\ X_{t_n} \end{pmatrix} \qquad t_1 \neq \dots \neq t_n \in \mathbb{Z}$$

$$\mathbb{E}[X] = \begin{pmatrix} \mathbb{E}[X_{t_1}] \\ | \\ \mathbb{E}[X_{t_n}] \end{pmatrix}_{(n,1)} = \begin{pmatrix} \mu_X(t_1) \\ | \\ \mu_X(t_n) \end{pmatrix}$$

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])(X' - \mathbb{E}[X'])]$$

$$= \left(\mathrm{cov}(X_{t_i}, X_{t_j}) \right)_{(n,n)}$$

$$= \left(\gamma_X(t_i, t_j) \right)_{(n,n)}$$

$\Rightarrow$ symétrique

$(*)\ a'\mathbb{V}[X]a = \big(\_\_\big)\underbrace{\big(\_\_\big)^{(1)}}$

$a \in \mathbb{R}^n$

$a'X = \sum a_i X_{t_i}$

on sait $\mathbb{V}[a'X] \geq 0$

et $\mathbb{V}[a'X] = a'\mathbb{V}[X]a$

$$(1) = \begin{pmatrix} \sum_{j=1}^{n} a_j \gamma_X(t_1, t_j) \\ \vdots \\ \sum_{j=1}^{n} a_j \gamma_X(t_n, t_j) \end{pmatrix}$$

$$(a_1 \ \cdots \ a_n)\,(1) = \sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j \gamma_X(t_i, t_j)$$

$$\overset{(*)}{=} \sum_{i,j} a_i a_j \underbrace{\gamma_X(t_i - t_j)}_{} \geq 0$$

$= Cov(X_t, X_{t+t_i-t_j})$

$= Cov(X_{t_j}, X_{t_i})$ avec $t = t_j$

$\Rightarrow$ definie positive

Ex avec $t_1 = 1 \ \cdots \ t_n = n$

$$\mathbb{V}[X] = \begin{pmatrix} \gamma_X(0) & \gamma_X(1) & \gamma_X(2) & \cdots & \gamma_X(n-1) \\ & \ddots & \ddots & \ddots & \gamma_X(2) \\ & & \ddots & \ddots & \gamma_X(1) \\ & & & & \gamma_X(0) \end{pmatrix} = V_n$$

matrice d'autocovariance

$$\rho[X] = \frac{R[X]}{\gamma_X(0)} = \begin{pmatrix} 1 & \rho_X(1) & \cdots & \rho_X(n-1) \\ & \ddots & \ddots & \vdots \\ & & \ddots & \rho_X(1) \\ & & & 1 \end{pmatrix} = R_n$$

matrice d'autocorrelat

$\det R_n \geq 0$

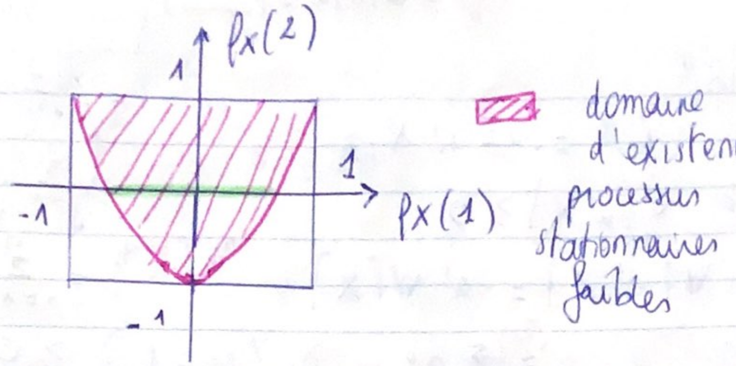
$$\det R_2 = \begin{vmatrix} 1 & \rho_X(1) \\ \rho_X(1) & 1 \end{vmatrix} = 1 - \rho_X^2(1) \geq 0 \quad \text{car } |\rho_X(h)| \leq 1$$

$$\det R_3 = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 \end{matrix} \\ \begin{vmatrix} 1 & \rho_X(1) & \rho_X(2) \\ \rho_X(1) & 1 & \rho_X(1) \\ \rho_X(2) & \rho_X(1) & 1 \end{vmatrix} \end{array} = \overbrace{(1 - \rho_X(2))}^{\geq 0} \times \underbrace{\left[1 + \rho_X(2) - 2\rho_X^2(1)\right]}_{\geq 0\ ?} \geq 0$$

$C_1 \leftarrow C_1 - C_3$

$$= \begin{vmatrix} 1 - \rho_X(2) & \rho_X(1) & \rho_X(2) \\ 0 & 1 & \rho_X(1) \\ \rho_X(2) - 1 & \rho_X(1) & 1 \end{vmatrix}$$

puis on développe selon la 1^{ere} colonne



$\rho_X(1)$ proche de $1 \Rightarrow \rho_X(2)$ proche de 1
―――――――― $-1 \Rightarrow$ ―――――――― 1
$\rho_X(2)$ nul $\Rightarrow \rho_X(1)$ ne pourra jamais être égal à 1 ou -1, il se balade là

Rmq

* fort + $\mathbb{E}[X^2] < +\infty \Rightarrow$ faible
$\cancel{\Leftarrow}$

* Σ temp gaussienne $\Leftrightarrow$ tous les sous vecteurs extraits sont gaussiens
$\Leftrightarrow$ pour tout sous vecteur fini, toute CL des composantes est gaussienne

$\Rightarrow$ fort = faible

Exemples de processus stationnaires

Définitions d'un $\underbrace{\text{bruit}}_{\text{parasite notre observation}}$ blanc :

- fort suite va iid centrées et de variance finie σ^2
  BBF $(0, \sigma^2)$

fort $\Rightarrow$ faible mais $\cancel{\Leftarrow}$

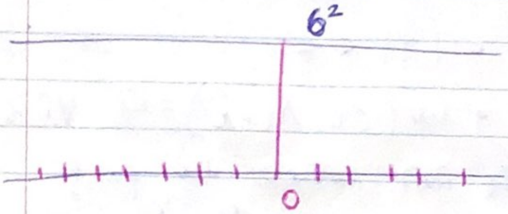
- faible suite va centrées, de variances finies et constantes et non corrélées

$\mathbb{E}[\varepsilon_t] = 0$, $\mathbb{V}[\varepsilon_t] = \sigma^2$, $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t'}) = 0 \ \forall t \neq t'$

BB $(0, \sigma^2)$ autocovariance $\gamma_\varepsilon(h) = \begin{cases} \sigma^2 & h = 0 \\ 0 & h \neq 0 \end{cases}$

$\gamma_\varepsilon(\pm 1) = 0$

$\Rightarrow \gamma_\varepsilon(\pm n) = 0$ (pour $n \neq 0$)



$$\rho_\varepsilon(h) = \begin{cases} 1 & \text{qd } h = 0 \\ 0 & \text{qd } h = \pm n \end{cases}$$

noté MA(1)

Definition d'un processus moyenne mobile d'ordre 1

$\varepsilon_t \sim BB(0, 6^2)$

$X_t = m + \varepsilon_t + \theta \varepsilon_{t-1}$, $\theta \in \mathbb{R}$, $m \in \mathbb{R}$ — X_t est une Σ temp

- $\mathbb{E}[X_t] = m$ ne depend pas du tps ✓
- $\mathbb{V}[X_t] = (1+\theta^2)6^2$ ne depend pas de t ✓

  $\mathbb{V}(a'X)$
- $cov(X_t, X_{t+1}) = cov(m + \varepsilon_t + \theta\varepsilon_{t-1}, m + \varepsilon_{t+1} + \theta \varepsilon_t)$

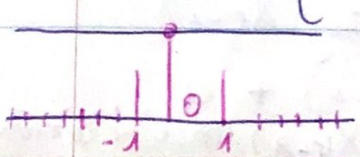
  $= \theta 6^2$ ne depend pas de t ✓
- $cov(X_t, X_{t+2}) = cov(m + \varepsilon_t + \theta\varepsilon_{t-1}, m + \varepsilon_{t+2} + \theta\varepsilon_{t+1})$

  $= 0$

$\forall t, \forall x > 1$ $cov(X_t, X_{t+x}) = 0$ ne dépend pas de t ✓

NB: X_t est bien stationnaire faible

d'où $$\gamma_X(h) = \begin{cases} (1+\theta)6^2 & \text{pour } h = 0 \\ \theta 6^2 & \text{pour } h = \pm 1 \\ 0 & \text{pour } h = \pm n \text{ pour } n \neq \{-1, 0, 1\} \end{cases}$$



$|\theta| \leq 1$ $$\rho_X(h) = \begin{cases} 1 & h = 0 \\ \frac{\theta}{1+\theta^2} & h = \pm 1 \\ 0 & h = \pm n \end{cases}$$

noté AR(1)

Definition d'un processus autorégressif d'ordre 1

$(\varepsilon_t)_{t\in\mathbb{Z}} \sim BB(0, \sigma^2)$

$X_t = \varphi X_{t-1} + \varepsilon_t$

- $|\varphi| < 1$
- $\operatorname{cov}(\varepsilon_t, X_{t-i}) = 0 \quad \forall i \geq 1$

ε_t non corrélé au passé de X_t

- (X_t) est stationnaire faible

$\star$ $\mathbb{E}[X_t] = \varphi\,\mathbb{E}[X_{t-1}] = \varphi\,\mathbb{E}[X_t]$ car faible

$\Leftrightarrow (1-\varphi)\,\mathbb{E}[X_t] = 0$

$\Leftrightarrow \underline{\mathbb{E}[X_t] = 0}$ car $|\varphi| < 1$ $\quad$ $\mathbb{E}$ ne dépend pas de t

$\star$ $\mathbb{V}[X_t] = \mathbb{V}[\varphi X_{t-1} + \varepsilon_t]$

$= \varphi^2\,\mathbb{V}[X_{t-1}] + \underbrace{\mathbb{V}[\varepsilon_t]}_{=\sigma^2} + 2\operatorname{cov}(\varphi X_{t-1}, \varepsilon_t)$

$= \varphi^2\,\mathbb{V}[X_{t-1}] + \sigma^2 + 2\varphi \times 0$ $\quad$ par hyp

$= \varphi^2\,\mathbb{V}[X_t] + \sigma^2$

car faible

$\Leftrightarrow (1-\varphi^2)\,\mathbb{V}[X_t] = \sigma^2 \underset{|\varphi|<1 \Rightarrow \varphi^2<1}{\Leftrightarrow} \underline{\mathbb{V}[X_t] = \dfrac{\sigma^2}{1-\varphi^2}}$

ne dépend pas de t

$\operatorname{cov}(X_t, X_{t-1}) = \operatorname{cov}(\varphi X_{t-1} + \varepsilon_t, X_{t-1})$

$= \varphi\,\mathbb{V}[X_{t-1}] + \underbrace{\operatorname{cov}(\varepsilon_t, X_{t-1})}_{=0}$

$= \varphi \dfrac{\sigma^2}{1-\varphi^2}$

$\operatorname{cov}(X_t, X_{t-2}) = \operatorname{cov}(\varphi X_{t-1} + \varepsilon_t, X_{t-2})$

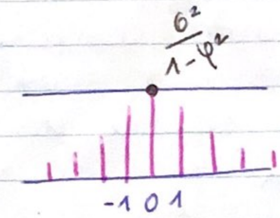
$= \varphi \underbrace{\operatorname{cov}(X_{t-1}, X_{t-2})}_{=\operatorname{cov}(X_t, X_{t-1})} + 0$ $\quad$ NB : c'est le même calcul

$$\Rightarrow \operatorname{cov}(X_t, X_{t-2}) = \varphi^2 \frac{\sigma^2}{1-\varphi^2}$$

$$\operatorname{cov}(X_t, X_{t-k}) = \varphi^k \frac{\sigma^2}{1-\varphi^2}$$ ne dépend pas de t

$k \geq 1$

$$\gamma_X(h) = \begin{cases} \dfrac{\sigma^2}{1-\varphi^2} & \text{qd } h = 0 \\ \varphi^{|h|} \dfrac{\sigma^2}{1-\varphi^2} & \end{cases}$$



Rmq
$$\begin{aligned} X_t &= \varphi X_{t-1} + \varepsilon_t \\ &= \varphi(\varphi X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \varphi^2 X_{t-2} + \varphi \varepsilon_{t-1} + \varepsilon_t \\ &\overset{\dots}{=} \varphi^k X_{t-k} + \sum_{i=0}^{k-1} \varphi^i \varepsilon_{t-i} \end{aligned}$$

qd $k \to +\infty$

$$X_t = \sum_{i=0}^{\infty} \varphi^i \varepsilon_{t-i}$$

mais toute cette remarque manque de précision et de définition mathématique

et
$$\begin{cases} \mathbb{E}[X_t] = 0 \\ \mathbb{V}[X_t] = \sum_{i=0}^{\infty} \varphi^{2i} \sigma^2 = \dfrac{\sigma^2}{1-\varphi^2} \end{cases}$$

Définition d'une marche aléatoire

$X_0 = 0$ condition initiale

$$X_t = X_{t-1} + \varepsilon_t = \sum_{i=1}^{t} \varepsilon_i$$

$\mathbb{E}[X_t] = 0$

$\mathbb{V}[X_t] = t\sigma^2$ dépend de t $\rightarrow$ pas stationnaire

$\operatorname{cov}(X_t, X_{t-1}) = (t-1)^2 \sigma^2$ dépend de t

⚠ pour AR(1) $X_t = \varphi X_{t-1} + \varepsilon_t$ et $|\varphi| < 1$

ici $\varphi = 1$ donc processus non stationnaire

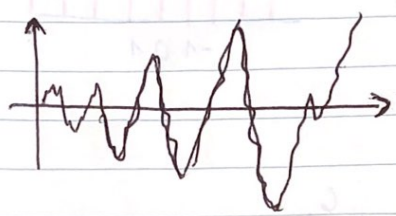
21/09/23

Séries temporelles

2 exemples à repérer pour la non stationnarité
↳ on donne une condition initiale et on regarde où le processus va

*** Marche aléatoire**

la variance est linéaire en t



il croise une ∞ de fois 0

il balaye tout l'espace

Comment le "stationnariser" = rendre stationnaire ce processus ? Noté $(X_t)_{t \in \mathbb{N}}$

on fait la différence (vu en TD 1)

$$Y_t = \Delta X_t = X_t - X_{t-1} = \varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$$

*** Modèle avec tendance et variance fonct° du temps**

Avec la tendance et la variance qui sont fonction du temps

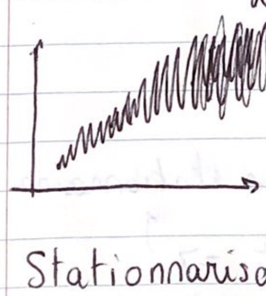
On est sur du non stationnaire déterministe

on crée un processus hétéroscédastique

$$X_t = \underbrace{m(t) + k(t)}_{\text{déterministes}} \varepsilon_t$$

$$\mathbb{E}[X_t] = m(t)$$

$$\mathbb{V}[X_t] = k^2(t)\,\sigma_\varepsilon^2$$



commentaires

double structure : + et ×

et variance ↗ au cours du temps

Stationnarisation de $(X_t)_{t \in \mathbb{N}}$

On centre réduit

$$Y_t = \frac{X_t - m(t)}{k(t)} = \varepsilon_t$$

on identifie $m(t)$ dans les données puis on le retranche à X_t
De même pour $k(t)$

Fin du chapitre 1, à savoir
* stationnarité faible, * Bruit Blanc Fort et faible
* identifier 2 formes de non stationnarité

Chapitre 2 Trend et saisonnalité

<u>Hyp</u> $X_t = m_t + S_t + Y_t$ dans ce chapitre

m_t : trend déterministe

S_t : saisonnalité déterministe périodique
$S_t = S_{t+d}$ $\forall t$
avec $d \in \mathbb{N}$ connu

Y_t : bruit (Blanc) aléatoire
$\mathbb{E}[Y_t] = 0$
$\mathbb{V}[Y_t] = \sigma_Y^2$
$cov(Y_t, Y_{t'}) = 0$ $t \neq t'$

on a donc d valeurs à trouver pour connaître totalement S_t

On a 1 observation de notre série temporelles avec un nb fini d'observations temporelles
à une réalisation de notre série

On veut estimer par moindres carrés, sans savoir les lois mais en ayant des informations sur les moments, 2 composantes m_t et S_t
(on trouve le bruit en retranchant m_t et S_t à l'observation)

méthodes de régression

régression globale vs régression locale

régression globale : 1 régression sur les données

régression locale : plusieurs modèles de régression sur plusieurs parties de l'échantillon
ça permet d'avoir une erreur plus faible mais

on ne pourra pas prédire à long terme, seulement à court terme / au voisinage de la dernière observation étudiée

méthodes de filtrage

par des opérateurs (comme en TD) pour se débarasser de la trend ou la saisonnalité et trouver l'autre composante

2. 1 Régression Buys Ballot

de période l dorénavant

décomposition additive

$$X_t = \underbrace{m_t}_{\sum_{i=1}^{k} b_i m_t^{(i)}} + \underbrace{S_t}_{l \text{ valeurs à trouver}} + Y_t$$

$$\sum_{j=1}^{l} c_j S_t^{(j)}$$

b_i : inconnu, paramètre à estimer ; $m_t^{(i)}$: fonctions connues

c_j : paramètre à estimer ; $S_t^{(j)}$: fonctions périodiques (l) connues

Exemple

m_t = polynôme quadratique

S_t = périodique de période 4

$$\underset{(T,1)}{X} = \begin{pmatrix} x_1 \\ | \\ x_T \end{pmatrix} = \begin{array}{c} \begin{array}{ccc} m^{(1)} & m^{(2)} & m^{(3)} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & \\ | & 2 & 4 & \\ | & 3 & 9 & \dots \\ | & | & | & \\ 1 & T & T^2 & \end{pmatrix} \end{array} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$+ \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ | & | & | & 0 \\ & & & 1 \end{pmatrix} \end{array} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} + \begin{pmatrix} Y_1 \\ | \\ Y_T \end{pmatrix}$$

On va noter

$$X_{(T,1)} = M_{(T,k)} \cdot b_{(k,1)} + S_{(T,l)}\, c_{(l,1)} + Y_{(T,1)}$$

$$X_{(T,1)} = \underset{(T,\,k+l)}{D}\ \underset{(k+l,\,1)}{\theta} + Y_{(T,1)} \quad \text{avec } D = (M, S) \quad \theta = \begin{pmatrix} b \\ c \end{pmatrix}$$

on met les paramètres à estimer dans le même vecteur θ

$$\mathbb{E}[Y] = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_{(T,1)} \qquad \mathbb{V}[Y] = \sigma_Y^2\, I_{d\,(T,T)}$$

On a X et D on cherche θ.

Est-ce que le modèle est identifiable ?

A partir des observations, on trouve un θ unique pour nos observations X

Avec $2\ \theta \neq$ on obtient 2 modèles $\neq$

$$\begin{cases} \theta_1 \neq \theta_2 \Rightarrow \mathbb{E}_{\theta_1}[X] \neq \mathbb{E}_{\theta_2}[X] \\ \qquad\qquad D\theta_1 \neq D\theta_2 \end{cases}$$

$\Leftrightarrow$ Par contraposée

$$\{ D\theta_1 = D\theta_2 \Rightarrow \theta_1 = \theta_2$$

$$\Leftrightarrow \quad D(\theta_1 - \theta_2) = 0 \Rightarrow \theta_1 - \theta_2 = 0 \Leftrightarrow \ker D = 0$$

$$\Leftrightarrow \operatorname{rg}(D) = k + l$$

$$\Leftrightarrow D'D \text{ est inversible}$$

On a écrit la condition d'identifiabilité

on a la somme des colonnes de S = 1^{re} colonne de M

donc D n'est pas de plein rang

donc le modèle n'est pas identifiable

D'où la contrainte d'identifiabilité

$$\sum_{j=1}^{l} c_j = 0 \Rightarrow c_l = -\sum_{j=1}^{l-1} c_j$$

pas d'effet de saisonnalité ?

un effet moyen à cycle nul

d'où $S_t = \sum_{j=1}^{l} c_j S_t^{(j)}$

$= \sum_{j=1}^{l-1} c_j \underbrace{\left(S_t^{(j)} - S_t^{(k)}\right)}_{= \tilde{S}_t^{(j)}}$

on réduit d'une dimension

D'où

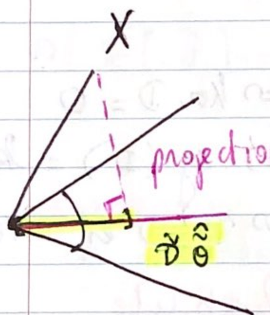
$$\tilde{S}_{(T,\, l-1)} = \begin{array}{c} \begin{array}{ccc} \tilde{S}_1 & \tilde{S}_2 & \tilde{S}_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \\ 1 & 0 & 0 \\ \vdots & 1 & 0 \\ & \vdots & \vdots \end{pmatrix} \end{array}$$

$$c_{(l-1,\, 1)} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

$$X = \tilde{D}\,\tilde{\theta} + Y \qquad \tilde{\theta} = \begin{pmatrix} b \\ c_1 \\ \vdots \\ c_{l-1} \end{pmatrix} \qquad \tilde{D} = (M, \tilde{S})$$

But réécrire pour rentrer dans un cadre classique de minimisation par moindres carrés

$\mathbb{R}^T$ = espace d'observation



projection orthogonale de X sur $\mathcal{L}(\tilde{D})$

$\mathcal{L}(\tilde{D})$ = e v généré par les colonnes de $\tilde{D}$

$= \mathbb{R}^{k+l-1}$

on cherche la C.L qui minimise la distance entre X_t et la CL des fonctions m_t et S_t

donc $\hat{\tilde{\theta}} = \underset{\mathbb{R}^{k+l-1}}{\operatorname{argmin}} \| X - \tilde{D}\tilde{\theta} \|^2 = MCO$

$$MCO \iff \underset{(k+l-1,\,T)}{\tilde{D}'}\underbrace{(X - \tilde{D}\hat{\tilde{\theta}})}_{(T,\,1)} = 0$$

condition d'orthogonalité

$$\text{D'où } \hat{\tilde{\theta}} = \underbrace{(\tilde{D}'\tilde{D})^{-1}}_{\text{inversible}} \tilde{D}'X$$

⟹ cela assure l'unicité et la bonne définition de $\hat{\tilde{\theta}}$

⟹ identifiabilité

Propriétés $\hat{\tilde{\theta}}$ BLUE

sans biais et de variance minimale parmi les estimateurs sans biais

$$\hat{\tilde{\theta}} = (\tilde{D}'\tilde{D})^{-1}\tilde{D}'(\tilde{D}\tilde{\theta} + Y)$$

$$= (\tilde{D}'\tilde{D})^{-1}\tilde{D}'\tilde{D}\tilde{\theta} + (\tilde{D}'\tilde{D})^{-1}\tilde{D}'Y$$

$$= \tilde{\theta} + (\tilde{D}'\tilde{D})^{-1}\tilde{D}'Y$$

$$\Rightarrow \mathbb{E}[\hat{\tilde{\theta}}] = \tilde{\theta} + \underbrace{(\tilde{D}'\tilde{D})^{-1}\tilde{D}'}_{\text{déterministe}}\underbrace{\mathbb{E}[Y]}_{=0 \text{ car bruit}}$$

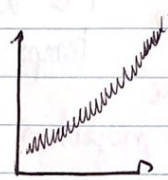
Avantages

on projette au delà de T

sensible par rapport à m_t

2.3 exemples

$\underline{\underline{CO_2}}$



trend linéaire croissant

Série 1 nos observations

Série 2 fonct° quadratique

Série 3 fonction périodique

Série 4 bruit

Série 4 = Série 1 - Série 2 - Série 3

Est que la série 4 est bien du bruit

$\mathbb{E} = 0$ ok

$\mathbb{V}$ constante ok

Bruit Blanc ? NON

on a pas l'indépendence

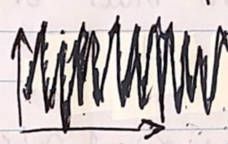
on a une autocorrélation forte

qd ≥ 0, le terme d'après a bcp de chances d'être ≥ 0

tandis que en $\perp\!\!\!\perp$, on s'attend à $X_n \geq 0 \Rightarrow X_{n+1} \geq 0$ ou ≤ 0 avec autant de probas

températures à Paris

On suppose trend linéaire décroissant

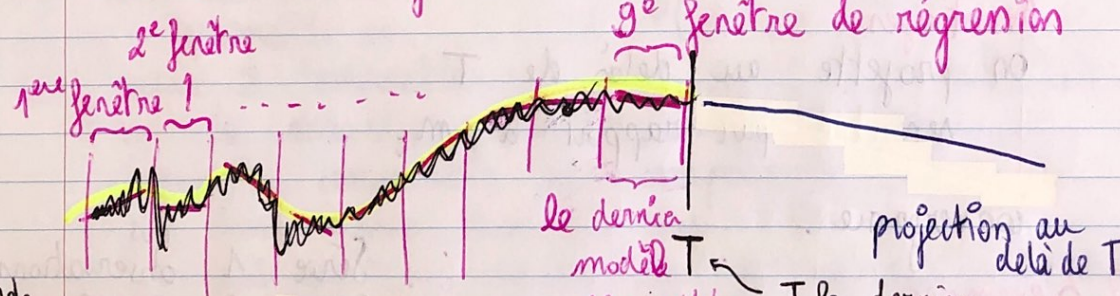


on alterne beaucoup plus fréquemment dans le bruit

it est plus blanc que l'exemple du CO_2

mais pour la projection on va pas utiliser ce modèle, d'où l'idée de faire des régressions locales

Méthode STL (régression locale)



Légende

régressions

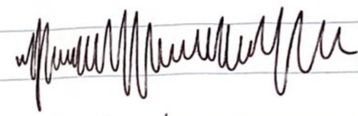
estimation (lissée)

avec STL

pour CO_2

trend

Bruit

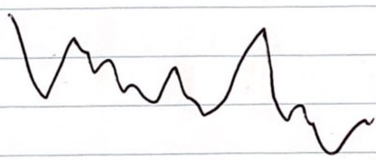


\+ Blanc

constantes de bruit ≥ 0 se suivent mais pas autant $\Rightarrow$ on a + d'alternance

pour températures à Paris

trend



? projection difficile à imaginer

Bruit bcp d'alternance donc bien blanc

2.2 Méthodes par moyenne mobile (filtrage)

ici filtre linéaire $\varphi : \begin{array}{l} \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}} \\ (X_t) \mapsto \varphi(X_t) \end{array}$

NB: j'utilise Σ comme raccourci pour dire « série ». Ce n'est pas une notation du prof.

φ est une transformation d'une Σ temporelle en une autre Σ temporelle

$$\varphi(\lambda X_t + \mu Y_t) = \lambda \varphi(X_t) + \mu \varphi(Y_t)$$

Pour la structure additive, on peut donc travailler composante par composante

On va devoir donc trouver les filtres en sachant qu'on cherche un filtre qui va annuler / effacer une composante pour étudier l'autre

Estimation de la saisonnalité

$$\varphi(X_t) = \varphi(m_t) + \varphi(S_t) + \varphi(Y_t)$$
$$= 0 \quad + \underbrace{S_t}_{\text{invariant}} + \underbrace{\quad}_{\text{on réduit à } 0}$$

φ 1er filtre

Estimation du trend

$$\Psi(X_t) = \Psi(m_t) + \Psi(S_t) + \Psi(Y_t)$$

$$= \underbrace{m_t}_{\text{invariant}} + 0 + \overset{\approx}{0}$$

Ψ est un 2^e filtre

NB : il est impossible d'annuler le bruit
il sera réduit à 0 / diminué le plus possible.

11/10/23

Séries temporelles

Slides 41/285 → 68/285

Rappel

$X_t = m_t + S_t + Y_t$

m_t : trend, déterministe

S_t : saisonnalité, déterministe, $\exists d \in \mathbb{N}^*$, $S_{t+d} = S_t$

Y_t : bruit, aléatoire, $\mathbb{E}[Y_t] = 0$, $\mathbb{V}[Y_t] = \sigma_Y^2$, $\mathrm{cov}(Y_t, Y_{t'}) = 0 \quad t \neq t'$

2.2 Méthodes par moyenne mobiles ou de filtrage

$\varphi : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ opérateur linéaire

↳ $\varphi(\lambda X_t + \mu Y_t) = \lambda \varphi(X_t) + \mu \varphi(Y_t)$

Objectifs :

→ Estimation du trend

$$\varphi(X_t) = \varphi(m_t) + \varphi(S_t) + \varphi(Y_t)$$
$$= m_t + 0 + \approx 0$$

élimine saisonnalité ; réduit fortement (impossible éliminer l'aléatoire)

→ Estimation de la saisonnalité

$$\varphi(X_t) = 0 + S_t + \approx 0$$

M comme moyenne mobile $M : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$

$$M X_t = \sum_{i=-m_1}^{m_2} \theta_i X_{t+i}$$

centrée = autant de coefficients dans le passé que dans le futur

+ symétrique avec autant de poids pour le passé que le futur (pondération uniforme)

$\mathrm{Ker}\, M = \{(X_t)_{t \in \mathbb{Z}} \mid M X_t = 0\}$ espace vectoriel sous espace propre associé à la valeur propre 0 les invariants

$$\mathrm{Inv}(M) = \{ (X_t)_{t\in\mathbb{Z}} , MX_t = X_t \}$$ ———— 1

$$= \mathrm{Ker}(\underbrace{M - Id}_{\text{aussi une moyenne mobile}})$$

Comment trouver le noyau d'une moyenne mobile ?

$$\mathrm{Ker}\, M = \{ (X_t)_t , \underbrace{\sum_{i=-m_1}^{m_2} \theta_i X_{t+i} = 0}_{} \quad \forall t \in \mathbb{Z} \}$$

$$\theta_{-m_1} X_{t-m_1} + \dots + \theta_{m_2} X_{t+m_2} = 0$$

$$\Leftrightarrow \alpha_0 z_n + \alpha_1 z_{n+1} + \dots + \alpha_p z_{n+p} = 0 \quad \forall n$$

Rappels

$p = 1$ $\quad \alpha_0 z_n + \alpha_1 z_{n+1} = 0 \Leftrightarrow z_{n+1} = \dfrac{\alpha_0}{\alpha_1} z_n$

$$= \beta z_n = \beta^{n+1} z_0$$

$p = 2$ $\quad \alpha_0 z_n + \alpha_1 z_{n+1} + \alpha_2 z_{n+2} = 0$

Polynôme associé $\quad P(z) = \alpha_0 + \alpha_1 z \quad p = 1$

$P(\beta) = 0$ pour $p = 1$

$P(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 \quad p = 2$

$\Delta > 0 \quad \beta_1, \beta_2$ racines réelles

$$\alpha_0 z^n + \alpha_1 z^{n+1} + \alpha_2 z^{n+2} = z^n(\alpha_0 + \alpha_1 z + \alpha_2 z^2)$$

$z_n = c_1 \beta_1^n + c_2 \beta_2^n \quad c_1, c_2 \in \mathbb{R}$ avec les conditions initiales pour trouver c_1 et c_2

espace des solutions est de dimension 2.

$z_0 = c_1 + c_2$

$z_1 = c_1 \beta_1 + c_2 \beta_2$

$\Delta = 0 \quad \beta$ racine double $\quad z_n = (c_1 + c_2 n)\beta^n, \ c_1, c_2 \in \mathbb{R}$

multiplicité 1

$\Delta < 0$ β_1, β_2 racines complexes conjuguées
$\beta_2 = \overline{\beta_1} = \overline{\beta}$

$\beta = \rho e^{i\theta}$ avec $\rho > 0$ $\theta \in [0, 2\pi[$

et $\overline{\beta} = \rho e^{-i\theta}$

on prend c et $\bar{c}$ pour avoir l'espace des solutions dans $\mathbb{R}$!

$z_n = c\beta^n + \bar{c}\,\overline{\beta}^n$ $c \in \mathbb{C}$ $c = \mu e^{i\omega}$

$= 2\,\mathrm{Re}(c\beta^n) = 2\mu\rho^n \cos(n\theta + \omega)$

$= \rho^n(A\cos n\theta + B\sin n\theta)$ $A, B \in \mathbb{R}$

cos a+b
cos a cos b + sin a sin b

multiplicité $p > 1$ z_n = polynôme deg $p-1$ $\times \rho^n(A\cos n\theta + B\sin n\theta)$

Invariant

Prop 1) Si $\sum_{i=-m_1}^{m_2} \theta_i = 1 \Rightarrow (a) \in \mathrm{Inv}\, M$ a série temporelle

NB $(X_t = a \;\forall t)$

$Ma = \left(\sum_{i=-m_1}^{m_2} \theta_i\right) a = a$

Dans l'invariant on a les constantes

les constantes sont conservées

2) (Si M symétrique et $\underline{\sum \theta_i = 1}$) $\Leftrightarrow$ $m_1 = m_2$, $\theta_1 = \theta_{-1}$

$=$ M conserve les constantes

$\Leftrightarrow$ les droites sont conservées dans l'invariant

c'est-à-dire $\Rightarrow$ $Mt = t$

on considère la série temporelle (t)
$(X_t = t \;\forall t)$

d'où $MX_t = \sum \theta_i X_{t+i}$

$= \sum \theta_i (t+i)$

$Mt = \sum_{i=-m}^{m} \theta_i (t+i)$

$= t + \sum_{i=-m}^{m} i\theta_i$

on a $-i\theta_{-i}$
$= -(i\theta_{-i})$
$= -(i\theta_i)$

impair donc la $\Sigma = 0$

Donc $Mt = t$

BBF après passage du filtre

on veut $\frac{\sigma^2_{M\varepsilon}}{\sigma^2_\varepsilon}$ le plus petit possible càd $\sum_{j=-m_1}^{m_2} \theta_j^2$

quelle est la moyenne mobile qui
- **conserve les constantes**
- **conserve les droites**
- **réduit le + possible $\frac{\sigma^2_{M\varepsilon}}{\sigma^2_\varepsilon}$**

celle qui satisfait pb d'optimisation suivant

$$\min \sum_{j=-m}^{m} \theta_j^2$$

sc

$$\sum_{j=-m}^{m} \theta_j = 1$$

c'est $\theta_j = \frac{1}{2m+1} \quad \forall j \in [\![1, m]\!]$

La moyenne mobile arithmétique MMA

$$MX_t = \frac{1}{2m+1} \sum_{i=-m}^{m} X_{t+i} \qquad \text{Ker } M ?$$

$$\frac{1}{2m+1} X_{t-m} + \cdots + \frac{1}{2m+1} X_{t+m} = 0$$

$$P(z) = \frac{1}{2m+1} \left(1 + z + \cdots + z^{2m} \right)$$

$$= \frac{1}{2m+1} \frac{1-z^{2m+1}}{1-z} \quad \left(= \frac{Q(z)}{R(z)} \right)$$

$$Q(z) = 0 \iff z^{2m+1} = 1 \iff z = e^{i\frac{2\pi j}{2m+1}}, \; j \in [\![0, 2m]\!]$$

d'où $P(z) = 0 \iff j \in [\![1, 2m]\!]$

0 et 2m sont conjuguées

on aura m racines

et m racines conjuguées

$\frac{2\pi}{2m+1}$

$-\frac{2\pi}{2m+1}$

donc $z_n = \sum_{j=1}^{m} \left(\alpha_j \cos\left(\frac{2\pi j}{2m+1} n\right) + \beta_j \sin\left(\frac{2\pi j}{2m+1} n\right)\right)$

$\Rightarrow$ $\forall$ les fonctions périodiques de période $2m+1$

MMA élimine les fonctions périodiques de période impaire.

Comment éliminer celles de période paire ?

avec une moyenne de moyenne mobile arithmétique qui conserve les droites et annule les fonctions périodiques de période paire.

↳ devient symétrique

Lissage

trend polynôme 2 ok

si° > deg 2 on veut lisser le plus possible.

4) output produit pour le trend soit le - erratique

5) simplicité des coefficients = nb de calculs

+ m est grand, + on perd de l'information au début et à $\underbrace{\text{la fin}}$

gênant car on veut faire de la prédiction

slide 53/285

En pratique

construction successive moyenne mobile pour éliminer le trend et la saisonnalité

$\hat{m}_t$ estimateur après le passage de MM

$= M_0 X_t$ (m_t + ≈ bruit)

$Id - M_0$ appliquée à M_1

puis à X_t

pour retrouver la constante

Exemple

nb termes θ_2 θ_1

ordre 4 $\Rightarrow m = 2$ $\quad \downarrow \quad$ θ_{-2} θ_{-1} θ_0

choix naturel $\left\{ [5] ; \frac{1}{8} [1, 2, 2] \right\}$

$$\frac{1}{4} \left(0{,}5 X_{t-2} + X_{t-1} + X_t + X_{t+1} + 0{,}5 X_{t+2} \right)$$

$$= \frac{1}{4} \, 0{,}5 \left(X_{t-2} + 2 X_{t-1} + 2 X_t + 2 X_{t+1} + X_{t+2} \right)$$

$$= \frac{1}{8} \left(\text{———} \right)$$

MM de Henderson (la meilleure pour le lissage) slides 69/285 à 115/285

12/10/23 Série temporelle

Chapitre 3 Processus Linéaires

Objectif : autre caractérisation des séries temporelles

$\mathcal{L}^2$ = va de carré intégrable

$= \{ X : \mathbb{E}[X^2] < +\infty \}$

3.1 L'espace $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$: définition et propriétés

Espace vectoriel $X, Y \in \mathcal{L}^2$

$$\mathbb{E}[(X+Y)^2] = \underbrace{\mathbb{E}[X^2]}_{<+\infty} + \underbrace{\mathbb{E}[Y^2]}_{<+\infty} + 2\underbrace{\mathbb{E}[XY]}_{<+\infty}$$

$$|\mathbb{E}[XY]| \leq \mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X]\,\mathbb{E}[Y]} < +\infty$$

Norme $\|X\| = \mathbb{E}[X^2]^{1/2}$

1) $\|X\| = 0 \Rightarrow X = 0$ ps
2) $\|\lambda X\| = |\lambda| \, \|X\|$
3) $\|X+Y\|^2 \leq \|X\|^2 + \|Y\|^2 + 2\sqrt{\|X\|^2 \|Y\|^2}$

$\leq (\|X\| + \|Y\|)^2$ inégalité triangulaire

On a donc un espace vectoriel normé

$d(X, Y) = \|X - Y\|$ distance

$\langle X | Y \rangle = \mathbb{E}[XY]$ produit scalaire

i) bilinéaire symétrique ok
ii) définie positive

orthogonal à

$X \perp Y$ si $\mathbb{E}[XY] = 0$

$\mathcal{L}^2$ est donc un espace vectoriel normé complet (Banach)

$X_n \xrightarrow{L^2} X$ si $\|X_n - X\| \underset{n\to+\infty}{\longrightarrow} 0$ ($\Rightarrow X \in L^2$)

déf de la convergence L^2

$X_n = \sum_{i=1}^{n} Y_i$, $Y_i \in L^2$ quelle est la limite de X_n ?

Toute série absolument convergente est convergente

$$\sum_{i=1}^{\infty} \|Y_i\| < +\infty \Rightarrow \sum_{i=1}^{n} Y_i \in \mathcal{L}^2$$

$Y = \sum_{j \in \mathbb{Z}} a_j X_j$ bien définie si $\sum_{j \in \mathbb{Z}} \| a_j X_j \| < +\infty$

avec $(a_j)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ $= \sum_{j \in \mathbb{Z}} |a_j| \, \| X_j \|$

Rmq $(X_t)_{t \in \mathbb{Z}}$ stationnaire (faible)

$$\| X_j \| = \sqrt{\mathbb{E}[X_j^2]} = \sqrt{\mathbb{V}[X_j] + \mathbb{E}[X_j]^2}$$

$$= \| X \|$$

Y existe si $\sum_{j \in \mathbb{Z}} |a_j| < \infty$

Définition moyenne mobile ∞

$$(X_t)_{t \in \mathbb{Z}} \xrightarrow{(a_j)_{j \in \mathbb{Z}} \,/\, \sum_{j \in \mathbb{Z}} |a_j| < +\infty} = (Y_t)_{t \in \mathbb{Z}}$$

$$Y_t = \sum_{j \in \mathbb{Z}} a_j X_{t-j}$$

en t on utilise ce qui se passe dans le futur (j négatif) $X_{t+1}, X_{t+2} \dots$
or on a que l'information dans le passé !

$(X_t)_{t \in \mathbb{Z}}$ stationnaire faible $\mathbb{E}[X_t] = \mu_X$

$\mathrm{Cov}(X_t, X_{t-h}) = \gamma_X(h)$

I T+ Fubini

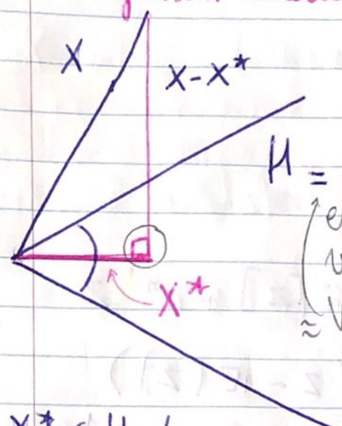
$$\mathbb{E}[|Y_t|] \leq \sum_{j \in \mathbb{Z}} |a_j| \underbrace{\mathbb{E}[|X_{t-j}|]}_{\text{ne dépend pas du temps}} < +\infty \text{ par hypothèse sur les } a_j$$

on peut donc permuter $\mathbb{E}$ et $\sum_{j \in \mathbb{Z}}$ par Fubini

$$\mathbb{E}[Y_t] = \sum_{j \in \mathbb{Z}} a_j \mathbb{E}[X_{t-j}] = \left(\sum_{j \in \mathbb{Z}} a_j \right) \mu_X$$

$$\mathrm{Cov}(Y_t, Y_{t-h}) = \mathrm{Cov}\left(\sum a_j X_{t-j}, \sum a_i X_{t-h-i}\right)$$
$$\underset{\text{Fubini}}{=} \sum_j \sum_i a_j a_i \gamma_X(h+i-j)$$

3.2 Régression linéaire ou affine théorique sur un nombre fini de retards



Sev fermé $\mathcal{L}^2$

projection d'une va sur L^2

$H = ev(Z_1, Z_2 \dots)$ (espace vectoriel $\approx$ Vect)

$X \in \mathcal{L}^2$

on veut la meilleure approximation linéaire de X dans H

on veut l'élément de H le + proche de X

$\exists!\, X^* \in H \,/$

$\min_{Y \in H} \|X - Y\| = \|X - X^*\|$

$Y \in \sum a_j Z_j$

avec $X^* = \sum a_j^* Z_j \,/\, \forall j \; (X - X^*) \perp Z_j$

$\iff \mathbb{E}[(X - X^*)Z_j] = 0$

On va trouver les a_j^*

$H = \mathrm{Vect}(Z_0, Z_1, \dots, Z_n)$ $\quad Z = \begin{pmatrix} Z_1 \\ | \\ Z_n \end{pmatrix} \in \mathbb{R}^n$

Z_0 est une constante ($\mathbb{V} = 0$) $= 1$

$X^* = \sum_{j=0}^{n} a_j^* Z_j$

Conditions d'orthogonalité $(n+1)$

$i = 0 \quad \mathbb{E}[(X - X^*)\,1] = 0 \iff \mathbb{E}[X] = \mathbb{E}[X^*]$
$$= a_0^* + \sum_{j=1}^{n} a_j^* \,\mathbb{E}[Z_j]$$
$$\overset{(1)}{=} a_0^* + (a^*)' \,\mathbb{E}[Z]$$

$i \in [\![1,n]\!] \quad \mathbb{E}[(X - X^*)Z_i] = 0 \iff \mathbb{E}[XZ_i] = \mathbb{E}[X^* Z_i]$
$$\overset{(2)}{=} a_0^* \,\mathbb{E}[Z_i] + \sum_{j=1}^{n} a_j \,\mathbb{E}[Z_j Z_i]$$

$(2) - \mathbb{E}[Z_i]\,(1)$

$$\Leftrightarrow \quad Cov(X, Z_i) = \sum_{j=1}^{n} a_j^* \, Cov(Z_i, Z_j)$$

$$\Leftrightarrow \quad \underbrace{Cov(X, Z)}_{(n,1)} = \underbrace{\mathbb{V}[Z]}_{(n,n)} \, \underbrace{a^*}_{(n,1)}$$

$$\Leftrightarrow \quad \underbrace{a^*}_{(n,1)} = \underbrace{\mathbb{V}(Z)^{-1}}_{(n,n)} \, \underbrace{Cov(Z, X)}_{(n,1)}$$

$$X^* = a_0^* + (a^*)' Z = \mathbb{E}[X] - (a^*)' \mathbb{E}[Z] + a^{*'} Z$$

$$= \mathbb{E}[X] + (a^*)' (Z - \mathbb{E}[Z])$$

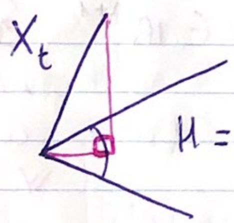
$$= \mathbb{E}[X] + Cov(X, Z) \, \mathbb{V}(Z)^{-1} (Z - \mathbb{E}[Z])$$

Rmq $H = ev(Z_1, -, Z_n)$ (sans la constante)

$$X^* = Cov(X, Z) \, \mathbb{V}(Z)^{-1} Z$$

si elles sont centrées pas besoin de la constante

↳ ce qu'on fera tout le temps ensuite



$H = ev(X_{t-1}, .., X_{t-p})$

$$X = \begin{pmatrix} X_t \\ | \\ X_t \end{pmatrix} (p,1)$$

$$Z = \begin{pmatrix} X_{t-1} \\ | \\ X_{t-p} \end{pmatrix}$$

$$X_t^* = \mathbb{EL}[X_t \mid X_{t-1}, .., X_{t-p}]$$

$$\mathbb{V}[Z] = \begin{pmatrix} \gamma_X(0) & \gamma_X(1) & .. & \gamma_X(p-1) \\ & \ddots & & | \\ & & \ddots & \gamma_X(1) \\ & & & \gamma_X(0) \end{pmatrix} (p,p)$$

$$Cov(X, Z) = \begin{pmatrix} \gamma_X(1) \\ | \\ \gamma_X(p) \end{pmatrix}$$

$(X_t) = (\varepsilon_t)$ Bruit Blanc Faible

non corrélée = pas d'information disponible du passé

$\mathbb{E}[\varepsilon_t] = 0$

$\mathbb{V}[\varepsilon_t] = \sigma_\varepsilon^2$

$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t'}) = 0 \quad \forall t \neq t'$

$\mathbb{EL}[\varepsilon_t \mid \varepsilon_{t-1}, .., \varepsilon_{t-p}] = 0$ ε_t pas prévisible ! car bruit

3. 3 Régression linéaire théorique sur un nombre ∞ de retards

Notations tout le passé

$\underline{X_{t-1}} = ev(X_{t-1}, X_{t-2}, X_{t-3}, ...)$

$\mathbb{EL}[X_t \mid \underline{X_{t-1}}] = \sum_{j=1}^{\infty} a_j^* X_{t-j}$

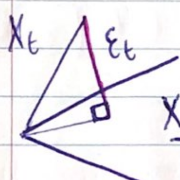
car $\lim_{p \to +\infty} \mathbb{EL}[X_t \mid X_{t-1}, .., X_{t-p}] = \sum_{j=1}^{\infty} a_j^* X_{t-j}$

(preuve non démontrée) par Banach

car $\sum_{j=1}^{\infty} |a_j| < +\infty$

Déf $\varepsilon_t = X_t - \mathbb{EL}[X_t \mid \underline{X_{t-1}}]$

erreur de prévision



$\underline{X_{t-1}}$ (le passé d'X_t)

ε_t - est l'innovation de $(X_t)_{t \in \mathbb{Z}}$ à la date t.

- est "non prévisible à partir du passé"

Prop i) $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc faible

ii) $\varepsilon_t \perp\!\!\!\perp \underline{X_{t-1}}$ (voir dessin)

on ne peut rien apprendre du passé si on a fait la meilleure prédiction

iii) Théorème de WOLD

tout processus non linéaire a une rep écriture linéaire

on peut donc toujours travailler en linéaire

$m_X + \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}$ mise à jour uniquement sur le passé !!

coefficient d'autocorrélation $\underbrace{\text{partielle}}$

car on enlève de l'information

$$\tilde{X}_t = X_t - \mathbb{EL}[X_t \mid X_{t-1}, .., X_{t-k-1}]$$

car on enlève l'information intermédiaire entre X_{t-k} et X_t

$r(k) = \mathrm{Cor}(\tilde{X}_t, \tilde{X}_{t-k})$ (Rappel $\rho_X(k) = \mathrm{Cor}(X_t, X_{t-k})$)

$k=1$ il n'y a rien entre $t-1$ et t
donc on projette sur rien

$r_X(1) = \rho_X(1)$

$$r_X(k) = \frac{\begin{vmatrix} & & \rho_X(1) \\ & // & | \\ & & \rho_X(k) \end{vmatrix}}{\begin{vmatrix} \rho_X(0) & \rho_X(1) & \text{—} & \rho_X(k-1) \\ & \ddots & \ddots & | \\ & & & \rho_X(1) \\ & & & \rho_X(0) \end{vmatrix}}$$

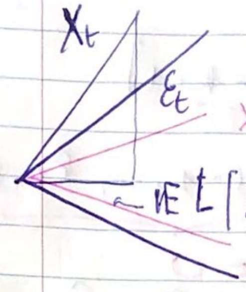
← même matrice sauf pour la dernière colonne

$k=2$

$$r_X(2) = \frac{\begin{vmatrix} \rho_X(0) & \rho_X(1) \\ \rho_X(1) & \rho_X(2) \end{vmatrix}}{\begin{vmatrix} \rho_X(0) & \rho_X(1) \\ \rho_X(1) & \rho_X(0) \end{vmatrix}}$$

$$\mathbb{EL}[X_t \mid X_{t-1}, .., X_{t-k-1}, X_{t-k}] = \sum_{j=1}^{k} a_j X_{t-j}$$

Prop $a_k = r(k)$



avec $\varepsilon_t \perp\!\!\!\perp$ ev $(X_{t-1}, .., X_{t-k})$

$X_{t-1}, .., X_{t-k-1}$ = le même espace vectoriel mais sans X_{t-k} !

$\mathbb{E}L[X_t | X_{t-1}, .., X_{t-k}]$

(1) $X_t = \sum_{j=1}^{k} a_j X_{t-j} + \varepsilon_t$

projection (linéaire)

projection de lui même est lui même

(2) $\mathbb{E}L[X_t | X_{t-1}, \dots X_{t-k-1}] = \sum_{j=1}^{k-1} a_j X_{t-j}$

$+ a_k \underbrace{\mathbb{E}L[X_{t-k} | X_{t-1}, \dots X_{t-k-1}]}$

$+ 0$

car ε_t orthogonal à n'importe quel élément sur lequel on projette / CL de l'espace n'importe quelle

(1) - (2) $\tilde{X}_t = \underbrace{a_k \tilde{X}_{t-k}} + \varepsilon_t$

on multiplie par $\tilde{X}_{t-k}$ puis on applique l' $\mathbb{E}$

$$\mathrm{Cov}(\tilde{X}_t, \tilde{X}_{t-k}) = a_k \mathbb{V}(\tilde{X}_{t-k}) + 0$$

$\mathrm{cov}(\varepsilon_t, \tilde{X}_{t-k})$

$$\Rightarrow a_k = \frac{\mathrm{Cov}(\tilde{X}_t, \tilde{X}_{t-k})}{\mathbb{V}[\tilde{X}_t]^{1/2}\, \mathbb{V}[\tilde{X}_{t-k}]} = \mathrm{Cor}(\tilde{X}_t, \tilde{X}_{t-k})$$

CL de $X_{t-1}, .., X_{t-k-1}$ et $\varepsilon_t \perp\!\!\!\perp X_{t-1} \vdots X_{t-k-1}$

$$\mathbb{V}[\tilde{X}_t] = \mathbb{V}[\tilde{X}_{t-k}]$$

$(X_t) = (\varepsilon_t)$ Bruit Blanc Faible

$$\tilde{\varepsilon}_t = \varepsilon_t - \underbrace{\mathbb{EL}[\varepsilon_t \mid \varepsilon_{t-1}, \dots, \varepsilon_{t-k-1}]}_{=0}$$

$$\tilde{\varepsilon}_{t-k} = \varepsilon_{t-k} - \underbrace{\mathbb{EL}[\varepsilon_{t-k} \mid \varepsilon_{t-1}, \dots, \varepsilon_{t-k-1}]}_{=0}$$

$$r_X(k) = \mathrm{Cov}(\tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-k}) = \mathrm{Cor}(\tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-k})$$
$$= 0 \quad \forall k \geq 1$$

C'est une autre façon de caractériser le bruit blanc faible
Seul lui a une covariance / corrélation partielle nulle

3.4 Densité spectrale et autocorrélations inverses

On prend $\mathbb{E}[X_t] = 0$

donc $\forall X_t$, on peut écrire $X_t = \sum_{j \in \mathbb{Z}} a_j \varepsilon_{t-j}$

avec $(\varepsilon_t) \sim BB(0, \sigma_\varepsilon^2)$

$\sum_{j \in \mathbb{Z}} |a_j| < \infty$.

* $\mathbb{E}[X_t] = 0$
* $\gamma_X(h) = \mathrm{Cov}(X_t, X_{t-h})$

$$= \mathrm{Cov}\left(\sum_j a_j \varepsilon_{t-j}, \sum_i a_i \varepsilon_{t-h-i}\right)$$

$$= \sum_j \sum_i a_j a_i \mathrm{Cov}(\varepsilon_{t-j}, \varepsilon_{t-h-i})$$

$$= \sum_j \sum_i a_j a_i \underbrace{\gamma_\varepsilon(h+i-j)}_{=0 \text{ sauf si } h+i-j=0} = \sum_i a_i a_{h+i} \sigma_\varepsilon^2$$

vu plus haut

(ça donne la variance)

on somme sur $\mathbb{Z}$ ici

$$\sum_h |\gamma_X(h)| = \sum_h \left| \sum_i a_i a_{i+h} \right| \sigma_\varepsilon^2$$
$$\leq \left(\sum_{i,h} |a_i| |a_{i+h}| \right) \sigma_\varepsilon^2$$
$$= \left(\sum_i |a_i| \right)^2 \sigma_\varepsilon^2 < +\infty$$

CDV $i+h=j$ puis on sépare les sommes

$$\boxed{f_X(\omega) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_X(h) e^{i\omega h} \quad \omega \in]-\pi, \pi[}$$

$$|f_X(\omega)| \leq \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} |\gamma_X(h)| < +\infty.$$

donc f_X est bien définie

f_X est réelle.

$$f_X(x) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \underbrace{\gamma_X(h)}_{\text{pair}} \Big(\underbrace{\cos(\omega h)}_{\text{pair}} + \underbrace{i \sin(\omega h)}_{\text{impair}} \Big)$$

$\sum_{h \in \mathbb{Z}} \gamma_X(h)\, i \sin(\omega h) = 0$ vont s'annuler car impaire

$f_X(\omega) = f_X(-\omega)$ pair

on peut donc juste la définir sur $[0, \pi[$ (sur $]-\pi, 0[$ on aura en faisant $-\omega$)

Prop $\quad \gamma_X(h) = \boxed{\int_{-\pi}^{\pi} f_X(\omega) \cos(\omega h)\, d\omega}^{(*)} \Big) + \underbrace{\int_{-\pi}^{\pi} i f_X(\omega) \sin(\omega h)\, d\omega}_{=0 \text{ car impair}}$

Preuve $\quad (*) = \int_{-\pi}^{\pi} f_X(\omega)\, e^{i\omega h}\, d\omega$

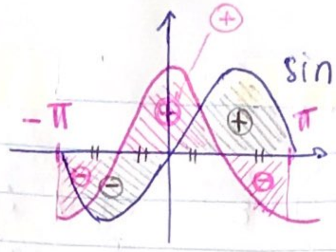
$$= \int_{-\pi}^{\pi} f_X(\omega)\, e^{-i\omega h}\, d\omega$$

CDV $x = -\omega$ mais f paire $f(x) = f(\omega)$

$$= \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \sum_k \gamma_X(k) e^{i\omega k} \right) e^{-i\omega h}\, d\omega$$

$$= \frac{1}{2\pi} \sum_k \gamma_X(k) \underbrace{\int_{-\pi}^{\pi} e^{i\omega(k-h)}\, d\omega}_{\substack{= 0 \text{ quand } k \neq h \\ = 2\pi \text{ quand } k = h}}$$

$$= \frac{1}{2\pi} 2\pi \gamma_X(h) = \gamma_X(h) \quad \blacksquare$$



on a $\int_{-\pi}^{\pi} e^{i\omega(k-h)} d\omega = 0$ quand $k \neq h$ car les intégrales en cos et sin sont nulles

sinus : $\int_{-\pi}^{\pi}$ = aire sous la courbe $= \oplus + \ominus = 0$

cosinus $\int_{-\pi}^{\pi}$ = aire sous la courbe $= \oplus + \ominus + \ominus = 0$
à droite de 0 à gauche de 0

Exemple 1

$$(X_t) = (\varepsilon_t) \qquad (\varepsilon_t) \text{ BB (faible)}$$
$$\gamma_\varepsilon(h) = 0 \quad h \neq 0$$

$$f_X(\omega) = \frac{1}{2\pi} \sum_h \gamma_\varepsilon(h) \cos(\omega h)$$
$$= \frac{1}{2\pi} \gamma_\varepsilon(0) = \frac{\sigma_\varepsilon^2}{2\pi}$$ ne dépend pas de ω !

Seul le bruit blanc faible a une densité spectrale constante

Exemple 2

MA(1) $\quad X_t = \varepsilon_t + \theta \varepsilon_{t-1} \quad (\varepsilon_t) \sim BB(0, \sigma_\varepsilon^2)$

$$\gamma_X(0) = (1+\theta^2)\sigma_\varepsilon^2$$
$$\gamma_X(1) = \theta \sigma_\varepsilon^2$$
$$\begin{cases} \gamma_X(2) = 0 & \text{car pas de terme en commun} \\ \gamma_X(k) = 0 & k > 2 \end{cases}$$

$$f_X(\omega) = \frac{1}{2\pi}\left(\gamma_X(0) + \gamma_X(1)\, 2\cos(\omega)\right)$$
$$= \frac{1}{2\pi}\sigma_\varepsilon^2\left((1+\theta^2) + 2\theta \cos\omega\right)$$
$$= A + B \cos\omega.$$

Seule une MA(1) a une densité spectrale $A + B\cos\omega$.

Exemple 3

$$(X_t) \xrightarrow{(a_i)} Y_t = \sum_{j \in \mathbb{Z}} a_j X_{t-j} \quad \text{avec } \sum_{j\in\mathbb{Z}} |a_j| < +\infty$$

$$f_X \xleftrightarrow{\text{relations ?}} f_Y$$

$$\gamma_X \xrightarrow[\text{en fonction de } \gamma_X]{\text{on peut écrire } \gamma_Y} \gamma_Y(h) = \sum_{i,j} a_i a_j \gamma_X(h+j-i)$$

$$f_Y(\omega) = \frac{1}{2\pi} \sum_h \gamma_Y(h) e^{i\omega h}$$

$$= \frac{1}{2\pi} \sum_h \sum_\ell \sum_j a_\ell a_j \gamma_X(h+j-\ell) e^{i\omega h} e^{i\omega(j-\ell)} \times e^{-i\omega j} \times e^{i\omega \ell}$$

$$= \boxed{\frac{1}{2\pi} \sum_h \boxed{\sum_\ell \sum_j a_\ell a_j} \gamma_X(h+j-\ell) e^{i\omega(h+j-\ell)}} e^{-i\omega j} e^{i\omega \ell}$$

$$= f_X(\omega) \underbrace{\sum_\ell a_\ell e^{i\omega \ell}}_{z} \times \underbrace{\sum_j a_j e^{-i\omega j}}_{\bar{z}}$$

$\rightarrow$ avec $z \in \mathbb{C}$: $z\bar{z} = |z|^2$

$$= f_X(\omega) \left| \sum_\ell a_\ell e^{i\omega \ell} \right|^2$$

norme

on pose $A(u) = \sum_\ell a_\ell u^\ell$ en série entière

$$= f_X(\omega) |A(e^{i\omega})|^2$$

Application MA(1) $X_t = \varepsilon_t + \theta \varepsilon_{t-1}$

$a_0 = 1 \quad a_1 = \theta \quad a_j = 0$ sinon

$$f_X(\omega) = f_\varepsilon(u) \left| 1 + \theta e^{i\omega} \right|^2$$

$$= \frac{\sigma_\varepsilon^2}{2\pi} \left((1 + \theta \cos(\omega))^2 + \theta \sin^2(\omega) \right)$$

$$= \frac{\sigma_\varepsilon^2}{2\pi} \left(1 + 2\theta \cos \omega + \theta^2 \right)$$

détails : $(1+\theta\cos\omega)^2 + \theta^2 \sin^2\omega$
$= 1 + 2\theta\cos\omega + \theta^2\cos^2\omega + \theta^2\sin^2\omega$
$\theta^2\cos^2\omega + \theta^2\sin^2\omega = \theta^2$

slides 84/285 à 129/285

25/10/23

Séries temporelles.

3, 4) Densité spectrale et autocorrélations inverses

innovation = erreur de prédiction basée sur tout l'historique du processus
c'est la ≠ entre la prédiction et toutes les informations du passé

transformée de Fourier = densité spectrale du processus stationnaire

on peut avoir la densité spectrale de X quand $X_t = \sum \psi_j Y_{t-j}$ avec $(Y_t)_{t\in\mathbb{Z}}$ de densité spectrale f_Y

$$f_X(\omega) = \underbrace{|\Psi(e^{i\omega})|^2}_{\text{norme au carré}} f_Y(\omega) \qquad \Psi(u) = \sum_j \psi_j u^j$$

Exemples

auto regressif d'ordre 1

a) AR(1) $X_t = \varphi X_{t-1} + \varepsilon_t$ avec $(\varepsilon_t)_{t\in\mathbb{Z}} \sim BB(0, \sigma_\varepsilon^2)$

$\mathbb{E}[X_t] = 0$ $\qquad |\varphi| < 1$

$\gamma_X(h) = \varphi^{|h|} \dfrac{\sigma_\varepsilon^2}{1-\varphi^2} \Rightarrow f_X(\omega) = \dfrac{1}{2\pi} \sum_{h\in\mathbb{Z}} \gamma_X(h) e^{i\omega h}$

calcul à faire !

on inverse le rôle entre le bruit et le processus

$\varepsilon_t = X_t - \varphi X_{t-1}$ ($\phi_0 = 1$, $\phi_1 = -\varphi$) on crée une série avec un nombre de termes défini

$\phi(u) = 1 - \varphi u$ (Ψ avec 2 termes : $\phi(u) = \sum_j \phi_j u^j$

$$\underbrace{f_\varepsilon(\omega)}_{\text{connu } = \frac{\sigma_\varepsilon^2}{2\pi}} = \underbrace{|\phi(e^{i\omega})|^2}_{=1-2\varphi\cos\omega+\varphi^2} f_X(\omega) \Rightarrow f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{1-2\varphi\cos\omega+\varphi^2}$$

$$= \frac{1}{A + B\cos\omega}$$

la densité spectrale caractérise AR(1)

Définition Autocorrélations inverses

$$(X_t)_{t\in\mathbb{Z}} \Leftrightarrow (\gamma_X(h))_{h\in\mathbb{Z}} \Leftrightarrow f_X(\omega) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \gamma_X(h) e^{i\omega h}$$

$$\Updownarrow$$

$$(Y_t)_{t\in\mathbb{Z}} \Leftrightarrow \gamma_Y(h) = \int_{-\pi}^{\pi} f_Y(\omega) e^{-i\omega h} d\omega \Leftrightarrow f_Y(\omega) = \frac{1}{f_X(\omega)}$$

$$\Downarrow$$

$$\rho_Y(h) = \mathrm{Corr}(Y_t, Y_{t-h}) = \rho^{i}_X(h)$$ autocorrélation inverse.

Exemple

MA(1)

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1} \Leftrightarrow \begin{cases} \gamma_X(0) = (1+\theta^2)\sigma^2_\varepsilon \\ \gamma_X(1) = \gamma_X(-1) = \theta\sigma^2_\varepsilon \\ \gamma_X(h) = 0 \text{ sinon} \end{cases}$$

$$\Leftrightarrow f_X(\omega) = \frac{\sigma^2_\varepsilon}{2\pi}(1+2\theta\cos\omega+\theta^2)$$

$$\Updownarrow$$

$$f_Y(\omega) = \frac{1}{f_X(\omega)} = \underbrace{\frac{1}{2\pi}}\ \frac{(2\pi)^2}{\sigma^2_\varepsilon}\ \frac{1}{1+2\theta\cos\omega+\theta^2}$$

on fait apparaître 2π au dénominateur

on a un AR(1)

$$\Updownarrow$$

$$Y_t = \underbrace{\varphi}\, Y_{t-1} + \eta_t = -\theta Y_{t-1} + \eta_t$$

en identifiant $\varphi = -\theta$

$\Rightarrow Y_t = -\theta Y_{t-1} + \eta_t$

où $(\eta_t) \sim BB(0, \sigma^2_\eta)$ en identifiant $\sigma^2_\eta = \frac{(2\pi)^2}{\sigma^2_\varepsilon}$

D'où $\rho^{i}_X(h) = \rho_Y(h) = (-\theta)^{|h|}$

fonction d'autocorrélation d'un AR(1)

La fonction d'autocorrélation inverse caractérise le processus MA(1)

(si $\rho^{i}_X(h) = \text{cste}^{|h|} \Rightarrow X$ MA(1))

En pratique, comment on fait ?

(X_t) IID $\mathbb{E}[X_t] = \mu_X$? on a nos observations $X_1, \dots, X_T$

notre estimateur sera la moyenne empirique

$$\hat{\mu}_{X,T} = \frac{1}{T}\sum_{t=1}^{T} X_t$$

SB $\Rightarrow \mathbb{E}[\hat{\mu}_{X,T}] = \mu_X$

converge $\Rightarrow \mathbb{V}[\hat{\mu}_{X,T}] \xrightarrow[T\to+\infty]{} 0$

$$\mathbb{V}[\hat{\mu}_{X,T}] = \frac{1}{T}\mathbb{V}[X_1] \text{ car iid.}$$

$$\xrightarrow[T\to+\infty]{} 0. \text{ (bien convergent)}$$

$$\sqrt{T}(\hat{\mu}_{X,T} - \mu_X) \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}(0, \mathbb{V}[X_1]) \text{ par TCL.}$$

résultat utile pour les intervalles de confiance et les tests d'égalité.

Rappel

- erreur de 1ère espèce : on accepte H_1 alors que c'est H_0 qui est vraie
- erreur de 2e espèce : on accepte H_0 alors que c'est H_1 qui est vraie

tout cela repose sur l'hyp iid

puissance du test $= 1 - \underset{\sim}{\alpha} = 1 -$ erreur de 2ème espèce

risque $= 5\%$

alors que nous on sera pas indépendant avec nos séries temporelles

3.3 Estimation de la moyenne et des autocovariances pour des processus linéaires

$X_t = \sum_j a_j \varepsilon_{t-j}$ avec $\sum_j |a_j| < +\infty$ et $(\varepsilon_t)_{t\in\mathbb{Z}} \sim BB(0, \sigma_\varepsilon^2)$

On a nos observations $X_1, \dots, X_T$

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t \xrightarrow[T\to+\infty]{L^2} 0 \iff \mathbb{V}[\bar{X}_T] \xrightarrow[T\to+\infty]{} 0$$

Calculons

$$\mathbb{V}[\bar{X}_T] = \frac{1}{T^2}\left(\sum_{t=1}^{T}\mathbb{V}[X_t] + 2\sum_{1\le t\le t'\le T} \mathrm{Cov}(X_t, X_{t'})\right)$$

matrice de variance covariance

$$\mathbb{V}\begin{pmatrix} X_1 \\ | \\ X_T \end{pmatrix} = \begin{pmatrix} \gamma_X(0) & \boxed{\gamma_X(1)} & \gamma_X(2) \dots & \gamma_X(T-1) \\ & & & \vdots \\ & & & \gamma_X(2) \\ & & & \gamma_X(1) \\ & & & \gamma_X(0) \end{pmatrix}$$

on somme en colonne

D'où $T\,\mathbb{V}[\bar{X}_T] = \gamma_X(0) + \frac{2}{T}\sum_{k=1}^{T}(T-k)\gamma_X(k)$

$$= \gamma_X(0) + \frac{2}{T}\sum_{k=1}^{T-1}\sum_{j=1}^{k}\gamma_X(j)$$

on pose $\beta_k = \sum_{j=1}^{k}\gamma_X(j) \xrightarrow[k\to+\infty]{} \sum_{j=1}^{+\infty}\gamma_X(j)$

D'où $T\,\mathbb{V}[\bar{X}_T] \xrightarrow[k\to+\infty]{} \gamma_X(0) + \frac{2}{T}\sum_{k=1}^{T-1}\beta_k \xrightarrow[T\to+\infty]{} \gamma_X(0) + 2\beta_\infty$

c'est le Lemme de Cesaro.

[si une Σ converge alors la moyenne des Σ cvg aussi et vers la m̂ limite]

avec $\beta_\infty = \sum_{h\in\mathbb{Z}}\gamma_X(h)$

$\Rightarrow \mathbb{V}[\bar{X}_T] \underset{T\to+\infty}{\sim} \frac{1}{T}\sum_{h\in\mathbb{Z}}\gamma_X(h) \xrightarrow[T\to+\infty]{} 0$ vitesse en $\frac{1}{T}$

on a donc un TCL

$$\sqrt{T}\,(\bar{X}_T - 0) \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, \sum_{h\in\mathbb{Z}}\gamma_X(h)\right)$$

on a $\gamma_X(h) = \sum_j a_j\, a_{j-h}\, \sigma_\varepsilon^2$

$\Rightarrow \sum_{h\in\mathbb{Z}}\gamma_X(h) = \sum_{h\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} a_j\, a_{j-h}\, \sigma_\varepsilon^2 = \left(\sum_{j\in\mathbb{Z}} a_j\right)^2\sigma_\varepsilon^2$

Exercice $X_t = \varepsilon_t - \varepsilon_{t-1}$

$\frac{1}{T}\sum_{t=1}^{T}X_t = \frac{1}{T}(\varepsilon_T - \varepsilon_0)$ $\quad \mathbb{V}\left[\frac{1}{T}\sum_{t=1}^{T}X_t\right] = \frac{1}{T^2}\,2\sigma_\varepsilon^2$

vitesse en $\frac{1}{T^2}$

L'estimateur de l'autocovariance n'est pas sans biais.
Exemples avec des processus stationnaires.
Exemples d'identification avec autocovariance partielle
autocovariance
autocorrélation ...
.....

★ BBF (Fort)
autocorrélogramme ACF
en 0 ça fait 1 OK
ailleurs nul théoriquement OK
La quasi totalité des valeurs calculées sont entre les pointillés bleus (intervalles de confiance) donc on accepte statistiquement que c'est nul.
autocorrelation partielle (pas défini en 0)
Partial ACF que 0. c'est normal
$H_0 : (= 0)$ on accepte H_0 car dans les pointillés bleus.

★ MA (1)
ACF
1 en 0 OK
$\neq 0$ en 1 et -1 OK
nul ailleurs statistiquement vrai donc OK

Partial ACF décroissance exponentielle
en pratique nul partout sauf 0.
alors qu'on est censé être $\neq$ de 0.

★ AR (1) ⇒ 1 seul bâton significatif dans Partial ACF
ACF décroissance exponentielle

Partial ACF $\neq$ en 0 : doit être égal à $\overbrace{0,5}^{=\varphi}$ $(r_X(1) = \varphi)$
le reste est nul (statistiquement vrai) OK

Exemples avec des Processus non stationnaires

1) marche aléatoire

on rend stationnaire en faisant la différence

2) $m_t + K_t \varepsilon_t$ on centre et réduit pour stationnariser

L'autocorrélogramme peut servir sur l'indication de la stationnarité.

méthode Buys Ballot

trend on décroît lentement

méthode STL

⇒ il reste trend

on a des vagues

il reste une saisonnalité !

il y a une dépendance au temps qui perdure (pendant au moins 2 périodes)

on a pas bien éliminé la saisonnalité.

Exemples séries réelles

variance change au cours du temps

modèle hétéroscédastique.

Rentabilités log.

ACF ⇒ on voudrait conclure BBF

Partial ACF ⇒ statistiquement nul

On fait log au carré.

ACF valeurs significatives & décroissance lente

⇒ dépendance forte et longue

c'est la volatilité des marchés financiers qui persiste

Conclusion

BB faible en regardant log et log au carré

Chapitre 4 Processus ARMA, ARIMA, SARIMA.

4.1) Polynômes retard et avance

Opérateur retard L : $LX_t = X_{t-1}$
opérateur avancé F : $FX_t = X_{t+1}$

$L = F^{-1} \Leftrightarrow L^{-1} = F \Leftrightarrow L \times F = Id$

$L^k X_t = X_{t-k}$ $\quad F^k X_t = X_{t+k}$ $\quad$ coefficients réels.

Polynôme associé à L $P(z) = \sum_{k=0}^{p} a_k z^k$ $\Rightarrow P(L) = \sum_{k=0}^{p} a_k L^k$

$A(L) = \sum_{k \in \mathbb{Z}} a_k L^k$

$\boxed{A(L) X_t = \sum_{k \in \mathbb{Z}} a_k X_{t-k} \quad (X_t)_t \text{ stationnaire existe si } \sum_{k \in \mathbb{Z}} |a_k| < +\infty.}$

Produit de 2 polynômes retards infinis $A(L)$ $B(L)$ est un polynôme retard infini

$A(L) B(L) = C(L) = \sum_{k \in \mathbb{Z}} c_k L^k \quad (\triangle\, c_k \neq a_k b_k)$

Preuve On a $A(L) B(L) = \sum_k a_k \sum_l b_l L^{k+l}$

$= \sum_j \underbrace{\left(\sum_k a_k b_{j-k} \right)}_{c_j} L^j$ $\quad$ produit de convolution

$j = k + l$

■

▷ Inversibilité des polynômes en L

$$A(L) = \sum_{k=0}^{p} a_k L^k \quad \exists B(L) \,/\, A(L)B(L) = B(L)A(L) = Id \; ?$$

Coefficients réels $\Rightarrow$ on peut factoriser par des monômes.

$$A(L) = a_p \prod_{k=1}^{p} (L - \mu_k) = a_0 \prod_{k=1}^{p} (1 - \lambda_k L)$$

μ_k racines avec $\lambda_k = \frac{1}{\mu_k}$

On doit maintenant regarder si les monômes sont inversibles.

$Q = (1 - \lambda L)$ inversible $\Leftrightarrow \exists B(L) \quad (1 - \lambda L) B(L) = Id.$

Cas 1 $|\lambda| < 1$

on cherche "$(1 - \lambda L)^{-1}$" $=$ "$\frac{1}{1 - \lambda L}$" $= \underbrace{\sum_{k=0}^{\infty} \lambda^k L^k}_{\text{bien défini car } \sum |\lambda|^k < +\infty.}$

on développe sur les termes passés

et $(1 - \lambda L) \underbrace{\sum_{k=0}^{\infty} \lambda^k L^k}_{} = \sum_{k=0}^{\infty} \lambda^k L^k - \sum_{k=1}^{\infty} \lambda^k L^k = \lambda^0 L^0 = 1.$

L'inverse existe et c'est ⑤

Cas 2 $|\lambda| > 1$

$$(1 - \lambda L) = \underbrace{-\lambda L}_{\text{inversible}} \underbrace{\left(1 - \frac{1}{\lambda} F\right)}_{\substack{\text{inversible} \\ \left|\frac{1}{\lambda}\right| < 1}}$$

$$\Rightarrow (1 - \lambda L)^{-1} = -\frac{1}{\lambda} F \sum_{k=0}^{\infty} \frac{1}{\lambda^k} F^k = -\sum_{k=1}^{\infty} \frac{F^k}{\lambda^k} = -\sum_{k=1}^{\infty} \frac{L^{-k}}{\lambda^k}$$

ici on développe sur le futur.

Cas 3 $|\lambda| = 1$ non inversible
Exercice à faire par raisonnement par l'absurde quand $\lambda = 1$.

$A(L)$ est inversible si toutes ses racines sont de module $\neq$ de 1.

Rmq : $A(L) = a_0 \prod_{k=1}^{p} (1 - \lambda_k L)$

Si toutes les racines sont de module > 1
$(|\lambda_k| < 1)$
$A^{-1}(L) = \sum_{k=0}^{\infty} b_k L^k$ avec $\sum |b_k| < \infty$.
$k \in \mathbb{N}$

S'il existe une racine de module >1 et 1 de module < 1
$A^{-1}(L) = \sum_{k \in \mathbb{Z}} b_k L^k$ on regarde dans le futur
mais on a pas les observations du futur

Ne pas retenir Methode 1
la méthode de calculer l'inverse de chaque monôme et d'en faire le produit
$\Rightarrow$ trop de produits de convolution à identifier.

Methode 2 on identifie les coefficients.
on connait les ϕ_i
$a_1 = -\phi_1$
$a_2 = -\phi_2 + \phi_1^2$
$\vdots$

Methode 3 ne marche qu'en multiplicité 1

4.2) Processus MA(q)

$(\varepsilon_t) \sim BB(0, \sigma_\varepsilon^2)$ bruit blanc faible.

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

on a q retards

on augmente la dépendance temporelle de X_t.

X_t, X_{t-1} ... X_t, X_{t-q} corrélés (2 à 2).

Processus stationnaire par construction

$$X_t = \underbrace{\Theta(L)}_{\text{définition d'un polynôme retard}} \varepsilon_t \quad \text{avec } \Theta(u) = 1 - \theta_1 u - \dots - \theta_q u^q$$

(ε_t) est-il le processus des innovations de (X_t) ?

càd $\varepsilon_t = EL[X_t \mid \underline{X_{t-1}}]$ avec

projection linéaire.

$\underline{X_{t-1}} = ev\{X_{t-1}, X_{t-2}, X_{t-3} \dots\}$

ensemble du passé du processus

$X_t \in \underline{\varepsilon_t}$ car CL des $\varepsilon_t, \varepsilon_{t-1}, \dots$ par définition

et $X_{t-1} \in \underline{\varepsilon_{t-1}} \subset \underline{\varepsilon_t}$ donc $\boxed{\underline{X_t} \subset \underline{\varepsilon_t}}$

Hypothèse les racines de Θ sont de module > 1

$\Rightarrow \Theta^{-1}(L) = \sum_{k \in \mathbb{N}} b_k L^k$

$$\boxed{\varepsilon_t = \Theta^{-1}(L) X_t = \sum_{k \in \mathbb{N}} b_k X_{t-k}} \Rightarrow \varepsilon_t \in \underline{X_t}$$

$\varepsilon_{t-1} \in \underline{X_{t-1}} \subset \underline{X_t}$

$\Rightarrow \boxed{\underline{\varepsilon_t} \subset \underline{X_t}}$

$\Rightarrow \underline{\varepsilon_t} = \underline{X_t}$

$$\mathbb{E}L[X_t \mid \underline{X_{t-1}}] \overset{\text{changement de base}}{=} \mathbb{E}L[X_t \mid \underline{\varepsilon_{t-1}}]$$

base orthogonale car non corrélé 2 à 2.

$$= \mathbb{E}L[\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \mid \underline{\varepsilon_{t-1}}]$$

$\varepsilon_t \perp \varepsilon_k$
$k \in [\![t-1 ; \dots]\!]$

$\mathbb{E}[\varepsilon_{t-1} \mid \underline{\varepsilon_{t-1}}]$
$= \varepsilon_{t-1}$ invariant car dans l'espace

$$= 0 - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

$$= X_t - \varepsilon_t .$$

donc on a bien $\varepsilon_t = X_t - \mathbb{E}L[X_t \mid \underline{X_{t-1}}]$

ε_t est le processus d'innovation de X_t.

29/11/23 SERIES TEMPORELLES

MA(q) $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$

$\varepsilon_t = X_t - \mathbb{E}L[X_t | \underline{X_{t-1}}]$?

Définition $\varepsilon_t = X_t - \mathbb{E}L[X_t | \underline{X_{t-1}}]$ ...
si (ε_t) est l'innovation.

Dans ce cas, $X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$ est l'écriture canonique.

$\Theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ est le polynôme retard

Prop L'écriture canonique existe si les racines de Θ sont de module > 1

Preuve

Soient μ_i les q racines de Θ, $\lambda_i = \frac{1}{\mu_i}$

$\Rightarrow \Theta(L) = \prod_{i=1}^{q} (1 - \lambda_i L)$

comme $|\mu_i| > 1 \Leftrightarrow |\lambda_i| < 1 \Rightarrow (1 - \lambda_i L)^{-1} = \sum_{k=0}^{\infty} \lambda_i^k L^k$

$\Rightarrow \Theta^{-1}(L) = \sum_{k=0}^{\infty} a_k L^k$

On a $X_t = \Theta(L) \varepsilon_t$ $\quad \mathcal{L}(X_t) \subset \mathcal{L}(\underline{\varepsilon_t})$

$\Rightarrow \mathcal{L}(\underline{X_t}) \subset \mathcal{L}(\underline{\varepsilon_t})$

NB
$\mathcal{L}$ = Historique / Histoire

$\varepsilon_t = \Theta^{-1}(L) X_t = \sum_{k=0}^{\infty} a_k X_{t-k}$ $\quad \mathcal{L}(\varepsilon_t) \subset \mathcal{L}(\underline{X_t})$

$\Rightarrow \mathcal{L}(\underline{\varepsilon_t}) \subset \mathcal{L}(\underline{X_t})$

D'où $\mathcal{L}(\underline{\varepsilon_t}) = \mathcal{L}(\underline{X_t})$

$\mathbb{E}L[X_t | \underline{X_{t-1}}] = \mathbb{E}L(X_t | \underline{\varepsilon_{t-1}})$

$= \mathbb{E}L(\underbrace{\varepsilon_t}_{\notin \underline{\varepsilon_{t-1}}} - \theta_1 \underbrace{\varepsilon_{t-1} \dots - \theta_q \varepsilon_{t-q}}_{\in \underline{\varepsilon_{t-1}}} | \underline{\varepsilon_{t-1}})$

$= 0 - \theta_1 \varepsilon_{t-1} \dots - \theta_q \varepsilon_{t-q}$

D'où $\boxed{X_t - \mathbb{EL}[X_t \mid \underline{X_{t-1}}] = \varepsilon_t}$

Maintenant on considère que

$$\Theta(L) = \prod_{\substack{|\lambda_i|<1 \\ |\mu_i|>1}} (1-\lambda_i L) \prod_{\substack{|\lambda_i|>1 \\ |\mu_i|<1}} (1-\lambda_i L)$$

$$\Theta^*(L) = \prod_{|\lambda_i|<1} (1-\lambda_i L) \prod_{|\lambda_i|>1} \left(1-\frac{1}{\lambda_i} L\right)$$

Proposition

On se retrouve avec 2 écritures équivalentes

$$X_t = \Theta(L)\varepsilon_t \qquad \varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$$

$$X_t = \Theta^*(L)\eta_t \qquad \eta_t \sim BB(0, \sigma_\eta^2)$$

$$\sigma_\eta^2 = \sigma_\varepsilon^2 \prod_{|\lambda_i|>1} |\lambda_i|^2$$

Preuve

$$X_t = \Theta(L)\varepsilon_t \rightsquigarrow f_X(\omega) = \underbrace{|\Theta(e^{i\omega})|^2}_{(*)} \frac{\sigma_\varepsilon^2}{2\pi}$$

X est une transformation linéaire de ε_t qui est stationnaire

$$X_t = \Theta^*(L)\eta_t \rightsquigarrow f_X(\omega) = \overbrace{|\Theta^*(e^{i\omega})|^2}^{(**)} \frac{\sigma_\eta^2}{2\pi}$$

$$(*) = \prod_{|\lambda_j|<1} |1-\lambda_j e^{i\omega}|^2 \prod_{|\lambda_j|>1} |1-\lambda_j e^{i\omega}|^2$$

$$(**) = \prod_{|\lambda_j|<1} |1-\lambda_j e^{i\omega}|^2 \prod_{|\lambda_j|>1} \left|1-\frac{1}{\lambda_j} e^{i\omega}\right|^2$$

On regarde pour $|\lambda_j|>1$, ce que vaut $\dfrac{|1-\lambda_j e^{i\omega}|^2}{\left|1-\frac{1}{\lambda_j} e^{i\omega}\right|^2}$

$$= \frac{\cancel{|1-\lambda_j e^{i\omega}|^2}}{\left|\frac{1}{\lambda_j} e^{i\omega}\right|^2 \cancel{|1-\lambda_j e^{-i\omega}|^2}} = \frac{1}{\left|\frac{1}{\lambda_j} e^{i\omega}\right|^2} = |\lambda_j|^2 \text{ indépendant de } \omega$$

sont conjugués donc égaux en module

en posant $\sigma_\eta^2 = \sigma_\varepsilon^2 \prod_{|\lambda_j|>1} |\lambda_j|^2$

On a bien que nos 2 écritures spectrales pour X sont égales. ■

D'où (η_t) est l'innovation

$\Rightarrow X_t = \Theta^*(L)\eta_t$ est l'écriture canonique

$\Rightarrow \eta_t = (\Theta^*(L))^{-1} X_t = \sum_{k=0}^{\infty} a_k X_{t-k}$

4.3 Processus autorégressifs d'ordre p : AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

avec $(\varepsilon_t)_{t\in\mathbb{Z}} \sim BB(0, \sigma_\varepsilon^2)$

$\Leftrightarrow \phi(L) X_t = \varepsilon_t$ avec $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$

Rmq $X_t = \mu + \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} + \varepsilon_t$

Si $(X_t)_t$ est stationnaire alors

$$\mathbb{E}[X_t] = \mu + \phi_1 \underbrace{\mathbb{E}[X_{t-1}]}_{=\mathbb{E}[X_t]} - \dots - \phi_p \underbrace{\mathbb{E}[X_{t-p}]}_{=\mathbb{E}[X_t]} + \underbrace{0}_{\mathbb{E}[\varepsilon_t]=0 \text{ car } BB}$$

car stationnaire

$\Rightarrow \mathbb{E}[X_t] = \dfrac{\mu}{1-\phi_1 - \dots - \phi_p}$ si $\underbrace{1-\phi_1 - \dots - \phi_p}_{=\phi(1)} \neq 0$

si 1 n'est pas racine de ϕ

(X_t) existe si $X_t = \underbrace{\phi^{-1}}(L)\varepsilon_t$

$= \sum_{k\in\mathbb{Z}} b_k L^k$ avec $\sum_{k\in\mathbb{Z}} |b_k| < \infty$.

$\Leftrightarrow$ toutes les racines de ϕ sont de module $\neq$ de 1

Exemple $X_t = X_{t-1} + \varepsilon_t \qquad \phi_1 = 1$

$\Leftrightarrow (1-L) X_t = \varepsilon_t$

$\phi(L) = 1 - L \qquad$ racine $= 1$

ϕ est non inversible donc (X_t) non stationnaire

On peut le rendre stationnaire en faisant la différence.

Proposition $\phi(L) X_t = \varepsilon_t$ est l'écriture canonique

$\Leftrightarrow \varepsilon_t = X_t - \mathbb{E}L[X_t \mid \underline{X_{t-1}}]$ (ε_t est l'innovation)

$\Leftrightarrow$ les racines de ϕ sont de module > 1

Preuve

$\varepsilon_t = \phi(L) X_t \Rightarrow \mathcal{L}(\varepsilon_t) \subset \mathcal{L}(\underline{X}_t)$

$\Rightarrow \mathcal{L}(\underline{\varepsilon}_t) \subset \mathcal{L}(\underline{X}_t)$

$X_t = \phi^{-1}(L)\varepsilon_t \Rightarrow \mathcal{L}(X_t) \subset \mathcal{L}(\underline{\varepsilon}_t)$

$= \sum_{k=0}^{\infty} b_k \varepsilon_{t-k} \qquad \Rightarrow \mathcal{L}(\underline{X}_t) \subset \mathcal{L}(\underline{\varepsilon}_t)$

$\longrightarrow \mathcal{L}(\underline{\varepsilon}_t) = \mathcal{L}(\underline{X}_t)$

racines module $\neq 1$
et même module > 1
d'où l'inverse est

$$\mathbb{E}L[X_t \mid \underline{X_{t-1}}] = \mathbb{E}L[\phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \mid \underline{X_{t-1}}]$$

$$= \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + 0$$

$\mathbb{E}[\varepsilon_t \mid \underline{X_{t-1}}]$
$= \mathbb{E}[\varepsilon_t \mid \underline{\varepsilon_{t-1}}]$
$= 0$ car ε_t BB donc 2 à 2 orthogonaux

D'où $X_t - \mathbb{E}L[X_t \mid \underline{X_{t-1}}] = \varepsilon_t$

Conclusion $(\varepsilon_t)_t$ est l'innovation. ∎

Proposition $\phi(L) X_t = \varepsilon_t \qquad \phi^*(L) X_t = \eta_t$

$$\phi(L) = \prod_{|\lambda_i|<1} (1-\lambda_i L) \prod_{|\lambda_i|>1} (1-\lambda_i L)$$

$$\phi^*(L) = \prod_{|\lambda_i|<1} (1-\lambda_i L) \prod_{|\lambda_i|>1} \left(1-\frac{1}{\lambda_i} L\right)$$

Pour que la densité spectrale soit identique

$$\sigma_\eta^2 = \frac{1}{\prod_{|\lambda_i|>1} |\lambda_i|^2} \sigma_\varepsilon^2$$

$\phi(L) X_t = \varepsilon_t$ forme canonique $\quad (\mathbb{E}[X_t] = 0)$

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t$$

$\times X_t$ et on prend $\mathbb{E}$

$$\mathbb{V}[X_t] = \gamma_X(0) = \phi_1 \gamma_X(1) + \cdots + \phi_p \gamma_X(p) + \underbrace{\mathbb{E}[X_t \varepsilon_t]}_{= \mathrm{Cov}(X_t, \varepsilon_t)}$$

on fait $\times \varepsilon_t$ et on prend l'$\mathbb{E}$

$$\mathbb{E}[X_t \varepsilon_t] = \phi_1 \underbrace{\mathbb{E}[X_{t-1} \varepsilon_t]}_{=0} + \cdots + \phi_p \underbrace{\mathbb{E}[X_{t-p} \varepsilon_t]}_{=0} + \underbrace{\mathbb{E}[\varepsilon_t^2]}_{= \mathbb{V}(\varepsilon_t) = \sigma_\varepsilon^2}$$

car $\varepsilon_t \perp\!\!\!\perp \underline{X_{t-1}}$
par la forme canonique
car (ε_t) est l'innovation

$\Rightarrow \mathbb{E}[X_t \varepsilon_t] = \sigma_\varepsilon^2$

on fait $\times X_{t-h}$ et on prend l'$\mathbb{E}$

$$\boxed{\mathbb{E}[X_t X_{t-h}] = \gamma_X(h) = \phi_1 \gamma_X(h-1) + \phi_2 \gamma_X(h-2) + \cdots + \phi_p \gamma_X(h-p)} + \underbrace{\mathbb{E}[\varepsilon_t X_{t-h}]}_{=0}$$

Equation de récurrence linéaire

$$- \phi_p \gamma_X(h-p) - \phi_{p-1}\gamma_X(p-h+1) - \dots - \phi_1 \gamma_X(h-1) + \gamma_X(h) = 0$$

" $\alpha_0 z_n + \dots + \alpha_p z_{n+p} = 0$ "

Le polynôme associé est $P(z) = -\phi_p - \phi_{p-1} z - \dots - \phi_1 z^{p-1} + z^p$

$$\Rightarrow -\phi_p - \phi_{p-1} z - \dots - \phi_1 z^{p-1} + z^p = z^p \phi(z^{-1})$$

polynôme retard

Les racines de P sont les inverses des racines de ϕ donc elles sont de module < 1

de module > 1 car forme canonique

Conclusion $\gamma_X(h) \xrightarrow[h \to 0]{} 0$

$\gamma_X(h)$ converge exponentiellement vite vers 0

$\Rightarrow$ il n'y aura plus de cov entre X_t et X_{t-h}

$\Rightarrow$ on aura pas besoin de revenir loin en arrière dans l'historique pour faire de la prévision

La dépendance décroît exponentiellement vite vers 0

On écrit sous forme matricielle les équations de récurrence linéaire $\forall h$.

$$\begin{pmatrix} \gamma_X(1) \\ \gamma_X(2) \\ \vdots \\ \vdots \\ \gamma_X(p) \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma_X(0) & \gamma_X(1) & \gamma_X(2) & \dots & \gamma_X(p-1) \\ \gamma_X(1) & \gamma_X(0) & \gamma_X(1) & \dots & \gamma_X(p-2) \\ & \ddots & \ddots & \ddots & \vdots \\ & & & & \gamma_X(1) \\ & & & \gamma_X(1) & \gamma_X(0) \end{pmatrix}}_{\text{matrice d'autocovariance}} \begin{pmatrix} \phi_1 \\ \vdots \\ \vdots \\ \phi_p \end{pmatrix}$$

Système de Yule Walker

connue à partir de nos données et inversible

Autocorrélations partielles de $(X_t)_{t\in\mathbb{Z}}$

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

$$\mathbb{E}L\left[X_t \mid X_{t-1}, \dots, X_{t-k}\right] \quad \text{avec } k > p$$

$$= \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + 0$$

car ε_t est l'innovation

$r_X(k) = 0$ car le coefficient devant X_{t-k} est 0

$k = p \quad r_X(p) = \phi_p$

Autocorrélation inverse de $(X_t)_{t\in\mathbb{Z}}$

$$\phi(L) X_t = \varepsilon_t$$

$$|\phi(e^{i\omega})|^2 f_X(\omega) = f_\varepsilon(\omega)$$

$$\Rightarrow f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|\phi(e^{i\omega})|^2}$$

$$\Rightarrow \frac{1}{f_X(\omega)} = \frac{2\pi}{\sigma_\varepsilon^2} |\phi(e^{i\omega})|^2 = \underbrace{\frac{1}{2\pi} \frac{(2\pi)^2}{\sigma_\varepsilon^2} |\phi(e^{i\omega})|^2}_{= f_Y(\omega)}$$

avec $Y_t = \phi(L)\eta_t$ avec $(\eta_t) \sim BB\left(0, \frac{(2\pi)^2}{\sigma_\varepsilon^2}\right)$

$Y_t \sim MA(p)$

Conclusion $\rho_X^i(h) = \rho_Y(h) = 0$ pour $|h| > p$.

Exemple AR(1) $X_t = \phi_1 X_{t-1} + \varepsilon_t$

$\phi(L) = 1 - \phi_1 L$

- $|\phi_1| < 1$ forme canonique

$$X_t = \underbrace{\phi^{-1}(L)}_{|\phi_1| \neq 1 \text{ donc inversible}} \varepsilon_t = \left(\sum_{k=0}^{\infty} \phi_1^k L^k\right) \varepsilon_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k}$$

- $|\phi_1| > 1$ module $\neq 1$ donc inversible

$$X_t = \phi^{-1}(L)\varepsilon_t = -\left(\sum_{k=1}^{\infty} \phi_1^{-k} \underbrace{L^{-k}}_{=F^k}\right) \varepsilon_t$$

car $(-\varepsilon_t)_t$ $BB(0, \sigma_\varepsilon^2)$ $\overset{\downarrow}{=} \sum_{k=1}^{\infty} \phi_1^{-k} \varepsilon_{t+k}$

NB : si $(\varepsilon_t)_t$ BB $(0, \sigma_\varepsilon^2)$ $(-\varepsilon_t)_t$ BB aussi

$\mathbb{E}[-\varepsilon_t] = -\mathbb{E}[\varepsilon_t] = 0$

$\mathbb{V}[-\varepsilon_t] = \mathbb{V}[\varepsilon_t] = \sigma_\varepsilon^2$

$Cov(-\varepsilon_t, -\varepsilon_{t'}) = Cov(\varepsilon_t, \varepsilon_{t'}) = 0$

on change pour obtenir la forme canonique

$\phi^*(L) = 1 - \phi_1^{-1} L$

$X_t = \phi_1^{-1} X_{t-1} + \eta_t$ avec $\sigma_\eta^2 = \dfrac{\sigma_\varepsilon^2}{\phi_1^2}$

- $|\phi_1| < 1$

$\mathbb{E}[X_t] = 0$

$\mathbb{V}[X_t] = \dfrac{1}{1 - \phi_1^2} \sigma_\varepsilon^2$

- $|\phi_1| > 1$

$\mathbb{E}[X_t] = 0$

$$\mathbb{V}[X_t] = \sum_{k=1}^{\infty} \phi_1^{-2k} \sigma_\varepsilon^2 = \phi_1^{-2} \frac{1}{1 - \phi_1^{-2}} \sigma_\varepsilon^2 = \frac{1}{\phi_1^2 - 1} \sigma_\varepsilon^2$$

pas en forme canonique

et

on a la même variance

$\mathbb{E}[X_t] = 0$

$\mathbb{V}[X_t] = \dfrac{\sigma_\eta^2}{1 - \phi_1^{-2}} = \dfrac{\sigma_\varepsilon^2}{\phi_1^2 - 1}$ en forme canonique

- $|\phi_1| < 1$

$$X_t = \phi_1 X_{t-1} + \varepsilon_t$$

on multiplie par X_{t-1} et on prend l'$\mathbb{E}$

$$\gamma_X(1) = \phi_1 \gamma_X(0)$$

avec X_{t-h} au lieu de X_{t-1}

$$\gamma_X(h) = \phi_1 \gamma_X(h-1) = \phi_1^h \gamma_X(0) = \phi_1^h \underbrace{\frac{\sigma_\varepsilon^2}{1-\phi_1^2}}_{= \mathbb{V}[X_t]}$$

$$= \phi_1^2 \gamma_X(h-2)$$

$$= \phi_1^3 \gamma_X(h-3) \dots$$

$r_X(1) = \phi_1 = \rho_X(1)$ car $\phi_1 = \frac{\gamma_X(1)}{\gamma_X(0)} = \rho_X(1)$ par def

$r_X(h) = 0$ pour $h > 1$

$$f_X(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|1-\phi_1 e^{i\omega}|^2} \qquad X_t = \phi_1 X_{t-1} + \varepsilon_t \sim AR(1)$$

$$f_Y(\omega) = \frac{1}{f_X(\omega)} = \frac{1}{2\pi} \frac{(2\pi)^2}{\sigma_\varepsilon^2} |1-\phi_1 e^{i\omega}|^2 \qquad Y_t = \eta_t - \phi_1 \eta_{t-1} \sim MA(1)$$

$$\rho_X^i(h) = \rho_Y(h) = \frac{\gamma_Y(h)}{\gamma_Y(0)} = \begin{cases} \dfrac{-\phi_1}{1+\phi_1^2} & \text{pour } h = \pm 1 \\ 0 & \text{pour } |h| > 1 \end{cases}$$

7/12/2023

SÉRIES TEMPORELLES.

4.4 Processus ARMA(p, q)

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

avec (ε_t) l'innovation (donc orthogonal au passé de X_{t-1})
On multiplie par X_{t-h} et on prend l'$\mathbb{E}$
pour $h > q$

$$\gamma_X(h) = \phi_1 \gamma_X(h-1) + \cdots + \phi_p \gamma_X(h-p)$$

On écrit pour $h = q+1, \ldots, q+p$ pour obtenir p équations de Yule-Walker.

4.5 Processus ARIMA(p, d, q)

Marche aléatoire $X_t = X_{t-1} + \varepsilon_t$ $\quad (\varepsilon_t) \sim BB(0, \sigma_\varepsilon^2)$

$$\left.\begin{array}{l} X_0 \\ \underbrace{(1-L)}_{=\Delta} X_t = \varepsilon_t \end{array}\right\} \Rightarrow \text{NON STATIONNAIRE}$$

$$\Delta X_t = \varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$$

$(1-L)^d$ on différencie d fois.

ARIMA(p, d, q) $\quad \phi(L)\Delta^d X_t = \Theta(L)\varepsilon_t \quad d \geq 1$

degré $\phi = p \Leftrightarrow \phi_p \neq 0$
degré $\Theta = q \Leftrightarrow \Theta_q \neq 0$

non stationnaire
= homogène dans le temps

Donc on pose des conditions initiales car on ne peut pas laisser le processus venir de n'importe quel instant dans le temps.

4.6 SARIMA$(p, d, q)(P, D, Q)$

NB :
Sert pour saisonnalité

	Janv	Fév		...	Déc
	X_1	X_2	X_3	...	X_{12}
L^{12}	X_{13}				X_{24}
de l'année 1 sur 2	X_{25}				X_{36}

L →

dynamique temporelle avec une dépendance mensuelle
(ex vente de bonnets : si Déc froid, Janv sera sûrement froid aussi)

On a une double dynamique temporelle

effet annuel — effet mode = dynamique annuelle (magazines, ...)

polynôme retard d'une année sur l'autre

$$\phi(L^{12}) X_t = \Theta(L^{12}) U_t \quad ARMA(P, Q)$$

de l'année 1 sur 2

effet mensuel

$$\phi(L) U_t = \Theta(L) \varepsilon_t \quad ARMA(p, q)$$

$$(\varepsilon_t) \sim BB(0, \sigma_\varepsilon^2)$$

$$\Rightarrow \phi(L^{12}) \phi(L) X_t = \Theta(L^{12}) \Theta(\varepsilon_t)$$

$$SARMA_{12} \underbrace{(p, q)}_{\text{mensuel}} \underbrace{(P, Q)}_{\text{annuel}}$$

Exemple

* $SARIMA_{12}(0, 0, 1)(1, 0, 0)$ correspond à

$$X_t - \varphi X_{t-12} = \varepsilon_t + \theta \varepsilon_{t-1}$$

avec comme conditions $|\varphi| < 1$, $|\theta| < 1$ pour que X_t soit stationnaire

$$X_t = (1 - \varphi L^{12})^{-1} (1 + \theta L) \varepsilon_t$$

$$= \sum_{k=0}^{\infty} \varphi^k L^{12k} (1 + \theta L) \varepsilon_t$$

$$= \varepsilon_t + \theta \varepsilon_{t-1} + \varphi \varepsilon_{t-12} + \varphi \theta \varepsilon_{t-13} + \varphi^2 \varepsilon_{t-24} + \varphi^2 \theta \varepsilon_{t-25} + \dots$$

** $\mathbb{E}[X_t] = 0$

*** $\mathbb{V}[X_t] = (1+\theta^2)\frac{1}{1-\varphi^2}\sigma_\varepsilon^2$

(NB les ε_t sont de Cov = 0)

$\mathbb{V}[X_t]$ existe car processus stationnaires

on a $X_t = \varphi X_{t-12} + \varepsilon_t + \theta\varepsilon_{t-1}$ (*)

$\overset{\mathbb{V}}{\Rightarrow} \mathbb{V}[X_t] = \varphi^2\mathbb{V}[X_{t-12}] + (1+\theta^2)\sigma_\varepsilon^2 + \varphi\,\mathrm{Cov}(X_{t-12}, \varepsilon_t)$

$+ \varphi\theta\,\mathrm{Cov}(X_{t-12}, \varepsilon_{t-1}) + \theta\,\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-1})$

$\underset{\text{stationnaire}}{\Rightarrow} \mathbb{V}[X_t] = \frac{1}{1-\varphi^2}(1+\theta^2)\sigma_\varepsilon^2$

d'où $\mathbb{V}[X_t] = \mathbb{V}[X_{t-12}]$

(*) $\times X_{t-1}$ et on prend l'$\mathbb{E}$: $\gamma_X(1) = \varphi\gamma_X(11) + 0 + \theta\,\underbrace{\mathrm{Cov}(X_{t-1}, \varepsilon_{t-1})}_{=\sigma_\varepsilon^2}$

(*) $\times \varepsilon_t$ et on prend l'$\mathbb{E}$:

$\mathrm{Cov}(X_t, \varepsilon_t) = 0 + \sigma_\varepsilon^2 + 0$

(*) $\times X_{t-11}$ et on prend l'$\mathbb{E}$:

$\gamma_X(11) = \varphi\gamma_X(1)$

D'où $\gamma_X(1) = \frac{1}{1-\varphi}\theta\sigma_\varepsilon^2$

Pour avoir $\gamma_X(2)$:

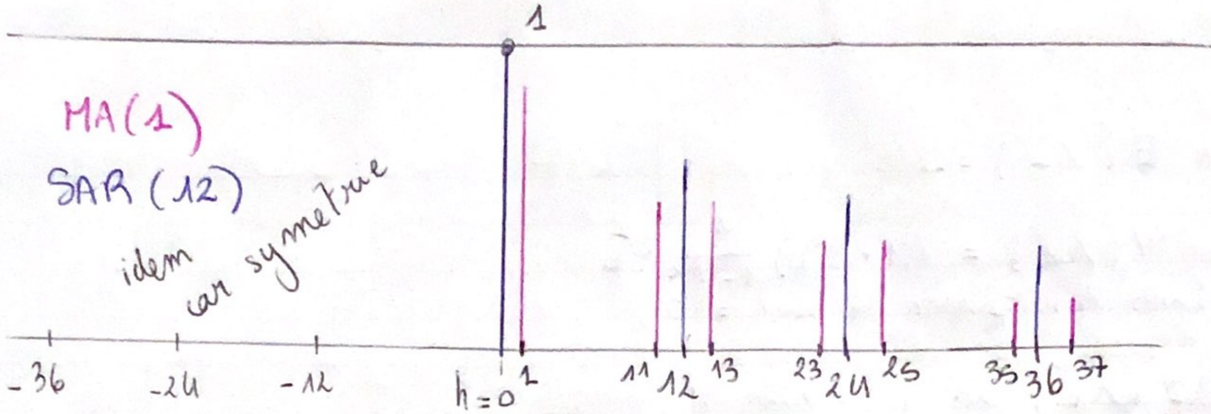
$\left.\begin{array}{l}(*) \times X_{t-2} \text{ et } \mathbb{E} : \gamma_X(2) = \varphi\gamma_X(10) \\ (*) \times X_{t-10} \text{ et } \mathbb{E} : \gamma_X(10) = \varphi\gamma_X(2)\end{array}\right\} \Rightarrow \begin{array}{l}\gamma_X(2) = 0 \\ \gamma_X(10) = 0\end{array}$

NB $\mathrm{Cov}(X_{t-h}, \varepsilon_{t-1})$ pour $h > 1$ $= 0$ car innovation

Ce qui donne $\rho_X(12h) = \varphi^h$

$\rho_X(12h-1) = \rho_X(12h+1) = \frac{\theta}{1+\theta^2}\varphi^h$

$\rho_X(h) = 0$ sinon



effet autorégressif tous les ans
= influence d'une année sur l'autre
décroissance en $\varphi^{h/12}$

effet MA
= influence d'un mois sur l'autre

Chapitre 5
Estimation ARMA et ARIMA

On doit estimer p, d et q qui sont des entiers naturels !
On ne peut pas estimer aussi facilement que des réels que l'on peut approcher au fur et à mesure et qu'on finira par atteindre avec TCL. Alors que pour les entiers naturels soit c'est bon soit c'est faux.
Il faut faire une identification a priori (de p, d et q) en proposant des valeurs réalistes / appropriées

Ensuite il faut déterminer les coefficients des polynômes.

Méthodes usuelles : maximum de vraisemblance, moindres carrées, méthode des moments

Vérification
* On doit tester que ϕ_p et θ_q soient significativement différents de 0 : on recommence le modèle avec un degré en moins etc...

* Vérifier que (ε_t) est un BB. de nombreux en estimant son autocorrélation (tests existent) en estimant d'abord ε_t en partant de

$$\phi(L) X_t = \Theta(L)\varepsilon_t \Rightarrow \Theta^{-1}(L)\phi(L)X_t = \varepsilon_t$$

Choix d'un modèle

* utiliser un choix de comparaison de modèles (si plusieurs bons modèles)

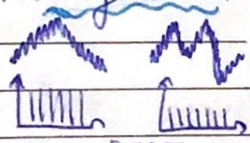
5.2

1ère étape identifier si stationnaire ou non ?

si non, car par ex | trend, saisonnalité → déterministes

on les enlève

∃ des tests pour vérifier si stationnaire mais le prof préfère d'abord qu'on fasse des analyses graphiques en regardant d'abord la série elle-même et sa forme, l'autocorrélogramme décroit très vite vers 0 exponentiellement quand stationnaire



NON STATIONNAIRE

et on aura des autocorrélogrammes comme ça (inverses et partielles)

⚠ à ne pas sur-différencier car on se rapproche du module = 1

par exemple $X_t = \varepsilon_t$

$$\Rightarrow \underset{\substack{\uparrow \\ MA(1)}}{\Delta X_t} = \varepsilon_t - \varepsilon_{t-1} = \underbrace{(1-L)}\varepsilon_t$$

racines du polynôme doivent être > à 1.

MA(q) $\rho_X(h) = 0 \quad |h| > q$ Test $H_0 : \gamma_X(q+1) = 0 \quad \gamma_X(q+2) = 0 \ldots$

$H_1 : \gamma_X(q+1) \neq 0$

AR(p) $\rho_X(h) = 0 \quad |h| > p$

$r_X(h) = 0 \quad h > p$

Loi asymptotique

$$\sqrt{T}\,(\hat{\rho}_X(h+q) - \underbrace{\rho_X(h+q)}_{h \geq 1}) \underset{T \to +\infty}{\sim} \mathcal{N}\left(0,\ 1 + \sum_{j=1}^{q} \rho_X^2(j)\right)$$

autocorrelogramme empirique tend vers autocorrelogramme théorique.

Région d'acceptation 5 %

$$\left[0,\ \underset{\downarrow\ 1{,}96^2}{4} / T \left(1 + \sum \underbrace{\hat{\rho}_X^2(j)}\right)\right]$$ Test du Khi 2.

on remplace ρ par son estimateur.

Si on a une seule autocorrélation en 0 $\Rightarrow$ MA(1)
($q = 1$)

Pour p (du AR(p)) 1 seule valeur de l'autocorrélation.
partielle ou inverse

$\Rightarrow$ AR(1) ($\Rightarrow p = 1$)

Mais si c'est un ARMA(p, q) comment estime t'on p et q ?
on va regarder si ce sont des AR(p^*) ou MA(q^*)

$$\phi(L) X_t = \Theta(L)\varepsilon_t \Rightarrow X_t = \phi^{-1}(L)\,\Theta(L)\,\varepsilon_t$$

$$= \sum_{k=0}^{\infty} \underbrace{b_k}\, \varepsilon_{t-k}$$

décroît exponentiellement vite vers 0.

$$\approx \sum_{k=0}^{q^*} b_k\, \varepsilon_{t-k}$$

$q^* \geq q$ $\quad p^* \geq p$ On doit trouver q^* et p^* de telle sorte que ARMA(p, q) = MA(q^*) et ARMA(p, q) = AR(p^*)
Comme ils sont forcément supérieurs à p et q on va descendre en valeurs pour arriver aux vrais p et q.

Estimation d'un AR(p) $X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t$

Exemple

AR(1) $X_t = \phi X_{t-1} + \varepsilon_t$ représentation canonique

$\Leftrightarrow |\phi| < 1$

$$\gamma_X(1) = \phi \gamma_X(0)$$

$$\phi = \rho_X(1)$$

$$\hat{\phi}_T = \frac{\frac{1}{T}\sum_{t=1}^{T-1} X_t X_{t-1}}{\frac{1}{T}\sum_{t=1}^{T} X_t^2}$$

on a $\sqrt{T}\,(\hat{\phi}_T - \phi) \xrightarrow[T\to+\infty]{} \mathcal{N}\left(0, \frac{\sigma_\varepsilon^2}{\gamma_X(0)}\right)$ car $\gamma_X(0)$

$\sqrt{T}\,(\hat{\phi}_T - \phi) \xrightarrow[T\to+\infty]{} \mathcal{N}(0, 1-\phi^2)$

$= \mathbb{V}[X_t]$

$= \dfrac{\sigma_\varepsilon^2}{1-\phi^2}$

Vraisemblance

Densité associée à $(X_1, \dots, X_T)$

$f_{(X_1,\dots,X_T)}(x_1,\dots,x_T)$ inconnue on suppose $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$

mais on connaît $f_{X_t | X_{t-1}}(x_t, x_{t-1})$

$X_t = \phi X_{t-1} + \varepsilon_t \Rightarrow X_t \,|\, (X_{t-1} = x_{t-1}) \sim \mathcal{N}(\phi x_{t-1}, \sigma_\varepsilon^2)$

d'où $f_{X_t|X_{t-1}}(x_t, x_{t-1}) = \dfrac{1}{\sqrt{2\pi}} \dfrac{1}{\sigma_\varepsilon} e^{-\frac{1}{2\sigma_\varepsilon^2}(x_t - \phi x_{t-1})^2}$

$X_t = \sum_{k=0}^{\infty} \phi^k \varepsilon_{t-k} \Rightarrow X_t \sim \mathcal{N}\left(0, \dfrac{\sigma_\varepsilon^2}{1-\phi^2}\right)$

pas la même loi $\mathcal{N}$

$f_{(X_1,\dots,X_T)}(x_1,\dots,x_T) = \underbrace{f_{X_T|X_{T-1},\dots,X_1}}_{= f_{X_T|X_{T-1}}} f_{X_{T-1},\dots,X_1}$ car $X_{T-2}, \dots, X_1$ n'apporte pas d'informations supplémentaires car $X_t = \phi X_{t-1} + \varepsilon_t$

Exemple

MA(1) $X_t = \varepsilon_t + \theta \varepsilon_{t-1}$

$\eta_t = Y_t + \theta Y_{t-1}$

$\hat{\theta}_T$ Estimateur du Maximum de Vraisemblance (EMV)

on a $(Y_t) \sim AR(1) \Rightarrow \mathbb{V}[Y_t] = \frac{\sigma_\eta^2}{1-\theta^2}$

$\sqrt{T}(\hat{\theta}_T - \theta) \rightarrow \mathcal{N}(0, 1-\theta^2)$

$X_t = \phi X_{t-1} + \varepsilon_t$