

Séries temporelles : régression, et modélisation ARIMA(p,d,q)

2 décembre 2012

Enseignant : Florin Avram

Objectif : La prédiction des phénomènes spatio-temporaux est une des préoccupations principales dans la science, par exemple en statistique, géostatistique, économétrie, météorologie, sciences environnementales, ..., etc.

Méthodes : Les méthodes principales sont la décomposition des moindres carrés déterministe, et la régression "statistique", qui va au delà de mc en offrant un "modèle stochastique", en analysant ses résidus, et en produisant des intervalles de confiance.

Le choix entre les divers modèles possibles n'est pas toujours évident, finalement il s'agit d'un art plutôt que d'une science.

Nous allons aborder ces thèmes dans le contexte des séries temporelles uni-dimensionnelles, en commençant par l'approche déterministe : inspection graphique, et identification de la tendance et de la partie saisonnière par moindres carrés. Déjà dans cette partie, l'utilisation des filtres, qui sont des opérateurs de moyenne mobile entre deux séries, s'avère utile.

En suite, nous examinons une approche itérative statistique qui consiste à raffiner des modèles ARIMA ou des modèles d'espace-état, choisis conformément aux tests pour les résidus. Dans cette partie, le concept de filtre est essentiel.

Compétences acquises : Les étudiants apprendront à utiliser des diverses méthodes de filtrage et prévision des séries temporelles, notamment par la modélisation ARMA, et à tester les résidus pour évaluer la fiabilité des modèles choisis.

Matériels :

1. Notes de cours WEB : Julien Jacques, Rossignol, A. Charpentier, M. Kratz, J-M. Dufour (en Français) et R.H. Smith, R. Weber (** En Anglais), etc
 - <http://www.crest.htfr/pageperso/lfa/charpent/charpent.htm#TS>
 - <http://www.math-info.univ-paris5.fr/kratz/cours.html>
 - <http://www.statslab.cam.ac.uk/rrw1/timeseries/index.html>
2. Notes de cours, qui utilisent parties des notes de C. Rabut (INSA-Toulouse) sur les splines, et de M. Lavielle, A. Korabinski et C. Paroissin sur les séries temporelles (toutes les coquilles sont de ma responsabilité).
3. A. C. Harvey, Time Series Models.
4. J. Durbin and S.J. Koopman, Time series analysis by state space methods.
5. C. Gourieroux et A. Monfort, Cours de séries temporelles.

Table des matières

1	Introduction	3
2	Premier abord aux séries temporelles/chroniques	4
2.1	Les composantes d'une chronique	4
2.2	Quelques types de décomposition	7
2.2.1	le modèle additif	7
2.2.2	le modèle multiplicatif(*)	7
2.2.3	modèles mixtes(*)	8
2.3	Recurrences et operateurs de retard et avance	9
3	Filtres/moyennes mobiles	9
3.1	Filtres qui enlèvent les composantes périodiques et saisonnières	10
3.2	Filtres de lissage	11
3.3	La reduction du bruit des observations	12
3.4	A quoi servent les filtres ?	13
3.5	Exercices : TD 1	14
4	Modélisation stochastique des séries temporelles	15
4.1	Introduction	15
4.2	Processus stochastiques stationnaires	17
5	Exemples des processus stationnaires	18
5.1	Le bruit blanc	18
5.2	Les moyennes mobiles MA(q) : l'exemple le plus simple de processus a composantes dépendantes	19
5.3	Les modèles autorégressifs AR(p) : des processus faciles de prédire	21
6	Fonctions de transfert et applications	22
6.1	Causalité des modèles AR(p)	22
6.2	La récurrence de Yule-Walker pour les coefficients de la fonction de transfert	23
6.3	Inversibilité des processus MA(q)	23
6.4	Équations de Yule-Walker pour les covariances/corrélations des modèles autorégressifs AR(p)	24
7	Contrôle continu d'entraînement	27
8	Les modèles ARMA(p,q)	29
8.1	Causalité et inversibilité des modèles ARMA(p,q)	30
8.2	(*)L'inversion des series des puissances et des filtres $\varphi(B)$	32
8.3	La prévision linéaire des processus ARMA(p,q)	33
8.4	Prévision des processus AR(p) par la récurrence de Yule-Walker	34
9	Les modèles ARIMA(p,d,q)	36
9.1	Prévision linéaire des modèles autorégressifs ARIMA(p,d,0)	37
9.2	Prévision des modèles à bruit moyenne mobile MA(1)	38
9.3	Prévision des processus ARMA(p,1)	39

9.4	Prévision des processus ARIMA(p,d,q)	40
9.5	Exercices : TD 2	40
10	L'estimation des modèles ARIMA(p,d,q)	43
10.1	Équations de Yule-Walker pour les covariances/corrélations des processus ARMA(p,q)	43
10.2	La détermination de l'ordre d'un modèle autorégressif	44
10.3	Intervales de confiance pour la prédiction	45
10.4	Exercices : TD 3	45
11	Sommaire des définitions et résultats dans les séries temporelles	48
11.1	Filtres	48
11.2	Causalité et inversibilité des modèles ARMA(p,q)	48
11.3	Équations Yule Walker	49
11.4	Prévision linéaire	49
12	Examens d'entraînement	50
13	Examen d'entraînement 1	50
13.1	Examen d'entraînement 3	52
13.2	Examen d'entraînement 4	54
13.3	Examen d'entraînement 2	56
14	TP Series Temporelles	58
14.1	Tendance et composante périodique	58
14.2	Arima, Regression	59
14.3	Projets	60

1 Introduction

Définition 1.1 Une série chronologique (ou temporelle) est une succession d'observations au cours du temps : $\{U_t : t = 1, 2, \dots, n, \dots\} = (U_1, U_2, \dots, U_n, \dots)$

Par rapport aux autres types de données statistiques, la particularité des séries chronologiques tient à la présence d'une relation d'antériorité qui ordonne l'ensemble des informations. Les dates d'observations sont souvent équidistantes les unes des autres : on a des séries mensuelles, trimestrielles, etc, dans quel cas on peut les indexer par $t \in \mathbb{N}$.

Exemples : a) Nombre des moutons par année en Angleterre, entre 1867 et 2003. b) Nombre de voyageurs par mois (SNCF) entre 1990 et 2003. c) Nombre de voitures vendues par un garage, par trimestre entre 1995 et 1999. d) Taux de mortalité, per age, entre 55 et 104 (c'est le premier exemple d'utilisation de splines, par Whittaker (1923)).

Les séries temporelles sont le plus simple exemple d'une thématique plus large : l'estimation et prévision des processus stochastique, i.e. des familles des variables aléatoires $U(x)$. Pour les séries temporelles/chronologiques, on s'intéresse en $x \in \mathbb{N}, \mathbb{Z}$ ou \mathbb{R}_+ , pendant que dans la statistique spatiale, (par exemple en géostatistique) on s'intéresse dans le cas $x \in \mathbb{Z}^d$ ou $x \in \mathbb{R}^d$.

On se propose d'estimer la valeur de la variable $U(x)$ en un point x quelconque connaissant les valeurs $U(x_i)$ aux points de mesure donnés x_i , pour $i = 1, \dots, N$. Le but principal est le choix d'un modèle ("estimation") raisonnable, qui permettra à partir des valeurs connues la prédiction

des valeurs inobservables (comme les valeurs futures des séries temporelles, ou moins accessibles physiquement, coûteuses, etc). On veut à la fois : a) enlever du bruit d'observation éventuel et b) "extrapoler" du connu au inconnu.

Domaines d'application :

- Prospection et exploitation pétrolières et minières
- Traitement du signal
- Imagerie médicale
- Océanographie, météorologie, hydrogéologie, environnement, ...
- Séries temporelles, appliquées en économie, finances, météo, médecine, ...

2 Premier abord aux séries temporelles/chroniques

Une règle générale en statistique descriptive consiste à commencer par regarder ses données, avant d'effectuer le moindre calcul. Ainsi, la figure 1 montre différentes séries chronologiques, qui méritent quelques commentaires.

- La consommation des ménages en Allemagne et le Produit Intérieur Brut en France semblent avoir augmenté régulièrement.
- Le taux de chômage en Allemagne semble avoir globalement augmenté depuis 1960, mais avec une alternance de baisses et de hausses soudaines. Le taux de chômage des Etats-Unis ne semble pas évoluer globalement, mais présente également cette alternance de baisses et de hausses.
- Les ventes de champagnes, tout comme la production industrielle semblent exhiber un caractère périodique (ventes importantes de champagne en fin d'année, baisse de la production industrielle en été, ...).
- D'autre part, les variations de ces 2 séries (indice de production industrielle et ventes de champagne) ont une amplitude qui semble augmenter au cours du temps.
- Toutes ces séries ont un aspect irrégulier. Ces fluctuations irrégulières ont parfois une amplitude anormalement élevée (PIB et production industrielle en France au second trimestre 1968, consommation en Allemagne en 1991).

Cette liste de remarques n'est bien sûr pas exhaustive. Elles traduisent simplement quelques comportements que l'on retrouve sur la plupart des séries chronologiques. Puisque notre ambition est de décrire et d'analyser ce genre de chroniques, il nous faut donc proposer des modèles qui intègrent les différentes caractéristiques que nous venons de relever.

2.1 Les composantes d'une chronique

Dans un premier temps, l'examen graphique de la série étudiée ($y_i, 1 \leq i \leq n$) permet de dégager, lorsqu'on envisage une période de temps suffisamment longue, un certain nombre de composantes fondamentales de l'évolution de la grandeur étudiée.

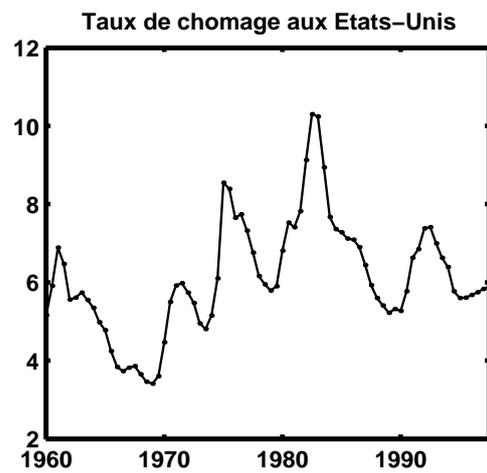
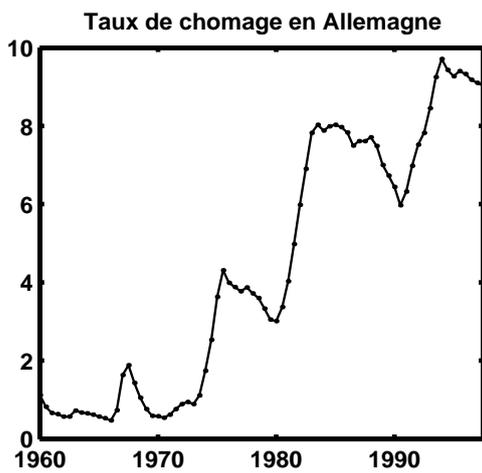
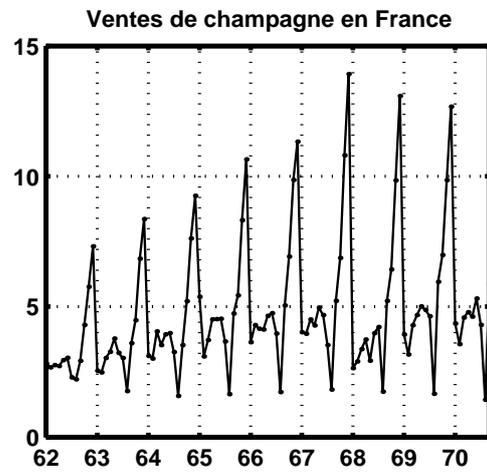
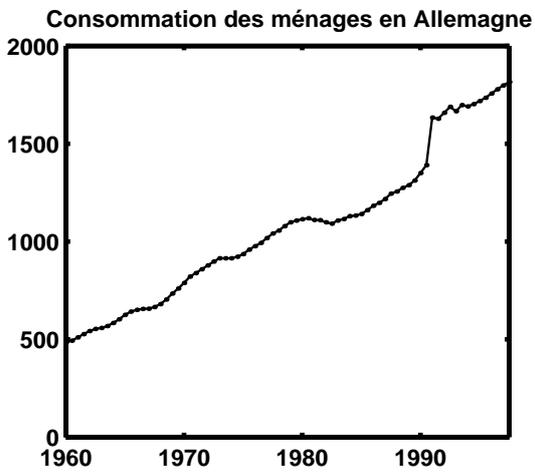
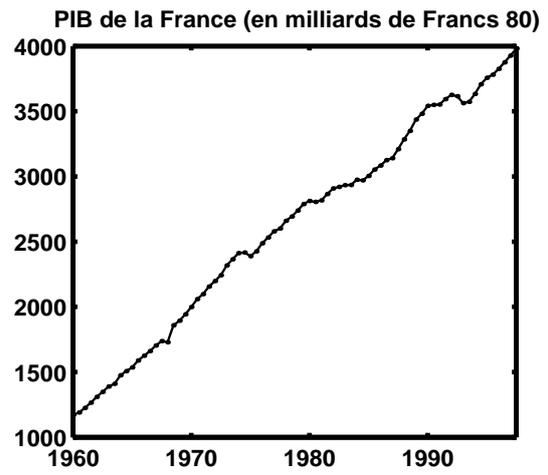
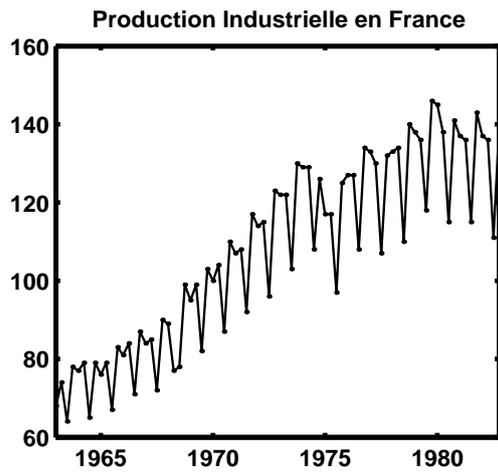


FIGURE 1 – Quelques exemples de séries chronologiques

Il faut alors analyser ces composantes, en les dissociant les unes des autres, c'est-à-dire en considérant une série comme résultant de la combinaison de différentes composantes, tel que chacune d'elles ait une évolution simple.

1. *La tendance* ($f_i, 1 \leq i \leq n$) représente l'évolution à long terme de la grandeur étudiée, et traduit l'aspect général de la série. C'est une fonction monotone, souvent polynomiale.
2. *Les variations saisonnières* ($s_i, 1 \leq i \leq n$) sont liées au rythme imposé par les saisons météorologiques (production agricole, consommation de gaz, ...), ou encore par des activités économiques et sociales (fêtes, vacances, solde, etc).

Mathématiquement, ce sont des fonctions périodiques, c'est-à-dire qu'il existe un entier p , appelé *période*, tel que

$$s_i = s_{i+p}, \forall i \geq 1 \quad (1)$$

Cette composante est entièrement déterminée par ses p premières valeurs s_1, s_2, \dots, s_p .

On rencontre souvent aussi des phénomènes pour les quelles la période peut elle même varier. On parle alors de

3. *Cycles* ($c_i, 1 \leq i \leq n$), qui regroupent des variations à période moins précise autour de la tendance, par exemple les phases économiques d'expansion et de recession. Ces phases durent généralement plusieurs années, mais n'ont pas de durée fixe. Sans informations spécifiques, il est généralement très difficile de dissocier la tendance du cycle. Dans le cadre de ce cours, la composante appelée tendance regroupera pour la plupart du temps aussi les cycles.
4. *Les fluctuations irrégulières/résidues/bruit* ($e_i, 1 \leq i \leq n$) sont des variations de faible intensité et de courte durée, et de nature aléatoire (ce qui signifie ici, dans un cadre purement descriptif, qu'elles ne sont pas complètement expliquables). En effet, elles ne sont pas clairement apercevables dans les graphiques, à cause de leur faible intensité par rapport aux autres composantes. Elles apparaissent clairement seulement après "l'enlèvement du signal" ; la question qui se posera alors sera : est-ce qu'ils contiennent encore du signal, ou est-ce que c'est vraiment du "bruit" ?
5. *Les variations accidentelles/observations aberrantes* sont des valeurs isolées anormalement élevées ou faibles de courte durée. Ces variations brusques de la série sont généralement explicables (Mai 68, réunification de l'Allemagne, tempête, ...). La plupart du temps, ces accidents sont intégrés dans la série des bruits (les fluctuations irrégulières).
6. *Points de changement* Ce sont des points où la série change complètement d'allure, par exemple de tendance. Ils sont normalement explicables, et imposent une analyse séparée de la série, par morceaux.

En résumé, nous considérerons une série chronologique comme issue de la composition de 3 composantes :

($f_i, 1 \leq i \leq n$) la tendance (intégrant éventuellement un cycle),

($s_j, 1 \leq j \leq p$) les coefficients saisonniers,

($e_i, 1 \leq i \leq n$) les fluctuations irrégulières (intégrant éventuellement des accidents).

Exemple : Trouvez l'élément suivant de la série y_t ci-dessous, une équation de récurrence pour

$$y_t = \{1, 3, 7, 13, 21, 31, \dots\}$$

Obtenez une formule analytique pour y_t , en utilisant :

- a) la théorie des équations de récurrence à coefficients constants. R : $n^2 + n + 1$
 b) la méthode des fonctions génératrices, décomposition en fractions partielles et l'expansion en série des puissances :

$$\frac{1}{(a-z)^k} = \sum_{n=0}^{\infty} C_{k-1}^{n+k-1} \frac{z^n}{a^{n+1}}, \quad |z| \leq a$$

$$R : a(z) = \frac{1+z^2}{(1-z)^3} = \frac{2}{(1-z)^3} - \frac{2}{(1-z)^2} + \frac{1}{(1-z)}$$

2.2 Quelques types de décomposition

Après avoir détecté graphiquement quelles sont les composantes présentes, il faut proposer un modèle :

2.2.1 le modèle additif

$$y_i = f_i + s_i + e_i, \quad 1 \leq i \leq n. \quad (2)$$

Pour bien séparer la tendance de la composante saisonnière, et pour des raisons d'unicité dans la décomposition proposée, on impose que la somme des facteurs saisonniers soit nulle :

$$\sum_{j=1}^p s_j = 0. \quad (3)$$

Exemple : Imaginons que nous étudions la série des températures moyennes relevées chaque mois en un même site, depuis janvier 1990, et que la tendance (plutôt faible) a une allure linéaire. Le modèle additif est :

$$y_i = a + bi + \sum_{k=1}^{11} s_k 1_{i \cong k \pmod{12}} - \left(\sum_{k=1}^{11} s_k \right) 1_{i \cong 0 \pmod{12}} + e_i$$

Les coefficients a, b, s_1, \dots, s_{11} et les résidus peuvent être déterminés en minimisant la somme des carrés des résidus $\sum_i e_i^2$, i.e. par régression.

Que peut-on dire des composantes présentes dans cet exemple ?

- la série (f_i) représente la tendance générale (réchauffement ? cycle ?).
- Les données étant mensuelles, la période est de un an, et donc $p = 12$.
- Des valeurs $s_1 = -10$ et $s_6 = +8$ signifient que le mois de janvier est plus froid de 10° par rapport à l'ensemble de l'année, alors que juin est plus chaud de 8° .
- Une fluctuation irrégulière $e_{14} = -2$ signifie qu'il a fait 2° de moins que prévu pour un mois de février, en 1991 (c'est-à-dire ce que nous laissaient prévoir la tendance et l'effet saisonnier pour février 1991).

2.2.2 le modèle multiplicatif(*)

$$y_i = f_i(1 + s_i)(1 + e_i), \quad 1 \leq i \leq n. \quad (4)$$

Là encore, on impose que la somme des facteurs saisonniers soit nulle : $\sum_{j=1}^p s_j = 0$.

Dans ce modèle, on considère maintenant que les amplitudes des fluctuations dépendent du niveau. Considérons le nombre d'entrées quotidiennes dans un cinéma. Des valeurs $s_4 = -0.5$ et $s_6 = +0.8$ signifient ici que la fréquentation de cette salle diminue de 50% le jeudi et augmente de 80% le samedi (par rapport à l'ensemble de la semaine). Une valeur $e_9 = +0.2$ signifie que le nombre d'entrée du deuxième mardi a été de 20% supérieur au chiffre attendu pour ce jour là.

Remarque : Le modèle multiplicatif est généralement utilisé pour des données de type économique.

2.2.3 modèles mixtes(*)

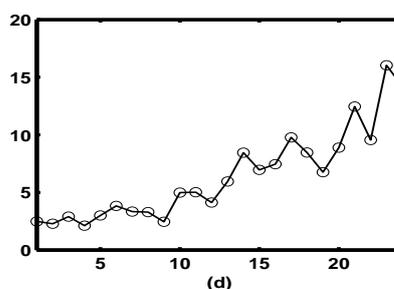
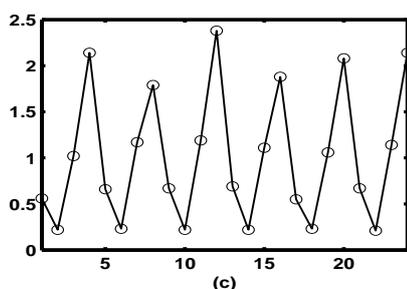
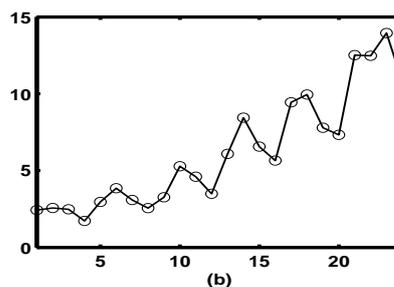
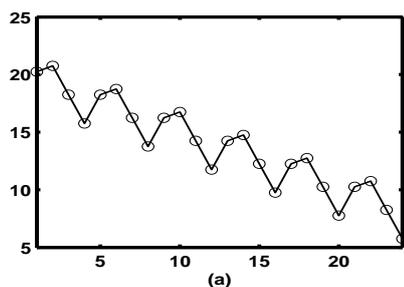
Il s'agit là des modèles où addition et multiplication sont utilisées. On peut supposer, par exemple, que la composante saisonnière agit de façon multiplicative, alors que les fluctuations irrégulières sont additives :

$$y_i = f_i(1 + s_i) + e_i, \quad 1 \leq i \leq n. \quad (5)$$

(toutes les autres combinaisons sont également possibles ...).

La modélisation stochastique des séries temporelles commence en observant leur graphique et en cherchant une décomposition additive ou multiplicative.

Exercice 2.1 Pour chacune des quatre séries suivantes,



- écrire un modèle qui vous semble convenir, en précisant le type du modèle, la tendance et la période
- Exprimez le modèle choisi sous la forme d'une équation vectorielle **linéaire** dans les paramètres inconnues, et donnez la formule de la régression qui permet à déterminer ces paramètres.

Nous étudierons dans la suite surtout le modèle additif (le modèle multiplicatif revient à un modèle additif pour le log des données).

Remarque 2.1 Une fois un modèle est obtenu, il peut être utilisé pour la prédiction des valeurs futurs.

En examinant des graphes des séries temporelles comme ci-dessus, les pères fondateurs du domaine (Wiener, Kolmogorov, ...) ont eu l'intuition que le souci de dégager les composantes périodiques et la tendance était similaire au problème d'enlèvement du bruit dans les appareils de radio, qui avait été résolu par l'utilisation des filtres ("low-pass" et "high-pass"). En effet, les prochains chapitres montreront les multiples utilisations des filtres/moyennes mobiles pour la prédiction des séries temporelles.

2.3 Recurrences et opérateurs de retard et avance

Nous venons de voir que les composantes périodiques (1) et périodiques de somme nulle (3) satisfont des recurrences à coefficients constants, et par conséquent ont des formules explicites comme sommes des puissances de racines de l'unité.

Pour manipuler les recurrences, il est convenable d'introduire un **opérateur de retard** B (ou encore de "rétro-décalage") défini sur l'ensemble des suites par

$$BX_t = X_{t-1} \quad \text{donc } B^i X_t = X_{t-i}, \quad \text{et finalement } \sum_{i=-k_1}^{k_2} \theta_i X_{t-i} = \theta(B)X_t$$

où $\theta(B)$ dénote la somme des puissances $\sum_{i=-k_1}^{k_2} \theta_i B^i$. Alternativement, on peut travailler avec $F = B^{-1}$.

Les équations (1), (3) deviennent

$$(I - B^p)\mathbf{s} = 0, (I + B + B^2 + \dots + B^{p-1})\mathbf{s} = 0$$

Dans cette écriture, on voit immédiatement que la deuxième équation implique la première (en multipliant par $(I - B)$, et donc les périodiques de somme nulle sont périodiques.

3 Filtrage/moyennes mobiles

Souvent il semble une bonne idée de remplacer l'information locale fournie par un "pixel" d'une série/champs aléatoire par l'information obtenue en ajoutant ses voisins, ce qui suggère de construire des "moyennes mobiles".

Définition 3.1 La série Y_t s'appelle une **moyenne mobile** de X_t ou **filtre** si

$$Y_t = \sum_{i=-k_1}^{k_2} \theta_i X_{t-i} \tag{6}$$

où $k_1, k_2 \geq 0$. L'ordre du filtre est $k = k_1 + k_2$.

Remarque 3.1 La notation des polynômes de retard ramène (6) à la forme :

$$Y_t = \theta(B)X_t$$

Un filtre est supposé réduire le bruit, mais laisser invariants des tendances constantes ou polynomiales.

Le cas des θ_i à somme égale à 1 s'appelle **lissage**, celui avec des θ_i égaux s'appelle **moyenne arithmétique**, et le cas d'une moyenne arithmétique avec $k_1 = k_2 = q$ sera appelé **moyenne arithmétique symétrique ou centré**.

$\theta(B)$ est appelé **opérateur de récurrence** (sur l'espace des séries). La série de Laurent associée $\theta(z) = \sum_{i=-k_1}^{k_2} \theta_i z^i$ sera appelée le symbole de l'opérateur.

Définition 3.2 a) Le symbole d'un filtre $\theta(B)$ est la fonction $\theta(z) : \mathbb{C}^- \rightarrow \mathbb{C}$.

b) En écrivant le filtre comme

$$\theta(z) = z^{-k_1} p(z)$$

ou $p(z)$ est un polynôme on obtient le **polynôme caractéristique**, et en renversant les coefficients du polynôme caractéristique on obtient le **polynôme dual**.

3.1 Filtres qui enlèvent les composantes périodiques et saisonnières

Définition 3.3 a) Une série s_t sera appelée périodique de période p ssi

$$s_{t+p} = s_t \iff (1 - B^p)s_t = 0, \quad \forall t \quad (7)$$

b) Une série s_t sera appelée saisonnière de période p ssi

$$\sum_{i=1}^p s_{t+i} = 0 \iff \left(\sum_{i=0}^{p-1} B^i \right) s_t = 0, \quad \forall t \quad (8)$$

Exercice 3.1 1. Donnez une formule explicite décrivant une série arbitraire appartenant à l'espace d'états vectoriel des séries périodiques. Quelle est la dimension de cet espace ?

2. Même questions pour une série arbitraire appartenant à l'espace d'états vectoriel des séries saisonnières.
3. Montrez qu'une série saisonnière est périodique, et que chaque série périodique p_t est la somme d'une série saisonnière et d'une série constante.
4. Trouvez une base de l'espace vectoriel des séries périodiques d'ordre p .
5. Trouvez une base réelle de l'espace vectoriel des séries saisonnières d'ordre p , et ensuite une base de l'espace des séries périodiques, qui la contient.
6. Formulez en forme matricielle un problème de régression (moindre carrés) qui permet de décomposer une série comme une somme d'une tendance linéaire, une partie saisonnière de période 4, et des résidus.

Nous allons identifier maintenant les filtres, ou opérateurs de récurrence qui "enlèvent/annulent" les composantes périodiques et saisonnières. Remarquons qu'un opérateur de récurrence qui annule une suite caractérise la suite en tant que combinaison des puissances des racines du polynôme caractéristique, ce qui rend encore plus évidente l'importance d'identifier des tels opérateurs.

Exercice 3.2 Montrez qu'un filtre $\theta(z)$ qui est divisible par $1 + z + \dots + z^{p-1}$, i.e. de la forme $\theta(z) = (1 + z + \dots + z^{p-1})\theta_1(z)$, "enlèvent" les composantes saisonnières de période p , i.e. :

$$\theta(B)s(t) = 0 \quad \forall t$$

pour chaque série s_t satisfaisant (8).

En effet, la réciproque est aussi vraie (admis) :

Théorème 3.1 Un filtre $\theta(B)$ annule (ou enlève) les composantes saisonnières d'ordre p ssi son symbole $\theta(z)$ est divisible par $1 + z + \dots + z^{p-1}$ (donc si $\theta(z) = 0$, pour toutes les racine d'ordre p de l'unité, sauf $z = 1$).

Exemples : Pour enlever les composantes saisonnières d'ordre 4, on peut utiliser donc la moyenne mobile arithmétique d'ordre 4, pour une périodicité mensuelle on peut utiliser la moyenne mobile arithmétique d'ordre 12, etc... En général, en utilisant un filtre arithmétique d'ordre p on peut enlever la partie saisonnière de cet ordre, pour mieux deceler ensuite la tendance.

3.2 Filtres de lissage

Un filtre de lissage (à $\sum_i \theta_i = 1$)

$$Y_t = \sum_{i=1}^k \theta_i X_{t-i} := \hat{X}_t$$

peut être utilisé pour la prédiction de X_t . Remarquez que le fait que la prédiction est "non-biaisée pour les séries stationnaires", i.e. :

$$\mathbb{E}\hat{X}_t = \mathbb{E} \sum_{i=1}^k \theta_i X_{t-i} = \left(\sum_{i=1}^k \theta_i \right) \mathbb{E}X_1$$

est assuré par la condition $\sum_{i=1}^k \theta_i = 1$.

Cette condition assure aussi qu'une série égale à $\mathbf{1}$ sera "prédite" exactement, i.e. $\theta(B)\mathbf{1} = \mathbf{1}$, et en fait chaque série constante $X_t = c$ sera prédite exactement :

$$\theta(B)c = c \cdot (\theta(B)\mathbf{1}) = c \cdot \mathbf{1} = c$$

Il est possible en fait, en choisissant les coefficients θ_i d'un filtre, d'assurer qu'il laisse invariants toutes les séries polynomiales p_t d'un degré donné. Le théorème suivant nous donne un critère pour identifier le degré maximal des polynômes laissés invariants par un filtre $\theta(B)$; autrement dit, de déterminer le degré maximal des polynômes inclus dans l'espace invariant des séries Z_t satisfaisant $\theta(B)Z_t = Z_t$:

Théorème 3.2 L'espace invariant d'un filtre contient les polynômes de degré $\leq n$ ssi

a) 1 est une racine d'ordre au moins $n + 1$ de l'équation $\theta(z) = 1$, i.e. $\theta(1) = 1, \theta'(1) = 0, \theta''(1) = 0, \dots, \theta^{(n)}(1) = 0$, où $\theta(z)$ est le symbole

b) 1 est une racine d'ordre au moins $n + 1$ de l'équation $p(z) - z^{k_1} = 0$, i.e. $p(1) = 1, p'(1) = k_1, p''(1) = k_1(k_1 - 1), \dots, \theta^{(n)}(1) = (k_1)_n$.

Exercice 3.3 *Démontrez le théorème*

Exercice 3.4 a) *Montrez qu'une moyenne arithmétique symétrique d'ordre $2q + 1 = 3$, donnée par*

$$\theta(B) = \frac{1}{3}(1 + B + B^{-1})$$

conserve (laisse invariante) les tendances linéaires $p_t = a + bt$. b) Généraliser pour q quelconque.

Après le choix d'une forme appropriée pour la tendance et une période spécifique, selon le graphe, on peut déterminer au même temps les coefficients de la tendance et de la partie périodique par une régression linéaire.

Exercice 3.5 *Montrez que le filtre $\frac{1}{9}(-B^2 + 4B + 3 + 4B^{-1} - B^{-2})$ laisse invariants les polynômes de troisième degré, et enlève les composantes saisonnières d'ordre 3.*

3.3 La réduction du bruit des observations

Nous verrons maintenant un résultat désirable de l'application des filtres de lissage : la réduction de la variance des observations.

Exercice 3.6 *Montrez qu'une moyenne arithmétique symétrique d'ordre $2q + 1$ diminue la variance σ^2 d'un bruit blanc (=série i.i.d. de moyenne 0) par $2q + 1$.*

En conclusion, si la série observée est de la forme

$$X_t = p_t + \epsilon_t$$

où $p_t = a + bt$ est une tendance linéaire, que l'opération de prendre une moyenne arithmétique symétrique d'ordre q n'affecte pas la tendance, i.e. $\theta(B)p_t = p_t$, mais a un effet de diminution du bruit stochastique ϵ_t , ramenant à :

$$\hat{X}_t = \theta(B)(p_t + \epsilon_t) = p_t + (\theta(B)\epsilon_t) = p_t + \frac{\epsilon_{t+q} + \dots + \epsilon_t + \dots + \epsilon_{t-q}}{2q + 1} := p_t + \epsilon'_t$$

avec un nouveau bruit $\epsilon'_t = \frac{\epsilon_{t+q} + \dots + \epsilon_t + \dots + \epsilon_{t-q}}{2q + 1}$ de variance inférieure à celle de ϵ_t .

Donc, si on constate une tendance linéaire dans le comportement d'une chronique dans un voisinage, on peut estimer la tendance dans ce voisinage en prenant des moyennes mobiles arithmétiques symétriques, car ça va réduire (atténuer) le bruit et mettre en évidence la tendance linéaire. L'effet du lissage augmente en augmentant q .

Exercice 3.7 *Montrez que la droite obtenue en lissant $2q + 2$ observations avec des moyennes mobiles arithmétiques symétriques d'ordre $2q + 1$ est :*

$$y - \frac{\sum_{i=1}^{2q+1} X_i}{2q + 1} = (x - (q + 1)) \frac{X_{2q+2} - X_1}{2q + 1}$$

3.4 A quoi servent les filtres ?

1. Les filtres causaux

$$\theta(B) = \sum_{i=0}^k \theta_i B^i$$

sont importants dans la prediction

2. Les filtres symétriques

$$\theta(B) = \sum_{i=-k}^k \theta_i B^i$$

sont importants dans le lissage (l'enlèvement du bruit). Les **filtres arithmétiques**

$$\theta(B) = (2k + 1)^{-1} \sum_{i=-k}^k B^i$$

minimisent le plus la variance (donc enlève bien le bruit).

Aussi, les filtres fournissent un langage adéquat pour décrire les séries temporelles. Rapellons que l'approche de décomposition additive suggère de rompre une série donnée

$$Y_t = s_t + m_t + \epsilon_t$$

dans un "signal" $s_t + m_t$ et un bruit stationnaire ϵ_t . La partie la plus importante d'une série est bien sur le signal, et son depistage peut se faire par l'approche classique de regression. Une fois depisté, la prediction du signal est evidente (par sa formule!).

Par contre, la prediction du bruit (corrélé) demande des approches plus sophistiquées, comme la modelisation ARMA, où l'utilisation des filtres devienne essentielle.

Il est naturel d'essayer de depister le signal par un filtre de lissage causal $l(B)$, donc $m_t = l(B)Y_t$, tq ce filtre "detruit le bruit mais laisse passer le signal". Il en suit que le bruit aussi est donné par un filtre causal

$$\epsilon_t = Y_t - m_t = (I - l(B))Y_t := \pi(B)Y_t$$

où $\pi(B)$ "detruit le signal mais laisse passer le bruit".

On peut voir donc l'analyse des séries temporelles comme la recherche pour un filtre qui transformera notre série en bruit. Pour cela, il va être important de définir des tests pour décider quand une série est un bruit. Nous allons examiner plus tard des statistiques comme les corrélations, corrélations partielles, le periodogramme, etc., issues de la modelisation probabiliste des séries temporelles.

Pour l'instant, les prochains paragraphes nous donnent quelques outils pour juger le comportement des filtres appliqués aux séries déterministes.

3.5 Exercices : TD 1

1. Trouvez l'élément suivant des séries y_t ci-dessous, ainsi que des équations de récurrences qu'elles satisfont et leurs solutions analytiques :

$$2, 6, 12, 20, 30, 42, \dots \quad (9)$$

$$4, 10, 20, 36, 62, 104, \dots \quad (10)$$

$$3, 2, 1, 6, 3, 2, 1, \dots \quad (11)$$

$$0, -1, -2, 3, 0, -1, -2, \dots \quad (12)$$

Indication. a), b) Calculez les séries différenciées : $z_t = \Delta y_t = y_t - y_{t-1}$. La deuxième série admet deux continuations naturelles (au moins).

2. Une série vérifie la récurrence $y_t - 2y_{t-1} + y_{t-2} = (-1)^{t-1}, t \geq 2, y_0 = 1, y_1 = 3$ Obtenez une formule analytique pour y_t , a) en éliminant le terme inhomogène

Sol : On arrive à $(B + 1)(B - 1)^2 y_t = 0$ La solution particulière doit être de la forme $y_t = A(-1)^{t-1}$. Comme $(B - 1)^2 (-1)^{t-1} = -4$, il suit que $A = \frac{-1}{4}$.

b) (*) en utilisant la méthode des fonctions génératrices, décomposition en fractions partielles, et l'expansion en série des puissances :

$$\frac{1}{(a - z)^k} = \sum_{n=0}^{\infty} C_{k-1}^{n+k-1} \frac{z^n}{a^{n+1}}, \quad |z| \leq a$$

Sol : $a(z) = \frac{1+2z}{(1+z)(1-z)^2} = \frac{3}{2(1-z)^2} - \frac{1/4}{(1-z)} - \frac{1/4}{(1+z)}, a(n) = (5 - (-1)^n + 6n)/4$

3. (a) Montrez que le filtre $P(B) = \frac{1}{3}(2 + B + B^2 - B^3)$ "enlève" les composantes saisonnières de période 3, i.e. qu'il transforme chaque fonction de période 3 dans une fonction constante.
 (b) Trouvez l'ordre de la tendance polynômiale maximale conservée (laissée invariante) par ce filtre.

Sol : Ordre 1.

4. Trouver un filtre $1 + \alpha B + \beta B^2 + \gamma B^3$ qui laisse passer un tendance affine sans distortion et élimine les périodicités d'ordre 2. Indication : Trouver un système des $2 + 1 = 3$ équations et résoudre.
 5. Trouvez un filtre $f(B)$ qui conserve les polynômes de degré ≤ 1 , et qui enlève les composantes saisonnières d'ordre 4, et déduisez que pour une série ayant une composante périodique d'ordre 4 et une tendance linéaire m_t , la tendance est donné par $m_t = f(B)Y_t$.

Sol : $\frac{1+B+B^2+B^3}{4} \frac{5-3B}{2}$

6. On considère la série suivante :

t_i	1	2	3	4	5	6	7	8	9	10
y_i	58	40	31	15	18	15	9	9	10	8

a) Représenter graphiquement cette série.

b) On se propose d'ajuster une tendance f de la forme $f(t) = \frac{1}{a + bt}$. Justifier ce choix.

- c) Déterminer les coefficients a et b , en utilisant un changement de variable approprié :
 - par la méthode des 2 points (en choisissant judicieusement les 2 points)
 - par la régression linéaire.

d) représenter les 2 tendances ainsi obtenues sur le graphique précédent et comparer les résultats. Est-ce que les résidus ont une allure irrégulière ?

7. On considère la série suivante

t_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_i	7.5	4.4	3.3	7.6	3.9	2.4	6.9	4.5	2.7	8.2	4.1	3.0	7.5	3.5	2.8

- Représenter graphiquement cette série.
 - Quel modèle proposeriez-vous pour cette série (justifier) ?
 - Calculer les facteurs saisonniers ($s_j, 1 \leq j \leq p$) ainsi que leur moyenne $p^{-1} \sum_{j=1}^p s_j$, en supposant une tendance constante $m_t = a$.
 - En notant ($e_i, 1 \leq i \leq n$) la série des fluctuations irrégulières, calculer e_1, e_2 et e_3 .
 - Proposer une méthode pour l'estimation des paramètres, en supposant cette fois une tendance linéaire $m_t = at + b$. Implementez le calcul en utilisant un logiciel. Proposez un teste pour choisir entre les deux modèles.
8. On considère un modèle simple où la tendance est une constante ($f(t) = a$).
- On considère tout d'abord le modèle sans composante saisonnière. Comment choisir a si le modèle est additif ? que peut-on alors dire sur les fluctuations irrégulières ? que se passe-t-il si le modèle est multiplicatif ?
 - On considère maintenant qu'une composante saisonnière ($s_j, 1 \leq j \leq p$) est présente. On suppose que le nombre d'observations n est un nombre entier L de périodes : $n = Lp$. Comment choisir a et (s_j) si le modèle est additif ? que peut-on alors dire sur les fluctuations irrégulières ? que se passe-t-il si le modèle est multiplicatif ?
 - * Reprendre la question b) lorsque le nombre d'observations n'est pas un nombre entier de périodes : $n = Lp + m$.
9. On considère une série ($y_i, 1 \leq i \leq n$) périodique, de période p . On suppose que le nombre d'observations n est un multiple de p : $n = Lp$. Montrer alors que les corrélations suivantes sont :

$$\rho(p) = \frac{L-1}{L} ; \rho(2p) = \frac{L-2}{L} ; \dots ; \rho(jp) = \frac{L-j}{L} \dots$$

4 Modélisation stochastique des séries temporelles

4.1 Introduction

Rappelons le modèle additif sans saisonnalité, qui cherche une décomposition de la forme :

$$Y_t = m_t + \epsilon_t \quad \text{où :}$$

- m_t représente la "tendance" (intuitivement un "mouvement lisse à long terme"), qui sera la composante la plus importante dans la prévision.
- $\epsilon_t = Y_t - m_t$ sont les "résidus" qui restent après qu'on enlève la partie structurée m_t . Elles représentent des "irrégularités/fluctuations imprévisibles", qui au début semblent inutilisables (à ignorer) pour la prévision (c'est correct du point de vue de la prévision ponctuelle, mais elles nous servent quand-même dans le calcul des intervalles de confiance).

On s'arrangera toujours tel que les résidus ont la moyenne 0, mais ça n'est pas suffisant pour qu'ils soient un bruit totalement sans structure="bruit blanc" (et s'il y a encore une partie structurée, elle devrait être incluse en m_t).

Le "bruit blanc" est notre premier exemple d'un processus stochastique : une formalisation du concept de séries temporelles, ayant des propriétés bien définies (voir prochaine chapitre). Inspirés par les propriétés de ce processus, on proposera des tests statistiques correspondant à ce modèle, qui nous permettront de décider si ϵ_t ont les propriétés de manque de structure désirées.

Pour tendance, plusieurs modèles se sont avérés utiles :

1. régression sur des prédicteurs exogènes ("covariates"), implémenté en logiciels comme R par "formules" :

$$m_t \sim X_t^{(1)} + X_t^{(2)} + \dots$$

2. modèles de superposition des chocs extérieurs/moyennes mobiles/FIR **inobservables** ϵ_t :

$$m_t = \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

3. modèles autoregressifs :

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots) + \epsilon_t$$

Dans le manque des prédicteurs exogènes, il est assez naturel d'adopter une modélisation autoregressive pour la tendance. Sous certaines conditions de régularité, ça ramènera à des prévisions autoregressives un pas en avant :

$$\hat{Y}_t = f(Y_{t-1}, Y_{t-2}, \dots)^1$$

Le modèle le plus simple est le processus AR(1) :

$$Y_t = \phi Y_{t-1} + b + \epsilon_t$$

Ce modèle est recommandable si on envisage une prévision

$$\hat{Y}_t = \phi Y_{t-1} + b \iff (\hat{Y}_t - a) = \phi(Y_{t-1} - a)$$

où $b = a(1 - \phi)$.

On vérifie que si la moyenne de Y_t est 0 on a $a = b = 0$; pour simplifier, on supposera normalement qu'on a déjà enlevé la moyenne de Y_t .

Pour utiliser ce modèle, on estime le paramètre ϕ par une régression linéaire des points

$$(Y_{t-1}, Y_{t-1}), t = 2, \dots, T$$

Le fait d'avoir enlevé la moyenne ramène à une droite passant par l'origine $y = \phi x$.

En suite, on utilise la valeur trouvée pour résoudre l'équation. On trouve

$$Y_t = \sum_{i=0}^{t-1} \phi^i \epsilon_{t-i} + \phi^t Y_0$$

1. La modélisation autoregressive permettra aussi des prédictions k pas en avant :

$$\hat{Y}_{t+k} = f_k(Y_t, Y_{t-1}, Y_{t-2}, \dots), \quad k = 1, 2, \dots$$

Les valeurs (positives) de k correspondent au futur et doivent être extrapolées/prévues. La fonction de prévision f_k représente une projection de Y_{t+k} sur l'espace engendré par $Y_t, Y_{t-1}, Y_{t-2}, \dots$. Plusieurs choix sont possibles par exemple extrapolation à partir d'un ajustement/interpolation polynomiale ou par splines.

et examiner solution stationnaire unique ssi : $|\phi| < 1$.

Indication : vous pouvez le faire en calculant la solution (i) par des substitutions répétées ou (ii) en utilisant des opérateurs, en posant $Y_t = (1 - \phi B)^{-1} \epsilon_t$, et en développant la fraction comme une série de puissances en B . En suite, calculez les covariances, pour montrer la stationnarité.

b) Montrez que l'équation : (21) a une solution stationnaire unique, qui dépend seulement du **bruit futur** ssi : $|\phi| > 1$.

En conclusion

1. pour $|\phi| < 1$, l'équation : (21) a une solution stationnaire unique causale, qui dépend seulement du **bruit passé**.
2. pour $|\phi| > 1$, l'équation : (21) a une solution stationnaire unique, qui dépend seulement du **bruit futur**.

Pour tester la validité des modèles, proposés, il faut d'abord préciser rigoureusement les propriétés désirées des résidus ou des chocs en tant que processus stochastiques ; en plus, les modèles seront utiles seulement si on peut vérifier leur "stationnarité", i.e une certaine uniformité de structure par rapport au temps. Ça nous ramène à considérer les processus stochastiques stationnaires, les distributions jointes de quelles ne varient pas avec le temps.

4.2 Processus stochastiques stationnaires

Définition 4.1 Soit X un processus aléatoire indexé par $T = \mathbb{N}$ ou \mathbb{Z} . On dit que X est **stationnaire (strict)** si pour toute famille finie d'instants $t_1 \dots t_r \in T$ et tout entier s , les lois jointes de $(X_{t_1} \dots X_{t_r})$ et de $(X_{t_1+s} \dots X_{t_r+s})$ sont les mêmes.

Définition 4.2 Soit X un processus aléatoire indexé par $T = \mathbb{N}$ ou \mathbb{Z} . On dit que X est **stationnaire à l'ordre 2** si la moyenne $m(t)$ et la covariance $\Gamma(s, t)$ sont invariantes par translation dans le temps, i.e. si la moyenne est constante :

$$\mathbb{E}X_t = m_t = m, \forall t$$

et si la covariance/corrélation dépend seulement de l'écart de temps $k = t - s$, i.e. il existe une fonction d'une variable $\gamma(k)$, paire, telle que :

$$\text{Cov}(X_t, X_s) = C(t, s) = \gamma(t - s) = \gamma(k), \quad \forall k = -2, -1, 0, 1, 2, 3, \dots$$

Comme la plupart de series n'est observable qu'une seule fois, l'utilité du concept de distributions et covariances théoriques n'est pas évidente pour les applications. Par contre, on peut toujours calculer des distributions et covariances empiriques, et sous l'hypothèse de stationnarité, les moyennes empiriques convergent vers les théoriques.

Théorème 4.1 Pour un processus stationnaire, les covariances empiriques

$$\gamma_n(k) = (n - k)^{-1} \sum_{t=1}^{n-k} (X_t - m_n)(X_{t+k} - m_n)$$

estimées à partir de n observations convergent vers les covariances théoriques quand $n \rightarrow \infty$.

D'ici l'importance du concept de stationnarité, qui justifie l'estimation des modèles statistiques observables une seule fois (le cas souvent dans les séries temporelles et la géostatistique!) : ceci est faisable ssi on a la chance d'avoir à faire avec un processus stationnaire.

Remarques :

1. La plupart des séries ne sont pas stationnaires, mais on peut essayer quand-même de se ramener à ce cas par des transformations (logarithmes, Box-Cox, etc).
2. Pour un processus du second ordre, la stationnarité stricte implique la stationnarité au sens large (à l'ordre 2). La réciproque est fautive. Une suite Y de v.a. indépendantes de même moyenne et même variance est toujours stationnaire à l'ordre 2 ; mais si les Y_n n'ont pas tous la même loi, Y n'est pas stationnaire au sens strict.
3. (*) La stationnarité à l'ordre 2 est bien plus facile à étudier et vérifier que la stationnarité stricte. Son importance pratique tient surtout aux problèmes de prédiction ou de régression. En effet, on se limite souvent à des critères de moindres carrés pour avoir des estimateurs calculables. Cela signifie alors utiliser des prédicteurs linéaires optimaux dont le calcul ne fait pas intervenir dans sa totalité la structure probabiliste du processus X observé, mais seulement la géométrie (angles et longueurs) de la suite (X_k) considérée comme suite de vecteurs dans l'espace de Hilbert $L^2(\Omega, P)$. Or, cette géométrie ne dépend que des moments d'ordre 2 de X ; la notion naturelle de stationnarité est donc l'invariance de ces moments d'ordre 2 par translation dans le temps.

5 Exemples des processus stationnaires

L'idéalisation probabiliste de la propriété que les résidus sont "completement irreguliers", ne retenant aucune structure, est le "bruit blanc" stationnaire.

Un deuxième exemple important des processus stationnaires "non-blancs" sont les "processus linéaires" $MA(\infty)$ $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ avec ϵ_t bruit blanc et leur cas particulier avec un nombre fini des coefficients ψ_i nonnuls, les "moyennes mobiles" $MA(q)$.

Le troisième exemple étudié seront les "processus autorégressifs" $AR(\infty)$ $\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$ avec ϵ_t bruit blanc et leur cas particulier avec un nombre fini des coefficients π_i nonnuls, les processus $AR(p)$.

5.1 Le bruit blanc

L'exemple le plus simple de modèle stochastique est le bruit blanc discret, la structure "revée" des résidus qui restent après qu'on enlève la tendance/moyenne d'un processus.

Définition 5.1 *Un processus $\epsilon_t, t \in \mathcal{T}$, où \mathcal{T} est un ensemble dénombrable quelconque, est appelé **bruit blanc stationnaire** si les variables ϵ_t sont **i.i.d.** (indépendants et identiquement distribués) à espérance $\mathbb{E}\epsilon_t = 0$. Il sera appelé **bruit blanc Gaussien** si la distribution de chaque v.a. ϵ_t est Gaussiennes.*

Un bruit blanc a la covariance

$$\gamma(s, t) = \mathbb{E}[\epsilon_s \epsilon_t] = 0, \forall s \neq t \quad \text{et donc le coefficient de corrélation} \quad (13)$$

$$\rho(s, t) = \frac{\gamma(s, t)}{\sigma_s \sigma_t} = \delta(s - t) \quad (14)$$

où $\delta(s - t)$ est le symbole du Kronecker.

Comme les tests d'indépendance et Gaussianité demandent beaucoup de données, qui ne sont pas toujours disponibles, il faut faire parfois avec un "idéale probabiliste moins structuré" : le "bruit blanc de second ordre" défini par les deux dernières formules équivalentes (13), (14).

Définition 5.2 Un processus $\epsilon_t, t \in \mathbb{N}$ ou $t \in \mathbb{Z}$ est appelé **bruit blanc de second ordre** s'il a la moyenne 0, la variance constante $\mathbb{E}\epsilon_t^2 = \sigma^2$ et une covariance $\gamma(s, t) = \mathbb{E}[\epsilon_s \epsilon_t] = 0, \forall s \neq t$ (et donc les coefficients de corrélation $\rho(s, t) = \delta(s - t)$).

Notes :

1. Le bruit blanc Gaussien est une structure probabiliste très naturelle, car la distribution Gaussienne possède plusieurs propriétés importantes, comme celle d'être invariante par rapport aux rotations, ce qui est évidemment une réquise pour un bruit aléatoire.
2. Le bruit blanc stationnaire est une idéalisation du processus des résidus de la régression linéaire, qu'on aimerait "indépendants". Mais, comme l'indépendance est un concept probabiliste, et les résidus sont le résultat déterminist d'une régression appliqué a une série observée une seule fois, il est difficile de la vérifier rigoureusement. Parmi les tests possibles, mentionnent celui de "turning points", qui demande de vérifier que la fréquence de ces points est environ 4/6, et le teste qui vérifie si la somme des corrélations empiriques est proche de 0. Si ces deux testes sont positives, on sait au moins "qu'on ne peut pas repousser l'hypothèse de l'indépendance". Il y aussi des tests distributionnels des résidus comme Fisher, Student, qui testent la Gaussianité.
3. Quand les tests des données rejettent l'hypothèse du bruit blanc, i.e. quand on a du **bruit corrélé**, la régression classique doit être remplacé par une analyse plus fine, appelée krigeage en géostatistique.

5.2 Les moyennes mobiles MA(q) : l'exemple le plus simple de processus a composantes dépendantes

Définition 5.3 Un processus Y_t sera appelé linéaire en ϵ_t s'il peut être représenté dans la forme :

$$Y_t = \sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i} \text{ avec } \sum \psi_i^2 < \infty \quad (15)$$

où ϵ_t est un bruit blanc.

Evidemment, du point de vue pratique (pour la prédiction), on ne s'intéresse que dans le cas –qui sera appelé **causal**– quand la représentation n'utilise pas "le bruit du futur" :

Définition 5.4 Un processus linéaire Y_t s'appelle **causal** s'il peut être représenté dans la forme :

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \quad (16)$$

où ϵ_t est un bruit blanc et $\sum \psi_i^2 < \infty$

Définition 5.5 On appelle processus MA(q) un processus linéaire $Z_t, t \in \mathbb{Z}$ vérifiant une relation :

$$Z_t = \sum_{i=0}^q \theta_i \epsilon_{t-i}, \forall t \in \mathbb{Z} \quad (17)$$

où ϵ_t est un bruit blanc de variance σ^2 et $\theta_0 = 1$.

La notation des polynômes de retard ramène (17) à la forme :

$$Z_t = \theta(B)\epsilon_t$$

Théorème 5.1 Un processus linéaire

$$Y_t = \sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i}$$

où $\sum \psi_i^2 < \infty$ est : a) bien défini dans L_2 (i.e. $\text{Var } Y_t < \infty$),
 b) à variance constante stationnaire $\text{Var } Y_t = \sigma_\epsilon^2 \sum_{i=-\infty}^{\infty} \psi_i^2$
 c) à autocovariance donnée par :

$$\gamma(t, t+k) = \sigma_\epsilon^2 \sum_{i=-\infty}^{\infty} \psi_i \psi_{i+k} < \infty \quad (18)$$

d) stationnaire à ordre deux.

Démonstration : a) En considérant $\text{Var}(Y_t)$, on voit que la condition est nécessaire et suffisante pour convergence. b),c),d) En suite, on voit qu'elle suffit pour la stationnarité, car elle assure que $\text{Cov}(Y_t, Y_{t+k})$ est bien définie par l'inégalité de Cauchy-Schwartz (qui est équivalente à $|\rho_k| \leq 1$) et ne dépend pas de t .

Exercice 5.1 1. Calculer la fonction d'autocovariance $\gamma(k)$ d'un processus MA(1).
 2. Calculer la fonction de covariance $\gamma(k)$ d'un processus MA(q)

Le fait que les fonctions de covariance et corrélation $\gamma(k)$, respectivement $\rho(k)$ d'un processus MA(q) s'annulent pour $k > q$ permet de reconnaître des séries qui peuvent être modélisées comme MA(q). Plus précisément, pour accepter l'hypothèse qu'une série est MA(q) pour un q donné, on vérifie que toutes les corrélations pour $k > q$ satisfont

$$|\rho_n(k)| \leq z_\alpha \sigma_q$$

où

$$\sigma_q^2 = \frac{1 + 2(\hat{\rho}(1))^2 + \hat{\rho}(2)^2 + \dots + \hat{\rho}(q)^2}{n}$$

(formule de Bartlett) et z_α , la "fractile" d'ordre α de la distribution Gaussienne, dépend du niveau de confiance α désiré (par exemple, $z_{.95} = 2$). Donc, si toutes les corrélations pour $k > q$ sont à l'intérieur de cette bande de confiance, on accepte l'hypothèse que la série est MA(q).

5.3 Les modèles autorégressifs AR(p) : des processus faciles de prédire

La prédiction d'une série est particulièrement simple quand elle peut être "bien approximée" par un modèle **autorégressif** paramétrique :

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots) + \epsilon_t \quad (19)$$

Dans ce cas il s'avère typiquement que la formule de prévision ponctuelle pour Y_t un pas en avant est simplement :

$$\hat{Y}_t = f(Y_{t-1}, Y_{t-2}, \dots)$$

Nous allons considérer ici surtout des modèles autorégressifs linéaires (où f est une fonction linéaire) AR(p) :

Définition 5.6 *Un processus stationnaire $Y_t, t \in \mathbb{Z}$ sera appelé **processus autorégressif linéaire d'ordre p : AR(p)** s'il existe un bruit blanc ϵ_t et des réels $\varphi_i, i = 1, \dots, p$ tels qu'une relation de récurrence :*

$$Y_t = \sum_{i=1}^p \varphi_i Y_{t-i} + \epsilon_t, \forall t \in \mathbb{Z} \quad (20)$$

est vérifiée.

La notation des polynômes de retard ramène (20) à la forme :

$$\varphi(B)Y_t = \epsilon_t \quad \text{où } \varphi(B) = 1 - \sum_{i=1}^p \varphi_i B^i$$

Définition 5.7 *Le polynôme*

$$\varphi(B) = 1 - \sum_{i=1}^p \varphi_i B^i$$

*sera appelé **polynôme caractéristique**, ou **symbole** du modèle (20).*

Rq : Les processus autorégressifs sont définis par une équation, qui à priori, peut ne pas avoir des solutions; comme "solution" de l'équation (20) nous aimerions avoir une représentation du processus Y_t par rapport au processus ϵ_t .

Exercice 5.2 processus AR(1). a) Montrez que l'équation :

$$Y_t = \phi Y_{t-1} + \epsilon_t \quad (21)$$

a une solution stationnaire unique **causale** (qui dépend seulement du **bruit présent et passé**) si : $|\phi| < 1$.

Indication : Calculer la solution de (21) par des substitutions répétées, ou en utilisant le calcul opérationnel de Heaviside (Arbogast), en posant $Y_t = (1 - \phi B)^{-1} \epsilon_t$, et en développant la fraction comme une série de puissances en B . En suite, montrez que la série est convergente dans le sens L_2 , et calculez les covariances, pour vérifier la stationnarité.

b) Montrez que l'équation : (21) a une solution stationnaire unique, qui dépend seulement du **bruit futur** si : $|\phi| > 1$.

Remarque 5.1 Dans le cas $|\phi| = 1$, les solutions de l'équation : (21) (appelées marches aléatoires) ne peuvent pas être stationnaires. Par contre, les **increments** $Y_t - Y_{t-1} = \epsilon_t$ sont stationnaires ; cette situation plus compliquée sera analysée dans le chapitre sur les processus ARIMA(p, d, q).

En conclusion

Théorème 5.2 Le processus (de Markov) AR(1) $Y_t = \phi Y_{t-1} + \epsilon_t$ a une représentation causale ssi $|\phi| \neq 1$ et cette représentation est causale ssi $|\phi| < 1$. Aussi

1. pour $|\phi| < 1$, on vérifie que ϵ_t est un bruit d'innovation, i.e. $E[\epsilon_t Y_{t-k}] = 0, \forall k \geq 1$.
2. pour $|\phi| > 1$, l'équation : (21) a une solution stationnaire unique, qui dépend seulement du bruit futur, et ϵ_t n'est pas un bruit d'innovation.

6 Fonctions de transfert et applications

6.1 Causalité des modèles AR(p)

Nous venons de voir qu'il y a des problèmes (manque de solutions stationnaires et non-causalité) avec le modèle AR(1) quand la racine $\lambda = \phi^{-1}$ de son polynôme $\varphi(z) = 1 - \phi z$ n'est pas à l'extérieur du cercle unitaire §.

On a la même situation de pour tous les modèles AR(p)

Théorème 6.1 a) Un processus AR(p) est causal, i.e. il peut être représenté sous la forme : $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ où $\sum \psi_i^2 < \infty$ ssi toutes les racines de son polynôme caractéristique $\varphi(z)$ sont à l'extérieur du cercle unitaire. Les coefficients ψ_i sont dans ce cas les coefficients de la série Taylor de $\psi(z) = \frac{1}{\varphi(z)}$

Définition 6.1 La fonction $\psi(z) = \frac{1}{\varphi(z)}$ intervenant dans la représentation

$$Y_t = \psi(D)\epsilon_t$$

est appelée fonction de transfert du modèle AR(p).

Remarque 6.1 La fonction de transfert permet de représenter le modèle AR(p) comme un modèle MA(∞).

Nous pouvons calculer les coefficients de la fonction de transfert par plusieurs méthodes. La première est l'identification des coefficients dans

$$\psi(z)\varphi(z) = 1$$

(utilisée aussi dans les approximations de Padé, etc...).

En appliquant cette méthode, on découvre une "récurrence fondamentale" à l'étude des processus ARMA.

§. Ce problème est lié à l'existence des plusieurs développements possibles pour la fonction $\varphi(z)^{-1} = \frac{1}{1-z\phi}$:

$$\frac{1}{1-z\phi} = \sum_{n=0}^{\infty} \phi^n z^n \quad \text{si } |\lambda| > 1, \text{ à l'intérieur du cercle } |z| \leq \lambda, \text{ mais}$$

$$\frac{1}{1-z\phi} = - \sum_{n=-\infty}^{-1} \phi^n z^n \quad \text{si } |\lambda| < 1, \text{ à l'extérieur du cercle unitaire, } |z| \geq \lambda$$

6.2 La récurrence de Yule-Walker pour les coefficients de la fonction de transfert

Exercice 6.1 Soit $\sum_{i=0}^{\infty} \psi_i z^i$ le développement limité autour de 0 de $\frac{1}{1-\varphi_1 z - \varphi_2 z^2}$.

1. Montrez que les coefficients ψ_i satisfont les équations Yule-Walker

$$\begin{aligned}\psi(0) &= 1, & \psi(1) &= \phi_1 \text{ (ou } \psi(-1) = 0) \\ \psi_i &= \phi_1 \psi_{i-1} + \phi_2 \psi_{i-2} \iff \varphi(B)\psi_i = 0, & i &\geq 2\end{aligned}$$

2. Donnez les formules explicites de $\psi_i, i = 0, 1, \dots, 3$.

Sol : $\{1, \phi(1), \phi(1)^2 + \phi(2), \phi(1)^3 + 2\phi(2)\phi(1)\}$.

3. Généraliser au développement limité autour de 0 de $\frac{1}{1-\varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p}$

Sol : Les solutions explicites se compliquent, mais la récurrence

$$\boxed{\varphi(B)\psi_i = 0, \quad i \geq p, \psi_0 = 1, \psi_{-1} = 0, \dots, \psi_{-p+1} = 0} \quad (22)$$

reste vraie.

Remarque 6.2 Il suit que si un processus $AR(p)$ $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \epsilon_t$ a une représentation stationnaire causale

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i},$$

(ce qui est le cas si le polynôme caractéristique $\varphi(z) = (1 - \lambda_1 z)(1 - \lambda_2 z) \dots (1 - \lambda_p z)$ a ses racines $\lambda_1^{-1}, \dots, \lambda_p^{-1}$ dehors le cercle unitaire, ou si le polynôme dual $\lambda^p \varphi(\lambda^{-1})$ a ses racines λ_i dedans le cercle unitaire), alors les coefficients ψ_i satisfont la **récurrence Yule-Walker** (22).

Exercice 6.2 Montrez que les coefficients $\psi_k = A_1 \lambda_1^k + A_2 \lambda_2^k$ d'un processus $AR(2)$ avec des racines distinctes sont donné par

$$\psi(k) = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2}.$$

Vérifier que dans le cas causal, $\sum_i \psi_i^2 \leq \infty$.

Exercice 6.3 (*) Trouvez les solutions stationnaires des processus $AR(2)$ avec au moins une des racines λ_1, λ_2 strictement dedans le cercle unitaire $|\lambda| < 1$.

6.3 Inversibilité des processus MA(q)

Définition 6.2 Une représentation causale

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$$

d'un processus stationnaire Y_t s'appelle **inversible** si on peut aussi représenter le bruit par une représentation causale :

$$\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i} \quad (23)$$

où $\sum \pi_i^2 < \infty$

Exemple 6.1 Le processus MA(1) $Y_t = \epsilon_t + \theta\epsilon_{t-1}$ est inversible ssi $|\theta| < 1$. En effet, comme dans la résolution de la récurrence AR(1), on voit que :

$$\epsilon_t = Y_t - \theta Y_{t-1} + \dots (-\theta)^{t-1} Y_1 + (-\theta)^t \epsilon_0$$

Pour $\theta < 1$, ça converge vers $\sum_{i=0}^{\infty} \pi_i Y_{t-i}$, où $\pi_i = (-\theta)^i$.

Exercice 6.4 On considère le processus aléatoire MA(2) suivant :

$$X_t = \epsilon_t - \frac{1}{3}\epsilon_{t-1} + \frac{1}{4}\epsilon_{t-2} = 45 + \epsilon_t$$

ou ϵ_t est BB(0, $\sigma^2 = 1$).

1. Calculer la variance de X_t , $\sigma^2 = \mathbb{E}X_t$, et les covariances.
2. Proposer une représentation autoregressive de ce processus. Est-ce qu'elle est causale est convergente ?

Théorème 6.2 Un processus MA(q) avec les racines du polynôme caractéristique $\theta(z)$ à l'extérieur du cercle unitaire est inversible, i.e. le bruit peut être représenté sous la forme : $\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$ où $\sum |\pi_i| < \infty$. Les coefficients π_i sont dans ce cas les coefficients de la série Taylor de $\pi(z) = \frac{1}{\theta(z)}$ ²

Remarque 6.3 Donc, ϵ_t appartient à l'espace linéaire engendré par le passé du signal observé

$$\epsilon_t \in \text{sp}\{Y_{t-i}, i = 0, 1, \dots\}$$

et les espaces engendrés par $\{Y_{t-i}, i = 0, 1, \dots\}$ et $\{\epsilon_{t-i}, i = 0, 1, \dots\}$ coïncident.

6.4 Équations de Yule-Walker pour les covariances/corrélations des modèles autorégressifs AR(p)

Le corrélogramme des processus autorégressifs AR(p) n'est pas aussi facile à calculer que celle des processus MA(q).

Théorème 6.3 Les covariances et les corrélations d'un processus AR(p) sont liées aux coefficients $\varphi = (\varphi_1, \dots, \varphi_p)$ par les équations de Yule-Walker.

2. La démonstration est basée sur un théorème concernant le développement en série des puissances des inverses des polynômes complexes –voir la prochaine section– et finalement sur le fait que l'ensemble des filtres est isomorphe à l'ensemble des fonctions complexes $\psi(z)$, l'isomorphisme étant "la transformée z " des suites. Cet isomorphisme explique quelques manipulations formelles avec les filtres (mais pas celles liés à l'inversion).

1. Pour les covariances, en multipliant la formule autoregressive de Y_{t+k} par Y_t et en prenant l'espérance, on obtient :

$$\begin{aligned} \gamma_k &= \sum_{i=1}^p \varphi_i \gamma_{k-i} \quad \text{pour } k \geq 1 \\ \gamma_0 &= \sum_{i=1}^p \varphi_i \gamma_i + \mathbb{E}Y_t \epsilon_t = \sum_{i=1}^p \varphi_i \gamma_i + \sigma^2 \quad \text{pour } k = 0 \end{aligned} \quad (24)$$

2. Pour les corrélations, on remarque d'abord, en divisant par γ_0 , qu'elles satisfont aussi la récurrence (24) :

$$\rho_k = \sum_{i=1}^p \varphi_i \rho_{k-i} \quad \text{pour } k \geq 1 \quad (25)$$

Alors, il suffit de déterminer les premières p corrélations, ce qu'on fait en appliquant la récurrence (25) pour $k = 1, \dots, p$, en tenant compte aussi de la symétrie de ρ_k . En général, pour le processus $AR(p)$ on arrive ainsi au **système Yule-Walker** pour les premières p corrélations $\boldsymbol{\rho} = (\rho(1), \dots, \rho(p))$:

$$\boxed{\mathbf{R}_p \boldsymbol{\varphi} = \boldsymbol{\rho}} \quad (26)$$

où \mathbf{R}_p est la matrice Toeplitz symétrique :

$$\mathbf{R}_p = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \dots & \dots & \dots & \dots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix}$$

Les équations (26) (un cas particulier d'"équations normales" de la régression) permettent de calculer les coefficients $\boldsymbol{\varphi}$ à partir des corrélations et viceversa (on calcule d'abord les premières p corrélations; en suite, on trouve aussi les autres corrélations en utilisant la récurrence).

Remarque 6.4 En prenant γ_0 comme facteur commun dans la deuxième équation en (24), on trouve γ_0 en fonction des corrélations ρ_i :

$$\gamma_0 = \frac{\sigma^2}{1 - \sum_i \varphi_i \rho_i}$$

Ça permet en suite d'obtenir les autres covariances, en partant des corrélations.

Dem : b) Començons par le cas $p = 2$. Les premières deux équations pour $k = 1, 2$ sont

$$\begin{aligned} \rho_1 &= \varphi_1 + \varphi_2 \rho_{-1} = \varphi_1 + \varphi_2 \rho_1 \\ \rho_2 &= \varphi_1 \rho_1 + \varphi_2 \end{aligned}$$

et la première donne $\rho_1 = \frac{\varphi_1}{1-\varphi_2}$. Pour trouver les φ_i à partir de ρ_i , on écrit le système comme :

$$\begin{cases} \varphi_1 + \varphi_2 \rho_1 = \rho_1 \\ \varphi_1 \rho_1 + \varphi_2 = \rho_2 \end{cases} \iff \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}$$

Considerons le cas $p = 3$. Les premières trois équations pour $k = 1, 2, 3$ sont

$$\begin{cases} \varphi_1 + \varphi_2 \rho_1 + \varphi_3 \rho_2 = \rho_1 \\ \varphi_1 \rho_1 + \varphi_2 + \varphi_3 \rho_1 = \rho_2 \\ \varphi_1 \rho_2 + \varphi_2 \rho_1 + \varphi_3 = \rho_3 \end{cases} \iff \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix}$$

et on voit que la structure Toeplitz précisée dans le théorème est correcte.

Remarque 6.5 Les matrices de corrélations \mathbf{R}_p sont positive définies et jouent un rôle fondamental dans l'estimation des modèles $AR(p)$ et $ARMA(p, q)$.

Finalement, il est intéressant de comparer trois solutions possibles pour obtenir les corrélations :

1. en résolvant le système et récurrence de Yule-Walker
2. par le développement limité de la fonction génératrice $f(z)$
3. en représentant d'abord le processus $AR(p)$ comme un processus $MA(\infty)$, par le développement limité $\psi(z) = 1/\varphi(z)$, et en utilisant en suite la formule (18) du corrélogramme des processus $MA(\infty)$.

Exercice 6.5 Calculer le corrélogramme de $\mathbf{AR}(1)$ avec $\sigma_\epsilon^2 = 1$ par :

1. les équations Yule-Walker
2. par le développement limité de la fonction génératrice $f(z)$
3. la représentation $MA(\infty)$.

Tracez la corrélogramme $\{\rho_k\}$ pour

(a) $\varphi = 0.5$, (b) $\varphi = 0$, and (c) $\varphi = -0.5$.

R : 1. Les équations Yule-Walker sont :

$$\rho_n = \varphi \rho_{n-1}, n \geq 1 \implies \rho_n = \varphi^n$$

et $\gamma_0 = \frac{1}{1-\varphi^2}$.

2. Décomposant en fractions simples :

$$\begin{aligned} f(z) &= \frac{z}{(1-\varphi z)(z-\varphi)} = \frac{1}{1-\varphi^2} \left(\frac{1}{1-z\varphi} + \frac{\varphi}{z-\varphi} \right) = \frac{1}{1-\varphi^2} \left(\frac{1}{1-z\varphi} + \frac{\varphi}{z} \frac{1}{1-\varphi/z} \right) \\ &= \frac{1}{1-\varphi^2} \left(\sum_{k=0}^{\infty} z^k \varphi^k + \frac{\varphi}{z} \left(\sum_{k=0}^{\infty} z^{-k} \varphi^k \right) \right) = \frac{1}{1-\varphi^2} \left(\sum_{k=0}^{\infty} z^k \varphi^k + \sum_{k=1}^{\infty} z^{-k} \varphi^k \right) \end{aligned}$$

Théorème 6.4 Formule générale des corrélations pour $AR(2)$: Pour le processus $AR(2)$

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \epsilon_t$$

le système de Yule-Walker (26) donne :

$$\rho_1 = \frac{\varphi_1}{1-\varphi_2}, \quad \rho_2 = \varphi_2 + \frac{\varphi_1^2}{1-\varphi_2} = \varphi_2 + \rho_1 \varphi_1 = \varphi_2 + \rho_1 \psi_1$$

(ce qui implique la restriction $\rho_1^2 < \frac{1}{2(1+\rho_2)}$).

1. Si les racines λ_1, λ_2 de $0 = \lambda^2 - \phi_1\lambda - \phi_2 = \lambda^2\varphi(\lambda^{-1})$ (qui sont dedans le cercle unitaire, par la causalité) sont distinctes, on obtient

$$\rho_k = \frac{\rho_1 - \lambda_2}{\lambda_1 - \lambda_2} \lambda_1^k + \frac{\lambda_1 - \rho_1}{\lambda_1 - \lambda_2} \lambda_2^k = \rho_1\psi_{k-1} + \varphi_2\psi_{k-2}, k \geq 1$$

Finalement, en utilisant $\varphi_1 = \lambda_1 + \lambda_2, \varphi_2 = -\lambda_1\lambda_2$, on arrive à

$$\rho_k = \frac{(1 - \lambda_2^2)\lambda_1^{k+1} - (1 - \lambda_1^2)\lambda_2^{k+1}}{(\lambda_1 - \lambda_2)(1 + \lambda_1\lambda_2)}, \quad k \geq 0.$$

termes de racines seulement.

2. Dans le cas de racines confondues $\lambda_i = \lambda$, on obtient :

$$\rho_k = \left\{ 1 + \left(\frac{1 - \lambda^2}{1 + \lambda^2} \right) k \right\} \lambda^k, \quad k \geq 0.$$

Exercice 6.6 AR(2) :

Vérifiez si les processus AR(2) :

a) $Y_t = -0.5Y_{t-1} + 0.14Y_{t-2} + \epsilon_t$ et

b) $Y_t = -0.6Y_{t-2} + \epsilon_t$ sont stationnaires causals. Montrez, en partant directement du système de Yule-Walker que leur corrélogrammes sont :

a) $\rho_k = \frac{17}{129}(0.2)^k + \frac{112}{129}(-0.7)^k, \quad k = 0, 1, 2, \dots$ et

b) $\rho_k = \frac{1}{2}i^k(0.6)^{k/2}\{1 + (-1)^k\} = (0.6)^{k/2} \cos(k\pi/2), \quad k = 0, 1, 2, \dots$

Exercice 6.7 Tracez les corrélogrammes pour les processus AR(2) avec :

(i) $\varphi_1 = 0.2, \quad \varphi_2 = 0.35$ et (ii) $\varphi_1 = -0.8, \quad \varphi_2 = -0.16$

7 Contrôle continu d'entraînement

1. Trouvez les coefficients ψ_j de la représentation comme processus MA(∞) du processus AR(2) définie par :

$$Y_t - \frac{1}{2}Y_{t-1} - \frac{3}{16}Y_{t-2} = \epsilon_t$$

Est-ce que le processus est causal ?

2. Déterminer un filtre (une moyenne mobile) causal $\theta(B) = \sum_{i=0}^q \theta_i B^i$ d'ordre q minimal, qui laisse passer une tendance quadratique sans distortion et qui enlève les composantes saisonnières d'ordre 3.

3. Soit $X_t = \sum_{i=0}^k \theta_i \epsilon_{t-i}$ un processus moyenne mobile, utilisé pour lisser un bruit blanc $\epsilon_t = BB(0, \sigma^2 = 1)$.

a) Quelle est la variance de X_t ?

b) Trouvez le filtre (i.e. les coefficients $(\theta_0, \dots, \theta_k)$) qui ramènent à un processus X_t à variance minimale, parmi toutes les moyennes mobiles d'ordre k qui laissent passer une tendance constante sans distortion.

4. On considère le processus aléatoire suivant :

$$X_t - .2 X_{t-1} - .35 X_{t-2} = 45 + \epsilon_t$$

ou ϵ_t est $BB(0, \sigma^2 = 1)$.

- (a) Calculer l'espérance de X_t , $m = \mathbb{E}X_t$, en supposant que le processus est stationnaire.
- (b) Quelle est l'équation de récurrence satisfaite par le processus $Y_t = X_t - m$?
- (c) Est-ce que notre processus Y_t est stationnaire causal ? Si oui, donner les coefficients ψ_k de sa représentation causale $Y_t = \sum_k \psi_k \epsilon_{t-k}$.
- (d) Donner les prévisions $\hat{X}_t(k)$, en fonctions des racines λ_i du polynôme dual de Y_t (définies par $\varphi(B) = (1 - \lambda_1 B)(1 - \lambda_2 B)$).
- (e) Donner les équations de Yule-Walker pour les autocorrélations du processus Y_t , calculer les 3 premières autocorrélations, ainsi que la variance.
- (f) (*) Calculer les 3 premières autocorrélations partielles.

Solutions :

1. $1 - \frac{1}{2}z - \frac{3}{16}z^2 = (1 - \frac{3}{4}z)(1 + \frac{1}{4}z)$.

$\psi_k = (\frac{3}{4})^{k+1} - (-\frac{1}{4})^{k+1}$

Oui, le processus est causal, car les racines duales $\lambda_1 = \frac{3}{4}, \lambda_2 = -\frac{1}{4}$ sont dedans le cercle unitaire $|\lambda| \leq 1$.

2. En demandant que $\theta(B) = \frac{1+B+B^2}{3}(a_0 + a_1 B + a_2 B^2)$ satisfait $\theta(1) = 1, \theta'(1) = 0, \theta''(1) = 0$, on trouve : $\theta(B) = 1/9(8 - 7B + 2B^2)(1 + B + B^2) = 8/9 + B/9 + B^2/3 - 5B^2/9 + 2B^4/9$.

3. a) $\text{Var } X_t = \sum_i \theta_i^2$. b) Elle est minimisé sous la contrainte $\sum_i \theta_i = 1$ par la moyenne arithmétique à coefficients égaux.

4. (a) Ce processus AR(2) pas centré peut s'écrire $(1 - .2B - .35B^2)X_t = (1 - .7B)(1 + .5B)X_t = 45 + \epsilon_t$ En prenant l'espérance on trouve $\mathbb{E}(X_t) = \frac{45}{1 - .2 - .35} = \frac{40}{.45} = 100$.

(b) Soit $Y_t = X_t - \mathbb{E}X_t$. Alors Y_t est un processus AR(2) centré satisfaisant $(1 - .2B - .35B^2)Y_t = (1 - .7B)(1 + .5B)Y_t = \epsilon_t$

(c) Pour un processus stationnaire causal (avec les racines du polynôme $\varphi(B)$ à l'extérieur du disque unité), les coefficients $\psi(k)$ satisfont les 'équations de Yule Walker

$$\psi_k = \varphi_1 \psi_{k-1} + \varphi_2 \psi_{k-2}, \psi_{-1} = 0, \psi_0 = 1.$$

Avec des racines duales distinctes λ_1, λ_2 , la solution est :

$$\psi(k) = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} = \dots$$

(d) En résolvant la récurrence (31), où par (33), on trouve ...

(e) La fonction d'autocovariance de Y_t (qui est la même que celle de X_t) est obtenue de l'équation : $\mathbb{E}(Y_t Y_{t-h}) = 0.2\mathbb{E}(Y_{t-1} Y_{t-h}) + 0.35\mathbb{E}(Y_{t-2} Y_{t-h}) + \mathbb{E}(\epsilon_t Y_{t-h})$, ce qui donne l'équation Yule Walker

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2}$$

Pour $k = 1, 2$ on a le système de Yule Walker :

$$\rho_1 = 0.2 + 0.35\rho_1$$

$$\rho_2 = 0.2\rho_1 + 0.35$$

La première équation donne $\rho_1 = \frac{2}{6.5} = .31$, et la seconde donne $\rho_2 = 0.2\rho_1 + .35 = .41$.
Finalement, $\gamma_0 = \frac{1}{1 - \sum \varphi_i \rho_i} = 2.52$.

- (f) (*) Les autocorrélations partielles $\hat{\rho}_i, i = 1, 2, 3$ se calculent à l'aide des déterminants des matrices d'autocorrélations et sont $.31, .35, \approx 0$. La troisième autocorrélation est en effet une erreur d'arrondissement, car le modèle AR(p) a les autocorrélations partielles nulles au delà du rang p.

8 Les modèles ARMA(p,q)

Définition 8.1 On appelle processus ARMA(p,q) un processus stationnaire $Y_t, t \in \mathbb{Z}$ vérifiant une relation de récurrence :

$$Y_t = \sum_{i=1}^p \varphi_i Y_{t-i} + \sum_{i=0}^q \theta_i \epsilon_{t-i}, \forall t \in \mathbb{Z} \quad (27)$$

où les φ_i, θ_i sont des réels et ϵ_t est un bruit blanc de variance σ^2 .

La notation des polynômes de retard ramène (27) à la forme :

$$\varphi(B)Y_t = \theta(B)\epsilon_t$$

Nous verrons dessous que **les processus ARMA(p,q) avec des polynômes caractéristiques $\varphi(B), \theta(B)$ à racines dehors le cercle unitaire ont deux autres représentations équivalentes :**

1. MA(∞), de Y_t en termes de ϵ_t (appelée aussi représentation linéaire causale), et
2. AR(∞), de ϵ_t en termes de Y_t (appelée aussi représentation inverse).

Ces représentations peuvent être obtenues par des **inversions formelles** de l'équation (27), suivies par un développement de la fraction correspondante dans une série des puissances :

$$Y_t = \frac{\theta(B)}{\varphi(B)} \epsilon_t = \left(\sum_{i=0}^{\infty} \psi_i B^i \right) \epsilon_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad \epsilon_t = \frac{\varphi(B)}{\theta(B)} Y_t = \left(\sum_{i=0}^{\infty} \pi_i B^i \right) Y_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$$

Par exemple, rappelons qu'un processus AR(1) a aussi une représentation causale MA(∞) ssi $|\varphi| < 1$ (obtenue : a) en résolvant la récurrence ou b) par l'inversion formelle du polynôme $\varphi(B) = 1 - \varphi B$). Donc, on a une représentation causale MA(∞) (en termes du bruit passé) du processus AR(1) ssi le **polynôme caractéristique** $\varphi(z) = 1 - \varphi z$ a sa racine à l'extérieur du cercle unitaire $|z| \leq 1$.

Exemple 8.1 ARMA(1,1) Trouver la représentation AR(∞) (i.e. $\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$) du processus ARMA(1)

$$Y_t = \epsilon_t + \theta \epsilon_{t-1} + \phi Y_{t-1}$$

Sol : $\{1, -\theta(1) - \phi(1), \theta(1)^2 + \phi(1)\theta(1), -\theta(1)^3 - \phi(1)\theta(1)^2\}$

8.1 Causalité et inversibilité des modèles ARMA(p,q)

Les problèmes de non-causalité et non-inversibilité des modèles ARMA(p,q) disparaissent quand toutes les racines de $\varphi(z)$ et $\theta(z)$ sont à l'extérieur du cercle unitaire :

Théorème 8.1 a) Un processus ARMA(p,q) avec toutes les racines du polynôme caractéristique $\varphi(z)$ à l'extérieur du cercle unitaire est causal, i.e. il peut être représenté sous la forme : $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ où $\sum |\psi_i| < \infty$ et donc Y_t appartient au espace linéaire engendré par le passé du bruit

$$Y_t \in sp\{\epsilon_{t-i}, i = 0, 1, \dots\}$$

Les coefficients ψ_i sont dans ce cas les coefficients de la série Taylor de $\psi(z) = \frac{\theta(z)}{\varphi(z)}$

b) Un processus ARMA(p,q) avec les racines du polynôme caractéristique $\theta(z)$ à l'extérieur du cercle unitaire est inversible, i.e. le bruit peut être représenté sous la forme : $\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$ où $\sum |\pi_i| < \infty$ et donc ϵ_t appartient au espace linéaire engendré par le passé du signal observé

$$\epsilon_t \in sp\{Y_{t-i}, i = 0, 1, \dots\}$$

Les coefficients π_i sont dans ce cas les coefficients de la série Taylor de la **fonction de transfert** $\pi(z) = \frac{\varphi(z)}{\theta(z)}$

Corollaire 8.1 Pour un processus ARMA(p,q) avec toutes les racines des polynômes caractéristiques $\varphi(z), \theta(z)$ à l'extérieur du cercle unitaire, les espaces linéaires engendrés par le bruit et le passé du signal coïncident :

$$sp\{Y_{t-i}, i = 0, 1, \dots\} = sp\{\epsilon_{t-i}, i = 0, 1, \dots\}$$

et

$$\boxed{\mathbb{E}Y_t \epsilon_{t+k} = 0, \forall k \geq 1}$$

Rémarque : Ce corollaire permettra un développement immédiat d'une approche de prévision (=régression) par projection dans l'espace de Hilbert engendré par le passé.

En conclusion, comme du point de vue pratique les développements Laurent sont inacceptables pour la prévision (parce-qu'elles impliquent les valeurs futures, imprévisibles du bruit), nous allons considérer désormais surtout les modèles ARMA(p,q) avec toutes les racines de $\varphi(z)$ et $\theta(z)$ à l'extérieur du cercle unitaire, qui sont causales et inversibles, et on s'appuyera sur le corollaire ci-dessus.

Exercice 8.1 ARMA(1,1)

1. Trouver la représentation MA(∞) d'un processus ARMA(1,1) causal

$$Y_t = \varphi Y_{t-1} + \epsilon_t + \theta \epsilon_{t-1}$$

2. Trouver la représentation AR(∞) d'un processus invertible ARMA(1,1).
3. Soit Y_t un processus ARMA(1,1) vérifiant l'équation $Y_t - 0.5Y_{t-1} = \epsilon_t + 0.4\epsilon_{t-1}$ avec ϵ_t un bruit blanc.

Précisez si le processus est stationnaire, causal et inversible, et calculez sa fonction d'autocovariance.

4. Trouvez les coefficients ψ_j de sa représentation comme processus $MA(\infty)$ et les coefficients π_j de sa représentation comme processus $AR(\infty)$ et précisez si ces représentations sont convergentes.

Mêmes questions pour le processus $ARMA(2, 1)$ défini par :

$$Y_t - 0.7Y_{t-1} + 0.1Y_{t-2} = \epsilon_t + 2\epsilon_{t-1}$$

Rémarque : Dans le cas le plus simple avec les racines λ_i de l'équation $\varphi(z) = 0$ distinctes, on obtient facilement des formules générales pour les coefficients ψ_n en commençant par un développement en fractions simples $\pi(z) = \frac{\varphi(z)}{\theta(z)} = \sum_i K_i \frac{1}{1-z/\lambda_i}$ où λ_i sont les racines du $\theta(z)$ et donc $K_i = -\frac{\theta(\lambda_i)}{\varphi'(\lambda_i)}$. On arrive à : $\psi_n = \sum_{i=1}^p \frac{K_i}{\lambda_i^{n+1}}$. Des formules pareilles existent pour π_n , et dans le cas des racines non-distinctes.

Dans le cas des racines non-distinctes et complexes, il est préférable d'abord le développement Taylor $\frac{\theta(z)}{\varphi(z)} = \psi(z)$ directement, en obtenant des équations de récurrence pour ψ_k , à partir des coefficients du développement $\varphi(z)\psi(z) = \theta(z)$. Cette méthode, est en effet applicable toujours :

L'exercice ci-dessus peut être généralisé :

Théorème 8.2 (*) a) Pour un processus $ARMA(p, q)$ $\varphi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\varphi(z)$ à l'extérieur du cercle unitaire, les coefficients $\psi_i = \sigma^{-2}\mathbb{E}Y_t\epsilon_{t-i}$ de la représentation causale $Y_t = \sum \psi_i\epsilon_{t-i}$ satisfont la récurrence

$$\varphi(B)\psi_k = \theta_k \iff$$

$$\begin{aligned} \psi_0 = 1, \quad \psi_k &= \theta_k + \sum_{i=1}^{\min[k, p]} \varphi(i)\psi(k-i), \quad 1 \leq k \leq q \\ \psi_k &= \sum_{i=1}^{\min[k, p]} \varphi(i)\psi(k-i), \quad k > q \end{aligned}$$

$$\iff \varphi(B)\psi(k) = \theta(k)$$

Note : Cette dernière équation est une **récurrence de Yule-Walker généralisé**

$$\text{Ainsi, } \begin{cases} \psi_1 = \theta_1 + \varphi_1 \\ \psi_2 = \theta_2 + \varphi_1(\theta_1 + \varphi_1) + \varphi_2 \\ \psi_3 = \theta_3 + \varphi_3 + 2\varphi_1\varphi_2 + \varphi_1^3 + (\varphi_2 + \varphi_1^2)\theta_1 + \varphi_1\theta_2 \\ \dots \end{cases}$$

b) Pour un processus $ARMA(p, q)$ $\varphi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\theta(z)$ à l'extérieur du cercle unitaire, les coefficients π_i de la représentation inverse $\epsilon_t = \sum \pi_i Y_{t-i}$ satisfont la récurrence

$$\begin{aligned} \pi_0 = 1, \quad \pi_k &= -\varphi_k - \sum_{i=1}^{\min[k, q]} \theta(i)\pi(k-i), \quad 1 \leq k \leq p \\ \pi_k &= -\sum_{i=1}^{\min[k, q]} \theta(i)\pi(k-i), \quad k > p \end{aligned}$$

$$\iff \theta(B)\pi(k) = -\varphi(k)$$

$$\text{Ainsi, } \begin{cases} \pi_1 = -\phi(1) - \theta(1) \\ \pi_2 = -\phi(2) + \theta(1)^2 + \phi(1)\theta(1) - \theta(2) \\ \pi_3 = -\phi(3) - \theta(1)^3 - \phi(1)\theta(1)^2 + 2\theta(2)\theta(1) + \phi(2)\theta(1) - \theta(3) + \theta(2)\phi(1) \end{cases}$$

Dem : Formellement, il suffit de projeter les relations $\varphi(B)\psi(B) = \theta(B)$, $\theta(B)\pi(B) = \varphi(B)$ sur la composante k .

Note : Les représentations inverse/causale permettent d'appliquer aux processus ARMA(p,q) les méthodes adaptés aux modèles AR(∞)/MA(∞). En particulier, on a la formule de prediction

$$\begin{aligned} \hat{X}_t &:= E[X_t/F_{t-1}] = -\sum_{i=1}^{\infty} \pi_i X_{t-i} = \\ &(\phi(1) + \theta(1))X_{t-1} + (\phi(2) - \theta(1)^2 - \phi(1)\theta(1) + \theta(2))X_{t-2} + \\ &(\phi(3) + \theta(1)^3 + \phi(1)\theta(1)^2 - 2\theta(2)\theta(1) - \phi(2)\theta(1) - \theta(2)\phi(1) + \theta(3))X_{t-3} + \dots \end{aligned}$$

La formule est plus compliquée a lire, mais, finalement, pour un logiciel symbolique, pas plus compliquée a deriver que son cas particulier AR(p)

$$\hat{X}_t = \phi(1)X_{t-1} + \phi(2)X_{t-2} + \phi(3)X_{t-3} + \dots$$

Théorème 8.3 (*) La fonction génératrice des covariances $f(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k$ d'un processus ARMA est

$$f(z) = \frac{\theta(z)\theta(z^{-1})}{\varphi(z)\varphi(z^{-1})}$$

Pour $z = e^{itx}$, $x \in \mathbb{R}$ on obtient la transformée Fourier des corrélations, aussi appelée densité spectrale. La densité spectrale est nonnegative, ce qui est equivalent au fait que les matrices Toeplitz \mathbf{R} des corrélations sont positives définies.

8.2 (*)L'inversion des series des puissances et des filtres $\varphi(B)$

Le résultat suivant est utile pour l'inversion des modèles AR(p), et aussi des ARMA(p,q), qui sont des processsus Y_t satisfaisant des équations de la forme : $\varphi(B)Y_t = \theta(B)\epsilon_t$.

Théorème 8.4 1. Soit $\sum_{i=0}^{\infty} \psi_i z^i$ le developpement limité autour de 0 de $\frac{1}{1-\varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p}$.

Alors les coefficients ψ_i satisfont la **réurrence Yule-Walker**

$$\begin{aligned} \psi_i &= \phi_1 \psi_{i-1} + \phi_2 \psi_{i-2} + \dots + \phi_p \psi_{i-p}, \quad i \geq p \\ \psi(0) &= 1, \quad \psi(-1) = \dots = \psi(-p+1) = 0 \end{aligned}$$

2. Pour un polynôme $\varphi(z) = \prod_{i=1}^p (1 - z/\lambda_i)$ qui a toutes ses racines λ_i à l'extérieur du cercle unitaire $|z| \leq 1$, $\frac{1}{\varphi(z)}$ a un développement en série de Taylor

$$\frac{1}{\varphi(z)} = \sum_{n=0}^{\infty} \psi_n z^n$$

qui est convergente à l'intérieur du cercle unitaire $|z| = 1$. Dans le cas le plus simple des racines λ_i distinctes, on a

$$\psi_n = \sum_{i=1}^p \frac{K_i}{\lambda_i^{n+1}} \quad (28)$$

où $K_i = -\frac{1}{\varphi'(\lambda_i)}$. (Dans le cas des racines confondues, on a des formules similaires qui utilisent dérivées de degré supérieur).

3. Pour un polynôme $\varphi(z) = \prod_{i=1}^p (1 - z/\lambda_i)$ qui a toutes ses racines λ_i à l'intérieur du cercle unitaire $|z| \leq 1$, $\frac{1}{\varphi(z)}$ a un développement en série de Laurent

$$\frac{1}{\varphi(z)} = \sum_{n=-1}^{-\infty} \psi_n z^n$$

qui est convergente à l'extérieur du cercle unitaire $|z| = 1$. Dans le cas le plus simple des racines λ_i distinctes, on a

$$\psi_n = -\sum_{i=1}^p K_i \lambda_i^{n+1} \quad (29)$$

où $K_i = -\frac{1}{\varphi'(\lambda_i)}$ c) Dans le cas mixte avec racines à l'intérieur et aussi à l'extérieur du cercle unitaire on a un mélange des formules ci-dessus.

Ce resultat justifie des manipulations formelles analogues qu'on fait avec des fonctions dans l'operateur B^3 .

8.3 La prévision linéaire des processus ARMA(p,q)

On se propose de donner au temps t une prévision $\hat{X}_t(k)$ de la valeur X_{t+k} d'un processus. Donc

1. t est le temps de prévision
2. $k > 0$ est l'écart de prévision

3. On peut approcher de manière rigoureuse les manipulations formelles comme l'inversion du polynôme $\varphi(B)$ par plusieurs démarches :

1. **Les fonctions génératrices.** Cette approche associe à chaque suite ψ_n avec $n \in \mathbb{N}$, $-n \in \mathbb{N}$ ou $n \in \mathbb{Z}$ la fonction $\psi(z) = \sum_n \psi_n z^n$. Dans le premier cas appelé série de puissances/Taylor, la série est convergente dans l'intérieur d'un certain "cercle de convergence", dans le deuxième cas, la série est convergente dans l'extérieur d'un certain "cercle de divergence" et dans le troisième cas, appelé série de Laurent, la série est convergente, mais a des expressions différentes dans l'intérieur des "anneaux de convergence" qui évitent les singularités. Le rôle joué par la convergence dans les calculs n'est pas crucial; on peut utiliser parfois même des séries divergentes partout, en les définissant comme objets isomorphes à un certain anneau algébrique.
2. **Les matrices Toeplitz.** On s'aperçoit que les operateurs sur les suites correspondant à des polynômes en B sont représentés par des matrices Toeplitz; on peut démontrer que il y a un isomorphisme entre l'anneau des matrices Toeplitz est celui des fonctions génératrices. Cet isomorphisme explique l'équivalence des deux approches. Formellement, la conclusion est que l'operateur B doit-être traité comme le scalaire $z = 1$ (qui est sa valeur propre), et donc "l'expansion correcte" pour les inversions $\frac{1}{\varphi(z)}$ en série des puissances dépendront du positionnement du point $z = 1$ par rapport aux racines.

3. $t + k$ est le temps à prédire.
4. $\hat{X}_t(k)$ est la prévision
5. $e_t(k) = X_{t+k} - \hat{X}_t(k)$ seront les erreurs de prévision.

Comme les processus ARMA(p,q) satisfont des contraintes linéaires, il est naturel de chercher une **prévision linéaire** $X_t(k)$ par une combinaison linéaire de valeurs passées ou du bruit blanc, à variance minimale, c'est à dire,

$$\begin{aligned}\hat{X}_t(k) &= \sum_{i=0}^{\infty} \pi_{t,k}(i) X_{t-i} \text{ ou} \\ \hat{X}_t(k) &= \sum_{i=0}^{\infty} a_{t,k}(i) \epsilon_{t-i}, k = 1, 2, \dots\end{aligned}$$

Théorème 8.5 *La prévision linéaire à variance minimale des processus ARMA(p,q) avec du bruit blanc Gaussien coïncide avec l'espérance conditionnelle*

$$\hat{X}_t(k) = \mathbb{E}[X_{t+k} | F_t]$$

où $F_t = \{X_t, X_{t-1}, \dots, X_0, X_{-1}, \dots\}$.

Donc, nous allons pouvoir profiter du fait que l'opérateur d'espérance conditionnelle $\hat{X}_t(k) = \hat{X}(t+k|t) = \mathbb{E}[X_{t+k} | F_t]$ (toutes ces notations sont rencontrées dans la littérature) est linéaire.

Toutes les trois représentations AR(∞), MA(∞) et ARMA(p, q) nous aideront dans la prévision, notamment la première.

Définition 8.2 *Si un processus Y_t défini par une autorégression (19) $Y_t = f(Y_{t-1}, Y_{t-2}, \dots) + \epsilon_t$ est tel que la différence $\epsilon_t = Y_t - f(Y_{t-1}, Y_{t-2}, \dots) = Y_t - \hat{Y}_t$ est indépendante du passé $\mathcal{F}_{t-1} = (Y_{t-1}, Y_{t-2}, \dots)$, il sera appelée **bruit d'innovation** (par rapport au passé).*

Dans ce cas,

$$\mathbb{E}[\epsilon_t | Y_{t-1}, Y_{t-2}, \dots] = 0 \tag{30}$$

et ϵ_t est noncorrélé avec Y_{t-1}, Y_{t-2}, \dots

Rq : On montre facilement que pour un bruit d'innovation, les innovations ϵ_t satisfont aussi $\mathbb{E}\epsilon_t = 0$ et $\mathbb{E}\epsilon_t \epsilon_{t-k} = 0$, pour $k \neq 0$. Donc, les innovations constituent un bruit blanc de deuxième ordre.

Le bruit blanc d'une équation autorégressive causale a la propriété très convenable de coïncider avec l'erreur de prédiction par rapport au passé $Y_t - \hat{Y}_t = Y_t - \mathbb{E}[Y_t | Y_{t-1}, Y_{t-2}, \dots]$:

Un résultat fondamental assure que le bruit d'un modèle (20) est d'innovation ssi le modèle est causal, et ssi le polynôme caractéristique a seulement des racines plus grandes en valeur absolue que 1. Dans ce cas, on a :

$$\hat{Y}_{t|t-1} = \mathbb{E}[Y_t | \{Y_{t-1}, Y_{t-2}, \dots\}] = f(Y_{t-1}, Y_{t-2}, \dots) = \sum_{i=1}^p \varphi_i Y_{t-i}$$

et donc la prédiction devient très aisée, car on applique simplement la relation d'autorégression, **en oubliant le bruit**.

8.4 Prévision des processus AR(p) par la récurrence de Yule-Walker

Nous considérons maintenant plus en détail la prévision des processus stationnaires dans la forme AR(p). La méthode sera "d'appliquer l'opérateur chapeau" dans l'équation linéaire définissant le modèle.

Théorème 8.6 Pour un modèle $AR(p)$

$$\varphi(B)X_t = \epsilon_t$$

tel que le symbole $\varphi(z)$ ne s'annule pas dans le cercle unitaire, le bruit ϵ_t est un bruit d'innovation, i.e. $\mathbb{E}[\epsilon_{t+k}|\mathcal{F}_t] = 0$ si $k > 0$, et les prévisions satisfont la récurrence Yule-Walker :

$$\hat{X}_t(k) := \mathbb{E}[X_{t+k}/\{X_t, X_{t-1}, \dots\}] = \sum_{i=1}^p \varphi_i \hat{X}_t(k-i), \quad \hat{X}_t(k) = X_{t+k}, \quad \text{si } k \in \{0, -1, -2, \dots\} \quad (31)$$

En particulier,

$$\hat{X}_t(1) = \mathbb{E}[X_{t+1}/\{X_t, X_{t-1}, \dots\}] = \sum_{i=1}^p \varphi_i \hat{X}_t(1-i) = \sum_{i=1}^p \varphi_i X_{t-i}$$

Démonstration immédiate, par la causalité et donc l'indépendance de ϵ_{t+k} de \mathcal{F}_t .

Exemple 8.2 La prévision linéaire $X_t(k)$ pour un processus $AR(1)$ à moyenne 0 satisfait la récurrence Yule Walker

$$X_t(k) = \varphi X_t(k-1), \quad X_t(0) = X_t,$$

étant donc simplement

$$X_t(k) = X_t \varphi^k$$

Pour un processus $AR(1)$ à moyenne connue μ elle est

$$X_t(k) - \mu = (X_t - \mu)\varphi^k \iff X_t(k) = \mu + (X_t - \mu)\varphi^k$$

Théorème 8.7 Soit

$$\varphi(B)X_t = (1 - \lambda_1 B)(1 - \lambda_2 B)X_t = \epsilon_t$$

un processus $AR(2)$, avec $\lambda_1 \neq \lambda_2$ étant les inverses des racines de $\varphi(z) = 0$. Montrez que les prévisions Box-Jenkins $X_t(k)$ au temps t sont de la forme :

$$X_t(k) = A_1(t)\lambda_1^k + A_2(t)\lambda_2^k = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} X_t - \lambda_1 \lambda_2 \frac{(\lambda_1^k - \lambda_2^k)}{\lambda_1 - \lambda_2} X_{t-1} \quad (32)$$

où encore

$$X_t(k) = \psi_k X_t - \lambda_1 \lambda_2 \psi_{k-1} X_{t-1}, \quad k = 0, 1, \dots \quad (33)$$

Démonstration : Les prévisions $\hat{X}_t(k)$ satisfont la récurrence Yule Walker :

$$\hat{X}_t(k) - \varphi_1 \hat{X}_t(k-1) - \varphi_2 \hat{X}_t(k-2) = \hat{X}_t(k) - (\lambda_1 + \lambda_2) \hat{X}_t(k-1) + \lambda_1 \lambda_2 \hat{X}_t(k-2) = 0, \quad (34)$$

et donc

$$\hat{X}_t(k) = \varphi_1 \hat{X}_t(k-1) + \varphi_2 \hat{X}_t(k-2),$$

avec valeurs initiales $\hat{X}_t(0) = X_t, \hat{X}_t(-1) = X_{t-1}$.

La solution générale d'une récurrence d'ordre 2 $\psi(k) = \varphi_1 \psi(k-1) + \varphi_2 \psi(k-2)$, pour valeurs initiales arbitraires $\psi(0) = x_0, \psi(1) = x_1$ est :

1. avec des racines distinctes λ_1, λ_2 :

$$\psi(k) = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} x_1 - \frac{(\lambda_1^k \lambda_2 - \lambda_2^k \lambda_1)}{\lambda_1 - \lambda_2} x_0$$

2. avec des racines confondues $\lambda_1 = \lambda_2 = \lambda$:

$$\psi(k) = x_0 \lambda^k + (x_1 - \lambda x_0) k \lambda^{k-1}$$

Aussi, en termes des valeurs initiales $\psi(0) = x_0, \psi(-1) = x_{-1}$, la solution quand $\lambda_1 \neq \lambda_2$ est :
 $\psi_k = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} x_0 - \frac{(\lambda_1^{k+1} \lambda_2 - \lambda_2^{k+1} \lambda_1)}{\lambda_1 - \lambda_2} x_{-1}$, d'où le résultat.

Remarque 8.1 Dans le cas causal, $\lambda_i < 1$, et donc la **prévision éventuelle**

$$\lim_{k \rightarrow \infty} X_t(k) = 0$$

est toujours 0.

Il est facile d'étendre cette approche pour tous les processus autorégressifs AR(p) à ordre p fini, et d'obtenir des formules explicites de prévision en termes de racines de l'équation $\varphi(z) = 1$, ou des dernières valeurs observées $X_t, X_{t-1}, \dots, X_{t-p+1}$.

Théorème 8.8 La fonction génératrice $X(z) := \sum_0^\infty z^k \hat{X}_t(k)$ des prévisions d'un processus AR(p) est

$$X(z) = \frac{X_t + z(\varphi_2 X_{t-1} + \varphi_3 X_{t-2} + \dots) + z^2(\varphi_3 X_{t-1} + \varphi_4 X_{t-2} + \dots) + \dots z^{p-1} \varphi_p X_{t-1}}{1 - \varphi(z)}$$

Exemple 8.3 Obtenir les prédictions $\hat{X}_t(k)$ pour le AR(2)

$$(1 - .2B - .35B^2)X_t = (1 - .7B)(1 + .5B)X_t = \epsilon_t$$

Sol : on a

$\hat{X}_t(1) = \varphi_1 X_t + \varphi_2 X_{t-1}, \hat{X}_t(2) = (\varphi_1^2 + \varphi_2) X_t + \varphi_1 \varphi_2 X_{t-1}, \dots$ et en général is s'agit des coefficients du développement limité du théorème 8.7

$$\frac{X_t + z\varphi_2 X_{t-1}}{1 - \varphi(z)} = X_t + z(\varphi_1 X_t + \varphi_2 X_{t-1}) + z^2((\varphi_1^2 + \varphi_2) X_t + \varphi_1 \varphi_2 X_{t-1}) + \dots$$

On peut aussi résoudre les équations YW, ou utiliser la formule (33).

9 Les modèles ARIMA(p,d,q)

Définition 9.1 On appelle processus ARIMA(p,d,q) un processus non stationnaire X_t pour le quel le processus différencié d'ordre d , $Y_t = (1 - B)^d X_t, t \in \mathbb{Z}$ est stationnaire, et vérifie une relation de récurrence ARMA(p,q) :

$$Y_t = \sum_{i=1}^p \varphi_i Y_{t-i} + \sum_{i=0}^q \theta_i \epsilon_{t-i}, \forall t \in \mathbb{Z} \quad (35)$$

où les φ_i, θ_i sont des réels et ϵ_t est un bruit blanc de variance σ^2 .

La notation des polynômes de retard ramène (35) à la forme :

$$\boxed{\varphi(B)(1-B)^d X_t = \varphi(B)Y_t = \theta(B)\epsilon_t},$$

où $\phi(B), \theta(B)$ sont des polynômes relativement primes dans l'opérateur de retard B à ordres p, q avec coefficient libre 1, et avec racines dehors le cercle unitaire.

Formellement, il s'agit des processus ARMA ayant aussi la racine 1, et nous verrons qu'en effet, la prévision des processus ARIMA(p,d,q) est donnée par les mêmes formules que celle des processus stationnaires ARMA(p,q).

9.1 Prévision linéaire des modèles autorégressifs ARIMA(p,d,0)

Les deux méthodes principales pour l'estimation des paramètres sont la méthode des moments et la maximisation de la vraisemblance. La première méthode s'appuie sur les formules théoriques des moments, en l'occurrence les corrélations.

Exemple 9.1 La prévision linéaire $X_t(k)$ pour le processus ARIMA(0,1,0) à moyenne c satisfait la récursion Yule Walker

$$X_t(k) = X_t(k-1)$$

et est donc constante

$$X_t(k) = X_t$$

(c'est un cas particulier de la formule dans l'exercice sur la prévision AR(1)).

Exemple 9.2 Déduisez la formule de prévision Box-Jenkins pour un processus ARIMA(1,1,0) Calculez la limite $\lim_{k \rightarrow \infty} X_t(k)$ pour un processus ARIMA(1,1,0)

En conclusion, nous voyons que le "type" de la fonction de prévision $X_t(k)$ dans le cas des bruits indépendants (sans partie MA) est déterminé complètement par la fonction $\varphi(z)$, et on verra que ça reste vrai pour les processus ARIMA(p,d,q), pour $k > q$.

Proposition 9.1 La fonction de prévision "eventuelle" de Box-Jenkins pour les processus ARIMA(p,d,q) est un élément de l'espace linéaire des solutions de la récursion $\varphi(B)X_t(k)$, pour $k > q$.

Par exemple, pour les processus ARIMA(0,d,q) la fonction de prévision "eventuelle" est un polynôme d'ordre $d-1$.

Exemple 9.3 On considère un processus $\{X_t\}$ pour lequel la série différenciée deux fois est un bruit blanc, c'est à dire $\{X_t\}$ est un processus ARIMA(0,2,0). Montrez que la fonction de prévision Box-Jenkins est donnée par

$$X_t(k) = X_t + k(X_t - X_{t-1}), \quad k \geq 0.$$

donc les prévisions se trouvent sur la droite qui passe par les deux derniers points.

Définition 9.2 Les dernières $p+d$ valeurs $(X_t(q), X_t(q-1), \dots, X_t(q-d-p+1))$ qui précèdent $X_t(q)$ (donc avant le point où la récursion de Yule Walker devient valable) s'appellent les **valeurs pivots**.

Il suit clairement que :

Proposition 9.2 *La prévision "eventuelle" de Box-Jenkins pour les processus ARIMA(p,d,q) est la fonction dans l'espace linéaire des solutions de la récursion $\varphi(B)X_t(k)$ qui passe par les valeurs pivots.*

Corollaire 9.1 *La prévision linéaire $X_t(k)$ pour le processus ARIMA(0,d,0) est donnée par le polynôme d'ordre $d-1$ qui passe par les d dernières points.*

9.2 Prévision des modèles à bruit moyenne mobile MA(1)

Pour la prévision linéaire $\hat{X}_t(k)$ des processus MA(q), l'idée est de convertir à la représentation autoregressive AR(∞) $\epsilon_t = X_t + \sum_{i=1}^{\infty} \pi_i X_{t-i}$, qui nous fournit une formule explicite :

$$\hat{X}_t(t|t-1) = \sum_{i=1}^{\infty} \pi_i X_{t-i} = (1 - \pi(B))X_t$$

Exemple 9.4 *Pour le processus MA(1), on vérifie facilement que $X_t(k) = 0$ pour $k = 2, 3, \dots$, (pour une généralisation, voir le théorème 10.3). Pour $k = 1$, la définition :*

$$X_{t+1} = \epsilon_{t+1} + \theta \epsilon_t$$

donne :

$$X_t(1) = \theta \epsilon_t$$

Pour se débarrasser de ϵ_t , on peut utiliser

$$\epsilon_t = X_t - \theta \epsilon_{t-1},$$

qui ramène à la représentation :

$$\epsilon_t = \sum_{i=0}^{t-1} (-1)^i \theta^{i+1} X_{t-i} - (-\theta)^{t+1} \epsilon_0 = \dots = \sum_{i=0}^{\infty} (-1)^i \theta^i X_{t-i} = X_t + \sum_{i=i}^{\infty} (-1)^i \theta^i X_{t-i}$$

Donc, $X_{t+1} = \epsilon_{t+1} + \sum_{i=1}^{\infty} (-1)^{i-1} \theta^i X_{t+1-i}$ et

$$X_t(1) = \hat{X}_t = \sum_{i=0}^{\infty} (-1)^i \theta^{i+1} X_{t-i}$$

$$X_t(k) = 0, k = 2, 3, \dots$$

Il est naturel de supposer que l'information est finie, i.e. $F_t = \{X_t, X_{t-1}, \dots, X_1\}$. La formule dans ce cas, obtenue par récursion, est :

$$\hat{X}_t = \mathbb{E}[X_{t+1}|F_t] = \sum_{i=0}^{t-1} (-1)^i \theta^{i+1} X_{t-i} - (-\theta)^{t+1} \epsilon_0$$

Comme ϵ_0 n'est pas connue, en pratique on utilise : $\hat{X}_t(1) = \sum_{i=0}^t (-1)^i \theta^{i+1} X_{t-i}$. Si $\theta < 1$ et t est large, la différence sera négligeable.

Donc, cet exemple montre déjà qu'une estimation du "bruit inobservable" $\epsilon_t, \dots, \epsilon_1, \epsilon_0, \dots$ est incontournable pour les modèles ARMA avec $q \geq 1$.

9.3 Prédiction des processus ARMA(p,1)

Exercice 9.1 On considère le processus ARMA(1,1) à moyenne 0 $(1 - \varphi B)Y_t = (1 + \theta B)\epsilon_t$ où $-1 < \varphi < 1$ et $-1 < \theta < 1$.

- Montrez que la fonction de prévision Box-Jenkins est donnée par $Y_t(k) = Y_t(1)\varphi^{k-1}$, $k \geq 1$, et que

$$\begin{aligned} Y_t(1) &= \varphi Y_t + \theta \epsilon_t \\ &= (\varphi + \theta)Y_t - \theta Y_{t-1}(1) \\ &= (\varphi + \theta)\{Y_t - \theta Y_{t-1} + \theta^2 Y_{t-2} + \dots\} \end{aligned}$$

- On utilise ce modèle pour ajuster une série et on obtient comme estimations des paramètres $\varphi = 0.8$, $\theta = 0.3$ et $c = ?$. Les dix dernières valeurs disponibles sont :

t :	51	52	53	54	55	56	57	58	59	60
y_t :	2.98	4.10	6.10	9.36	8.57	8.82	7.31	7.19	2.36	0.40

Donnez les prévisions des trois valeurs suivantes de la série. Quelle parmi les trois formules pour $Y_t(1)$ ci-dessus paraît la plus convenable à appliquer ?

Exercice 9.2 Est-ce que les résultats ci-dessus restent vrais si $\varphi = 1$, donc pour le processus ARIMA(0,1,1) (appelé aussi IMA(1,1)), défini par :

$$(1 - B)Y_t = (1 + \theta B)\epsilon_t?$$

Si $\theta < 1$, les coefficients de la représentation du bruit sont :

$$\pi_i = (1 + \theta)(-\theta)^{i-1}, i \geq 1,$$

(à vérifier).

- Montrez qu'on peut le représenter :

$$Y_t = \epsilon_t + (1 + \theta) \sum_{k=1}^{t-1} \epsilon_{t-k} + \theta \epsilon_0$$

- Montrez que $\bar{Y}_t = (1 + \theta)Y_t - \theta \bar{Y}_{t-1}$.

Note : La dernière formule est appelée lissage exponentiel, au moins quand $\theta \in (-1, 0)$ et donc $\alpha = 1 + \theta \in (0, 1)$. La formule donne une moyenne pondérée : $\bar{Y}_t = \alpha Y_t + (1 - \alpha)\bar{Y}_{t-1}$ α s'appelle constante de lissage.

Rémarques : 1) Plus α est petit, plus la nouvelle série est lisse et les valeurs passées ont un plus grand poids dans la prévision. Quand α est proche de 1 les valeurs les plus récentes ont le poids le plus important.

2) On peut voir la prévision Box-Jenkins comme une généralisation du lissage exponentiel, en utilisant des paramètres estimés à partir des données (au-lieu de ad-hoc).

9.4 Prévision des processus ARIMA(p,d,q)

En conclusion, pour la prévision linéaire $\hat{X}_t(k)$ des processus ARIMA(p,d,q), on aura toujours besoin d'une estimation de $\epsilon_{t-1}, \epsilon_{t-2}, \dots$, ou au moins de $\epsilon_{-1}, \epsilon_{-2}, \dots$, i.e. du "bruit inobservable passé" du modèle. On peut aussi recourir à la représentation AR(∞), dans quel cas on aura besoin de X_{-1}, X_{-2}, \dots , qui sont aussi inobservables. En plus, le resultat final demandera une approximation des valeurs precedant le debut d'observations 0 ; l'approximation la plus simple dans l'absence des moyennes est $\epsilon_k = Y_k = 0$ pour $k < 0$.

Théorème 9.1 Dans le cas d'un modèle ARIMA(p,d,q), la meilleure prévision linéaire au temps t est :

$$\hat{X}_t(k) = \mathbb{E}[X_{t+k}|F_t] = \sum_{i=1}^p \tilde{\varphi}_i \hat{X}_t(k-i) + \sum_{i=k}^q \theta_i \hat{\epsilon}_{t+k-i}$$

où les $\tilde{\varphi}_i$ sont les coefficients du polynôme $\varphi(B)(1-B)^d$ (dans le cas d'un modèle ARMA(p,q) $\tilde{\varphi}_i = \varphi_i$).

Pour $k > q$, cette formule est exactement la récurrence homogène Yule-Walker $\varphi(B)\hat{X}_t(k) =$, et donc la prévision sera donnée par la solution de cette équation qui passe par les $p + d$ points pivots.

Les inconnues $\hat{\epsilon}_{t-i}, i \geq 0$ peuvent être enlevés en utilisant la représentation inverse "π" du bruit en fonction de la série, ou en utilisant $\hat{\epsilon}_t = Y_t - \hat{Y}_{t-1}(1)$ (les dernières se calculent récursivement). Une estimation arbitraire de $\epsilon_0, \epsilon_{-1}, \dots$ sera nécessaire.

9.5 Exercices : TD 2

1. Calculer la fonction d'autocovariance du processus à valeurs en \mathbb{R}^2 :

$$Y_n = \begin{pmatrix} a_0 \epsilon_n + a_1 \epsilon_{n-1} \\ b_1 \epsilon_{n-1} + b_2 \epsilon_{n-2} \end{pmatrix}$$

où ϵ_n est un bruit blanc standard.

2. Restrictions sur les valeurs des coefficients d'autocorrélation pour les processus MA.

Trouvez, pour le processus MA(1), les valeurs maximales et minimales de la corrélation ρ_1 et les valeurs de θ pour les quelles ces valeurs sont atteintes.

3. a) Déterminez la corrélogramme des processus suivants :

(i) le processus MA(2) $Y_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$

(ii) le processus MA(3) $Y_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3}$

- b) Calculez et tracez la correlogramme pour les cas :

(i) MA(2) : $\theta_1 = -5/6, \theta_2 = 1/6$

(ii) MA(2) : $\theta_1 = 0.8, \theta_2 = 0.5$

(iii) MA(3) : $\theta_1 = 0.8, \theta_2 = -0.4, \theta_3 = -0.3$

4. Investiguez si les deux processus MA(2) sont inversibles, en examinant la récurrence obtenue par la méthode des coefficients indéterminés.

5. Soit Y_t un processus ARMA(1,1) vérifiant l'équation $Y_t - 0.5Y_{t-1} = \epsilon_t + 0.4\epsilon_{t-1}$ avec ϵ_t un bruit blanc.

- (a) Précisez si le processus est stationnaire, causal et inversible.
- (b) Trouvez les coefficients ψ_j de sa représentation comme processus $MA(\infty)$ et les coefficients π_j de sa représentation comme processus $AR(\infty)$ et précisez si ces représentations sont convergentes.
6. Mêmes questions pour les processus $ARMA(2, 1)$ et $ARMA(2, 2)$ définies par :
- a) $Y_t - \frac{1}{2}Y_{t-1} - \frac{3}{16}Y_{t-2} = \epsilon_t + 1.25\epsilon_{t-1}$ b) $(1 - B + \frac{B^2}{4})Y_t = (1 + B + B^2)\epsilon_t$
- Indication* : Remarquez que pour le processus $ARMA(p, q)$, la récurrence pour $k > q$ est exactement comme pour le processus $AR(p)$; mais, les premières équations obtenues en comparant les coefficients des puissances en $\boxed{\psi(B) \varphi(B) = \theta(B)}$ pour $k \leq q$ changent par l'ajout de θ_k de coté droite.
7. Soit le processus :

$$(1 - .8B + .16B^2)Y_t = (1 + \theta B)\epsilon_t$$

- (a) Est-ce que ce processus est stationnaire causal ? Si oui, obtenez la "représentation ψ " du Y_t par rapport au bruit ϵ_t .
- (b) Sous quelles conditions est ce processus inversible ? Obtenez la "représentation π " du bruit ϵ_t en termes de la série. De quel problème on s'aperçoit si le processus n'est pas inversible ?
8. a) Trouvez les inégalités (il y en a trois) qui définissent la région (triangulaire) du plan (θ_1, θ_2) pour laquelle un processus $MA(2)$ est inversible. Tracez la région sur un graphe. Indiquez le domaine des racines réelles et celui des racines complexes. *Indication* : Les conditions pour avoir racines de module plus grand que 1 sont différentes pour le cas des racines complexes et celui des racines réelles, et pour un polynôme $\theta(z) = 1 + \theta_1 z + \theta_2 z^2$, la condition pour racines réelles de module plus grand que 1 sont plus compliquées que les conditions (équivalentes) que le polynôme "reciproque" $\tilde{\theta}(z) = z^2 \theta(1/z) = z^2 + \theta_1 z + \theta_2$ ait des racines réelles z_i de module plus petit que 1. Pour ce dernier polynôme, les conditions sont :
- (a) racines complexes : $|z_i|^2 = |z_1 z_2| = |\frac{c}{a}| = |\theta_2| < 1$.
- (b) racines réelles : $\tilde{\theta}(1) = 1 + \theta_1 + \theta_2 > 0$, $\tilde{\theta}(-1) = 1 - \theta_1 + \theta_2 > 0$
- b) (*) Pour le processus $MA(2)$, trouvez un domaine S contenant toutes les valeurs possibles des coefficients d'autocorrélation ρ_1, ρ_2 tel que le processus soit inversible, et les valeurs de θ_1, θ_2 pour les quelles les valeurs sur la frontière de S sont atteintes.

Solution : La région d'inversibilité dans le domaine (θ_1, θ_2) :

$$\begin{aligned} \theta_2 &> -\theta_1 - 1 \\ \theta_2 &> \theta_1 - 1 \\ \theta_2 &< 1 \end{aligned}$$

est le triangle situé dessus les deux lignes $\theta_2 + \theta_1 = -1$, $\theta_2 = \theta_1 - 1$ et dessous la ligne $\theta_2 < 1$.

Les racines sont réelles/complexes dessous/dessus la parabole $\theta_2 = \frac{\theta_1^2}{4}$.

b) Pour passer de (θ_1, θ_2) à (ρ_1, ρ_2) on utilise

$$\rho_1 = \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2} \quad \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

Transformant les équations antérieures, on trouve :

$\theta_2 = 1$ implique $\rho_1 = \frac{2\theta_1}{2+\theta_1^2}$, $\rho_2 = \frac{1}{2+\theta_1^2}$, $\theta_1 = \frac{\rho_1}{2\rho_2}$, $\rho_2(2 + \frac{\rho_1^2}{4\rho_2^2}) = 2\rho_2 + \frac{\rho_1^2}{4\rho_2} = 1$ et donc $\rho_1^2 = 4\rho_2(1 - 2\rho_2)$. Finalement, on trouve

$$\begin{aligned} \text{dessous } \rho_1 &= 2\sqrt{\rho_2(1 - 2\rho_2)} \\ \rho_2 + 1/2 &\geq \rho_1 \\ \rho_2 + 1/2 &\geq -\rho_1 \end{aligned}$$

où les dernières deux inégalités viennent de l'inégalité entre les moyennes arithmétique et géométrique de $(1 + \theta_2)$, θ_1 .

9. (*) Trouver le domaine de causalité dans le plan (φ_1, φ_2) d'un processus $AR(2)$.
10. Obtenez, en partant directement du système de Yule-Walker, les premières cinq corrélations pour un processus $AR(2)$ avec : a) $\phi_1 = 0.6, \phi_2 = -0.2$ b) $\phi_1 = -0.6, \phi_2 = 0.2$ Calculez aussi la variance $\gamma(0)$. Tracez les corrélations.

Sol : Le domaine de causalité d'un processus $AR(2)$

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \epsilon_t$$

(beaucoup plus compliquée que pour le $AR(1)$), obtenu comme le domaine d'inversibilité du processus $Ma(2)$, est le triangle situé en dessous de $\varphi_2 + \varphi_1 < 1$, $\varphi_2 - \varphi_1 < 1$ et dessus $\varphi_2 = -1$.

11. a) Vérifiez si le processus $AR(2)$ $Y_t = -0.3Y_{t-1} + 0.10Y_{t-2} + \epsilon_t$ est stationnaire causal. Calculez son corrélogramme, en partant directement du système de Yule-Walker, et tracez le.
b) Même questions pour le processus $AR(2)$ $Y_t = -Y_{t-1} - 0.34Y_{t-2} + \epsilon_t$.
12. Calculez la fonction d'autocovariance et la fonction d'autocorrélation des processus dans les exercices antérieurs.
13. Une question d'unicité - est-ce que deux processus distincts peuvent avoir la même FAC (fonction d'autocovariance) ?

Soient $\{u_t, t \in \mathbb{Z}\}$ et $\{v_t, t \in \mathbb{Z}\}$ deux bruit blancs de variances respectives σ^2 et $\theta^2\sigma^2$, o $0 < |\theta| < 1$. On considère alors les processus aléatoires $\{X_t, t \in \mathbb{Z}\}$ et $\{Y_t, t \in \mathbb{Z}\}$ tels que :

$$\begin{aligned} X_t &= u_t + \theta u_{t-1} \\ Y_t &= v_t + \frac{1}{\theta} v_{t-1} \end{aligned}$$

Montrer que $\{X_t, t \in \mathbb{Z}\}$ et $\{Y_t, t \in \mathbb{Z}\}$ ont la même fonction d'autocovariance.

14. (*) Une question d'inversibilité - est ce qu'un processus à représentation MA noninversible peut aussi avoir une autre représentation inversible ? Soit $\{U_t, t \in \mathbb{Z}\}$ le processus aléatoire défini par l'équation

$$U_t = \epsilon_t + \frac{1}{\theta} \epsilon_{t-1} \text{ où } |\theta| < 1$$

et ϵ_t est bruit blanc.

- (a) Montrer que cette représentation du processus U_t n'est pas inversible.
- (b) On pose maintenant $w_t = \sum_{j=0}^{+\infty} \theta^j U_{t-j}$. Montrer que $\{w_t, t \in \mathbb{Z}\}$ est un bruit blanc dont on précisera la variance en fonction de σ^2 et θ
- (c) Montrer que $U_t = w_t + \theta w_{t-1}$ et que cette représentation de $\{U_t, t \in \mathbb{Z}\}$ est inversible.

10 L'estimation des modèles ARIMA(p,d,q)

10.1 Équations de Yule-Walker pour les covariances/corrélations des processus ARMA(p,q)

On obtient les mêmes équations de Yule-Walker $\gamma_k = \sum_{i=1}^p \varphi_i \gamma_{k-i}$ pour $k \geq N = q + 1$. Par contre, les premières équations pour $k < N$ deviennent plus compliquées. Soit $\gamma_k^{(Y,\epsilon)} = \mathbb{E}[Y_t \epsilon_{t-k}]$ les corrélations de la série avec le bruit, données par $\psi_k \sigma^2$ si $k \geq 0$ et par 0 outrement (par la causalité). On obtient :

$$\gamma_k = \sum_{i=1}^p \varphi_i \gamma_{k-i} + \sum_{0 \leq j \leq q} \theta_j \gamma_{k-j}^{(Y,\epsilon)} = \sum_{i=1}^p \varphi_i \gamma_{k-i} + \sigma^2 \sum_{k \leq j \leq q} \theta_j \psi_{k-j} \quad (36)$$

Pour appliquer la recursion, il faut obtenir les coefficients ψ_j , $j = 1, \dots, p$ et aussi les p valeurs initiales $\gamma_q, \dots, \gamma_{q-p+1}$, qu'on trouvera en utilisant les équations (36) et la symétrie de γ_k .

Rappel : Pour un processus ARMA(p,q) causal, les coefficients $\psi_i = \sigma^{-2} \mathbb{E} Y_t \epsilon_{t-i}$ de la représentation causale $Y_t = \sum \psi_i \epsilon_{t-i}$ satisfont la récurrence

$$\begin{aligned} \psi_0 &= 1, & \psi_k &= \theta_k + \sum_{i=1}^{\min[k,p]} \varphi(i) \psi(k-i), \quad 1 \leq k \leq q \\ \psi_k &= \sum_{i=1}^{\min[k,p]} \varphi(i) \psi(k-i), & k &> q \end{aligned}$$

Exemple 10.1 ARMA(1,1)

1. Montrez que $\rho_1 = \varphi + \frac{\theta \sigma^2}{\gamma_0}$, $\gamma_0 = \frac{\sigma^2(1+\theta^2+2\theta\varphi)}{1-\varphi^2}$, et $\{\rho_k\}$ est :

$$\rho_k = \frac{(\varphi + \theta)(1 + \varphi\theta)}{1 + 2\varphi\theta + \theta^2} \varphi^{k-1}, \quad k \geq 1.$$

2. Tracez la corrélogramme pour le cas : $\varphi = \pm 0.7, \theta = \pm 0.5$.

En général, nous trouvons :

Théorème 10.1 (*) Les premières $p + 1$ covariances s'obtiennent du système à $p + 1$ équations et $p + 1$ inconnues :

$$\mathbf{\Gamma} \begin{pmatrix} 1 \\ -\varphi_1 \\ -\varphi_2 \\ \dots \\ -\varphi_p \end{pmatrix} = \sigma^2 \mathbf{\Psi} \begin{pmatrix} 1 \\ \theta_1 \\ \theta_2 \\ \dots \\ \theta_q \end{pmatrix} \quad (37)$$

ou $\mathbf{\Gamma}$ est la matrice des covariances

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(p) \\ \gamma(1) & \gamma(0) & \dots & \gamma(p-1) \\ \dots & \dots & \dots & \dots \\ \gamma(p) & \gamma(p-1) & \dots & \gamma(0) \end{pmatrix}$$

Ψ est la matrice des dimensions $(p+1) \times (q+1)$:

$$\Psi = \begin{pmatrix} \psi(0) & \psi(1) & \dots & \psi(q) \\ 0 & \psi(0) & \dots & \psi(p-1) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \psi(0) \end{pmatrix}$$

et $\psi_i = \sigma^{-2} \mathbb{E}Y_t \epsilon_{t-i}$, $i = 0, \dots, q$ sont calculés par la récurrence

$$\psi_0 = 1, \quad \psi_k = \theta_k + \sum_{i=1}^{\min[k,p]} \varphi(i)\psi(k-i), \quad 1 \leq k \leq q$$

10.2 La détermination de l'ordre d'un modèle autorégressif

Supposons d'abord que p est fini et considérons le problème de déterminer la projection $\hat{X}_t(1) = \sum_{i=1}^l \pi_{t,l}(i)X_{t+1-i}$ dans l'espace $F_{t,k} = \{X_t, X_{t-1}, \dots, X_{t-l+1}\}$, où $l \geq p$. Soit $\boldsymbol{\pi}^{(t,l)} = (\pi_{t,l}(i), i = 1, \dots, l)$ le vecteur des coefficients de la régression. Par le théorème 8.5, il est nécessaire que $\boldsymbol{\pi}^{(t,l)} = (\varphi_1, \dots, \varphi_p, 0, \dots, 0)$.

Note : Comme nous sommes ici dans le cadre d'une régression classique avec du bruit indépendant, il est nécessaire que les coefficients $\boldsymbol{\pi}^{(t,l)}$ satisfont

$$\begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(l-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(l-2) \\ \dots & \dots & \dots & \dots \\ \gamma(l-1) & \gamma(l-2) & \dots & \gamma(0) \end{pmatrix} \boldsymbol{\pi}^{(t,l)} = \begin{pmatrix} \gamma(1) \\ \gamma(2) \\ \dots \\ \gamma(l) \end{pmatrix}$$

ou encore (en divisant par $\gamma(0)$),

$$\begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \dots & \dots & \dots & \dots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix} \boldsymbol{\pi}^{(t,l)} = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \dots \\ \rho(l) \end{pmatrix}$$

Pour $l = p$, on retrouve ainsi le système Yule-Walker, et pour $l > p$ il n'est pas difficile de voir que les vecteurs des coefficients φ_i étendu par zéros satisfont aussi des systèmes Yule-Walker.

Considérons maintenant le problème de déterminer si $p < \infty$; autrement dit, est-ce que les systèmes Yule-Walker ont toujours comme dernière composante un 0, à partir d'un point p ?

Définition 10.1 La suite $\varphi_n^{(n)}$ des dernières composantes des systèmes Yule-Walker d'ordre n est appelée la suite des **corrélations partielles**.

Par le théorème 8.5

Théorème 10.2 Une série stationnaire est $AR(p)$ avec $p < \infty$ ssi toutes les corrélations partielles sont 0 à partir du point $p+1$.

Le vrai ordre p du modèle est inconnu (et potentiellement infini). Pour le déterminer, on peut commencer en supposant $p \geq 1$; on calcule $\hat{\varphi}_1$ en supposant $p = 1$:

$$\hat{\varphi}_1^{(1)} = \rho_1$$

Si ce coefficient est "statistiquement" 0, ça suggère qu'on a à faire peut-être avec du bruit blanc. On continue de toute façon en investiguant $p \geq 2$; On calcule $\hat{\varphi}_1, \hat{\varphi}_2$ en supposant $p = 2$; on trouve alors :

$$\hat{\varphi}_2^{(2)} = \frac{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad (38)$$

Si le deuxième coefficient $\hat{\varphi}_2^{(2)}$ est "statistiquement" 0, on soupçonne qu'on a à faire avec un modèle $AR(1)$. On passe à

investiguer $p \geq 3$, en calculant le coefficient $\hat{\varphi}_3^{(3)} = \frac{\begin{vmatrix} 1 & \rho(1) & \rho(1) \\ \rho(1) & 1 & \rho(2) \\ \rho(2) & \rho(1) & \rho(3) \end{vmatrix}}{\begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix}}$ qui est l'estimation de φ_3 obtenue du système d'ordre

3, puis à $p \geq 4, \dots$

Donc pratiquement, on calcule un après l'autre quelques corrélations partielles, en essayant de voir s'il sont "statistiquement" 0. Si oui, à partir d'un point $p + 1$, on conclut qu'on a à faire avec un modèle $AR(p)$. Si non (et si les corrélations ne deviennent 0 non plus), on passe à chercher un modèle mixte $ARMA(p, q)$.

10.3 Intervalles de confiance pour la prédiction

Théorème 10.3 (*) Soit $e_t(k) = X_{t+k} - X_t(k)$ l'erreur de prévision de la prédiction linéaire optimale $X_t(k) = \sum_{i=0}^{\infty} a_{t,k}(i) \epsilon_{t-i}$ d'un processus stationnaire causal X_t , et soit $V(k) = \mathbb{E}[X_{t+k} - X_t(k)]^2$ sa variance. Alors, les coefficients de prévision ont donnés par

$$a_{t,k}(i) = \psi_{k+i}, i = 0, 1, \dots$$

et donc

$$X_t(k) = \sum_{i=0}^{\infty} \psi_{k+i} \epsilon_{t-i}, \text{ l'erreur est}$$

$$e_t(k) = \sum_{i=0}^{k-1} \psi_i \epsilon_{t+k-i}$$

et la variance de ce prédicteur est :

$$V(k) = \sigma^2 \sum_{i=0}^{k-1} \psi_i^2$$

Démonstration : Utilisant le développement linéaire en bruit de X_t , i.e. $X_t = \psi(B)\epsilon_t$ avec $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$ et ψ_i sont les coefficients de la représentation causale. On trouve que les coefficients $a_k(i)$ qui minimisent la variance de l'erreur : $V(k) = \mathbb{E}[X_{t+k} - X_t(k)]^2$ sont $a_{t,k}(i) = \psi_{k+i}$ et l'erreur de prévision peut s'exprimer comme $R_t(k) = \sum_{i=0}^{k-1} \psi_i \epsilon_{t+k-i}$.

Notes :

1. Ce résultat fournit des **intervalles de confiance pour la prévision**.
2. Pour $k=1$, $e_t(1) = \epsilon_t(1)$ et $V(1) = \sigma^2$
3. Ce résultat suggère que pour un processus non stationnaire, la variance de l'erreur de prévision converge vers infini :

$$\lim_{k \rightarrow \infty} V(k) = \infty$$

Pour les modèles $ARMA(p, q)$, au lieu d'utiliser les représentations $MA(\infty), AR(\infty)$, il est plus simple d'utiliser directement la définition du modèle. On verra ça dans le cadre plus général des processus **non stationnaires** $ARIMA(p, d, q)$.

10.4 Exercices : TD 3

1. Soit le processus :

$$(1 - .8B + .16B^2)Y_t = (1 + \theta B)Z_t$$

- (a) Est-ce que ce processus est stationnaire causal? Si oui, obtenez la "représentation ψ " du Y_t par rapport au bruit Z_t et trouvez la fonction d'autocorrélation de Y_t .
- (b) Sous quelles conditions est ce processus inversible? Obtenez la "représentation π " du bruit Z_t en termes de la série. De quel problème on s'aperçoit si le processus n'est pas inversible?
- (c) Donnez la prévision k pas en avant $\hat{Y}_t(k)$, en utilisant les valeurs Y_t, Y_{t-1}, Z_t .

2. Trouvez les coefficients d'autocorrélation et tracez la corrélogramme pour le processus ARMA(1,2)
 $Y_t = 0.6Y_{t-1} + \epsilon_t - 0.3\epsilon_{t-1} - 0.1\epsilon_{t-2}$
3. *Prévision du modèle ARIMA(2,1,0)*

$$(1 - \phi_1 B)(1 - \phi_2 B)(1 - B)X_t = \epsilon_t \quad \text{with } -1 < \phi_1 < \phi_2 < 1.$$

- (a) Vérifiez que :
- (i) $X_t(k) = (1 + \phi_1 + \phi_2)X_t(k-1) - (\phi_1 + \phi_2 + \phi_1\phi_2)X_t(k-2) + \phi_1\phi_2X_t(k-3)$, $k \geq 1$
- (ii) $X_t(k) = A_t + B_t\phi_1^k + (X_t - A_t - B_t)\phi_2^k$, $k = 0, 1, 2, \dots$ pour certaines A_t et B_t qui dépendent seulement du t .
- (b) Trouvez les poids des valeurs pivots X_t, X_{t-1}, X_{t-2} dans la prévision Box-Jenkins de X_t
- (c) Trouvez la limite $\lim_{k \rightarrow \infty} X_t(k)$
4. *Prévisions sous le modèle ARIMA(1,1,1)*

Considérons le processus ARIMA(1,1,1) $(1 - \varphi B)(1 - B)Y_t = (1 + \theta B)\epsilon_t$, avec $-1 < \varphi < 1$ et $-1 < \theta < 1$.

- (a) Montrez que $Y_t(1) = (1 + \varphi)Y_t - \varphi Y_{t-1} + \theta\epsilon_t$ et
 $Y_t(k) = (1 + \varphi)Y_t(k-1) - \varphi Y_t(k-2)$, $k \geq 2$.
- (b) Montrez que $Y_t(k) = A_t + B_t\varphi^k$ pour $k \geq 0$, et trouvez des expressions pour A_t et B_t en termes de $Y_t, Y_{t-1}, \epsilon_t, \varphi$ et θ , en utilisant $Y_t(0)[= Y_t]$ et $Y_t(1)$ du (a) ci-dessus. Montrez que :

$$Y_t(k) = Y_t + \varphi \frac{(1 - \varphi^k)}{1 - \varphi} (Y_t - Y_{t-1}) + \theta \frac{(1 - \varphi^k)}{1 - \varphi} \epsilon_t, \quad k \geq 0.$$

Trouvez la limite $\lim_{k \rightarrow \infty} Y_t(k)$

- (c) Montrez que $Y_t(1) = -\theta Y_{t-1}(1) + (1 + \varphi + \theta)Y_t - \varphi Y_{t-1}$ et
 $Y_t(k) = Y_{t-1}(k+1) + \psi_k \epsilon_t$.
- (d) Montrez que $Y_t(k)$ peut s'exprimer en fonction seulement des valeurs passées de la série.
[Indication : utilisez les π pour vous débarrasser de ϵ_t]
- (e) En utilisant le modèle $(1 - 0.6B)(1 - B)Y_t = (1 + 0.3B)\epsilon_t$ obtenez les prévisions des trois termes suivants de la série :

t :	1	2	3	4	5	6	7	8	9	10
y_t :	14.8	12.4	9.4	7.7	7.3	9.0	10.5	11.2	10.4	11.6

t :	11	12	13	14	15	16	17	18	19	20
y_t :	12.1	11.6	9.9	8.1	6.6	5.4	4.2	5.3	6.8	9.2

5. Considérons le processus ARIMA(1,1,2) :

$$(1 - \alpha B)(1 - B)Y_t = (1 + \theta_1 B + \theta_2 B^2)\epsilon_t$$

où $-1 < \alpha < 1$. Soit $Y_t(k)$ la prévision de Y_{t+k} au temps t .

- (a) Montrez que $Y_t(1) = (1 + \alpha)Y_t - \alpha Y_{t-1} + \theta_1 \epsilon_t + \theta_2 \epsilon_{t-1}$ et trouvez les expressions correspondantes pour $Y_t(2)$ et $Y_t(k)$ pour $k \geq 3$

- (b) Montrez que la fonction de prévision peut s'exprimer sous la forme $Y_t(k) = a_t + b_t \alpha^k$, $k \geq 1$, et donnez la formule de a_t, b_t comme fonctions de $Y_t, Y_{t-1}, \epsilon_t, \epsilon_{t-1}$.
- (c) Montrez que $Y_t(k)$ peut s'exprimer en fonction seulement des valeurs passées de la série.
- (d) Un statisticien a utilisé le modèle ARIMA (1,1,2) décrit ci-dessus pour une série (dénommée prix) qui exprime le prix d'une action à la bourse pour 100 jours consécutifs. En sachant que $Y_{98}(1) = 686,996$ et $Y_{99}(1) = 659,416$ et $\sigma_\epsilon = 2$, calculer les prévisions $Y_{101|100}, Y_{102|100}$ de Y_{101} et Y_{102} et donner les 95% intervalles de confiance associés avec ces prévisions.
- (e) Soit $X_m = \sin(2\pi s_1 m) + \sin(2\pi s_2 m) + \epsilon_m, m = 0, 1, \dots, N$. Donner une représentation autoregressive de ce processus. Calculer la densité spectrale, la variance $\sigma^2 = \mathbb{E}(X_t)^2$, et les covariances.
6. **Projet** : Il serait intéressant de déterminer analytiquement "la tendance asymptotique", i.e. le polynôme d'ordre $d - 1$ vers le quel les prévisions convergent asymptotiquement pour les processus ARIMA(p,d,q).
 Considerons par exemple ARIMA(p,2,0); ce modèle inclue une tendance lineaire, pour la quelle le premier candidat est la droite par les deux dernières points pivots (comme dans le cas $p = 0$, quand les pivots coincident avec les valeurs X_t, X_{t-1}). En général, les prévisions doivent encore converger asymptotiquement vers une droite. Pour $p = 0$, on commence déjà exactement sur la "droite de tendance" (due à l'absence d'autres racines dans la partie autorégressive); mais, pour $p \geq 1$, nous serons obligé de tenir compte d'autres valeurs pivots et donc de X_{t-2}, X_{t-3}, \dots . A priori donc, les p points qui precedent les 2 dernières point auront aussi une influence sur la "droite de tendance".

11 Sommaire des définitions et résultats dans les séries temporelles

11.1 Filtres

Définition 11.1 a) Une série s_t sera appelée saisonnière de période p ssi

$$\sum_{i=1}^p s_{t+i} = 0 \iff \left(\sum_{i=0}^{p-1} B^i \right) s_t = 0, \quad \forall t \quad (39)$$

b) Une série s_t sera appelée périodique de période p ssi

$$s_{t+p} = s_t \iff (1 - B^p) s_t = 0, \quad \forall t \quad (40)$$

Théorème 11.1 Un filtre $\psi(B)$ annule (ou "enlève") les composantes saisonnières d'ordre p ssi $\psi(z)$ est divisible par $1 + z + \dots + z^{p-1}$ (donc si $\psi(z) = 0$, pour toutes les racine d'ordre p de l'unité, sauf $z = 1$).

Théorème 11.2 L'espace invariant d'un filtre contient les polynômes de degré $\leq p$ ssi 1 est une racine d'ordre au moins $p+1$ de l'équation $\psi(z) = 1$, i.e. $\psi(1) = 1, \psi'(1) = 0, \psi''(1) = 0, \psi^{(p)}(1) = 0$.

11.2 Causalité et inversibilité des modèles ARMA(p,q)

Théorème 11.3 a) Un processus ARMA(p) $\varphi(B)Y_t = \theta(B)\epsilon_t$ est causal(inversible), i.e. il peut être représenté sous la forme : $Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ où $\sum \psi_i^2 < \infty$

($\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$ où $\sum \pi_i^2 < \infty$) ssi toutes les racines de son polynôme caractéristique $\varphi(z)$ ($\theta(z)$) sont à l'extérieur du cercle unitaire. Les coefficients ψ_i (π_i) sont dans ce cas les coefficients de la série Taylor de $\psi(z) = \frac{\theta(z)}{\varphi(z)}$ ($\pi(z) = \frac{\varphi(z)}{\theta(z)}$)

Théorème 11.4 (*) a) Pour un procesus ARMA(p,q) $\varphi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\varphi(z)$ à l'extérieur du cercle unitaire, les coefficients $\psi_i = \sigma^{-2} \mathbb{E}Y_t \epsilon_{t-i}$ de la représentation causale $Y_t = \sum \psi_i \epsilon_{t-i}$ satisfont la récurrence

$$\begin{aligned} \psi_0 &= 1, \quad \psi_k = \theta_k + \sum_{i=1}^{\min[k,p]} \varphi(i) \psi(k-i), \quad 1 \leq k \leq q \\ \psi_k &= \sum_{i=1}^{\min[k,p]} \varphi(i) \psi(k-i), \quad k > q \end{aligned}$$

Note : Cette dernière équation est appelée **recursion de Yule-Walker**.

b) Pour un procesus ARMA(p,q) $\varphi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\theta(z)$ à l'extérieur du cercle unitaire, les coefficients π_i de la représentation inverse $\epsilon_t = \sum \pi_i Y_{t-i}$ satisfont la récurrence

$$\begin{aligned} \pi_0 &= 1, \quad \pi_k = -\varphi_k + \sum_{i=1}^{\min[k,q]} \theta(i) \pi(k-i), \quad 1 \leq k \leq p \\ \pi_k &= \sum_{i=1}^{\min[k,q]} \theta(i) \pi(k-i), \quad k > p \end{aligned}$$

Exemple 11.1 Ainsi, $\psi_1 = \varphi_1 + \theta_1, \psi_2 = \varphi_2 + \varphi_1^2 + \theta_1\varphi_1 + \theta_2, \psi_3 = \varphi_3 + 2\varphi_1\varphi_2 + \varphi_1^3 + (\varphi_2 + \varphi_1^2)\theta_1 + \varphi_1\theta_2 + \theta_3, \dots$

11.3 Équations Yule Walker

1. Les premières p corrélations $\boldsymbol{\rho} = (\rho(1), \dots, \rho(p))$ peuvent être obtenues du système réduit à p équations et p inconnues :

$$\mathbf{R} \boldsymbol{\varphi} = \boldsymbol{\rho} \quad (41)$$

ou R est la matrice Toeplitz symétrique :

$$\mathbf{R} = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \dots & \dots & \dots & \dots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix}$$

2. En suite, pour $k > p$ on utilise la récurrence : $\rho_k = \sum_{i=1}^p \varphi_i \rho_{k-i}$
3. La variance est $\gamma_0 = \frac{\sigma^2}{1 - \sum_i \varphi_i \rho_i}$ (et en suite, on obtient les covariances par $\gamma_k = \rho_k \gamma_0, k > 1$).
4. **Exemple : AR(2)** Pour le processus AR(2)

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \epsilon_t$$

avec racines λ_1, λ_2 de $0 = \lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 \varphi(\lambda^{-1})$ dedans le cercle unitaire (pour assurer la causalité), on obtient, en résolvant (41) :

$$\rho_1 = \frac{\varphi_1}{1 - \varphi_2} = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 \lambda_2}, \quad \rho_2 = \varphi_2 + \frac{\varphi_1^2}{1 - \varphi_2}, \dots$$

- (a) Si les racines λ_1, λ_2 sont distinctes, on obtient

$$\rho_k = \rho_1 \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} - \lambda_1 \lambda_2 \frac{\lambda_1^{k-1} - \lambda_2^{k-1}}{\lambda_1 - \lambda_2}$$

Finalement, en utilisant $\varphi_1 = \lambda_1 + \lambda_2, \varphi_2 = -\lambda_1 \lambda_2$, on arrive à

$$\rho_k = \frac{(1 - \lambda_2^2) \lambda_1^{k+1} - (1 - \lambda_1^2) \lambda_2^{k+1}}{(\lambda_1 - \lambda_2)(1 + \lambda_1 \lambda_2)} = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{(\lambda_1 - \lambda_2)(1 + \lambda_1 \lambda_2)} - \lambda_1^2 \lambda_2^2 \frac{\lambda_1^{k-1} - \lambda_2^{k-1}}{(\lambda_1 - \lambda_2)(1 + \lambda_1 \lambda_2)}, \quad k \geq 0$$

en termes de racines seulement.

- (b) Dans le cas de racines *confondues* $\lambda_i = \lambda$, on obtient :

$$\rho_k = \left\{ 1 + \left(\frac{1 - \lambda^2}{1 + \lambda^2} \right) k \right\} \lambda^k, \quad k \geq 0.$$

11.4 Prédiction linéaire

$$\hat{Y}_t(k) = \mathbb{E}[Y_{t+k} | F_t] = \sum_{i=1}^p \tilde{\varphi}_i \hat{Y}_t(k-i) + \sum_{i=k}^q \theta_i \hat{\epsilon}_{t+k-i}$$

Dans le cas d'un modèle ARIMA(p,d,q), les $\tilde{\varphi}_i$ sont les coefficients du polynôme $\varphi(B)(1-B)^d$, et dans le cas d'un modèle ARMA(p,q) $\tilde{\varphi}_i = \varphi_i$. Les $\hat{\epsilon}_t$ peuvent être enlevés en utilisant la représentation "π" du bruit en fonction de la série. La récurrence satisfaite par les π_i est $\varphi(B)\hat{Y}_t(k) = \theta(B)\epsilon_{t+k}$ - voir Thm 8.2.

12 Examens d'entraînement

13 Examen d'entraînement 1

Examen : 10 Janvier 2005, UPPA, Dépt. de Mathématiques
Module : Séries temporelles : TMUZ44U
Durée : deux heures

1. Déterminer une moyenne mobile causale $\theta(B) = \sum_{i=0}^q \theta_i B^i$ d'ordre q minimal, qui laisse passer une tendance quadratique sans distortion et qui enlève les composantes saisonnières d'ordre 4.
2. Soit $X_t = \sum_{i=0}^3 \theta_i \epsilon_{t-i}$ un processus moyenne mobile, utilisé pour lisser un bruit blanc $\epsilon_t = BB(0, \sigma^2 = 1)$.
 - a) Quelle est la variance de X_t ?
 - b) Trouvez le filtre (i.e. les coefficients $(\theta_0, \theta_1, \theta_2, \theta_3)$) qui ramènent à un processus X_t à variance minimale, parmi toutes les moyennes mobiles d'ordre 3 qui laissent passer une tendance affine $X_t = a + bt$ sans distortion.
3. *Stationarité des processus* Soit

$$X_{t+1} = \phi X_t + \epsilon_{t+1}, t = 0, 1, 2, \dots,$$

où ϵ_t est bruit blanc Gaussien de distribution $N(0, \sigma^2)$, un processus $AR(1)$ défini seulement à partir du temps $t = 0$. Supposons aussi que X_0 est Gaussien de distribution $N(0, \sigma_0^2)$. Déterminez la valeur de σ_0^2 qui assure que le processus X_t est stationnaire.

4. Considérez le **modèle ARIMA(1,1,0)**

$$(1 - \lambda B)(1 - B)X_t = \epsilon_t$$

où $|\lambda| < 1$ et ϵ_t est bruit blanc Gaussien de distribution $N(0, 1)$.

- (a) Donnez une formule de récurrence et ensuite une formule générale pour les coefficients ψ_k de sa représentation causale $X_t = \sum_k \psi_k \epsilon_{t-k}$.
- (b) Est-ce que ce processus est stationnaire causal ?
- (c) Donner une formule de récurrence et ensuite une formule générale pour les autocorrélations ρ_k du processus X_t . Est-ce que $\sum_k \rho_k < \infty$?
- (d) Donnez une formule de récurrence et ensuite une formule générale pour la prévision Box-Jenkins $\hat{X}_t(k) = \mathbb{E}[X(t+k)|X_t]$ de $X(t+k)$, $k = 0, 1, 2, \dots$
- (e) Trouvez la limite $\lim_{k \rightarrow \infty} \hat{X}_t(k)$.

Solutions :

1. En demandant que $\theta(B) = \frac{1+B+B^2+B^2}{3}(a_0 + a_1B + a_2B^2)$ satisfait $\theta(1) = 1, \theta'(1) = 0, \theta''(1) = 0$, on trouve : $\theta(B) = \frac{15-16B+5B^2}{4} \frac{1+B+B^2+B^3}{4} = \frac{15}{16} - \frac{1}{16}B + \frac{4}{16}B^2 - \frac{4}{16}B^3 - \frac{11}{16}B^4 + \frac{5}{16}B^5$.
2. a) $\text{Var } X_t = \sum_i \theta_i^2$. b) Elle est minimisée sous les contraintes $\sum_i \theta_i = 1, \sum_i i\theta_i = 0$ par $(7/10, 4/10, 1/10, -2/10)$.
3. $\sigma_0^2 = \frac{\sigma^2}{1-\varphi^2}$
4. (a) Les coefficients $\psi(k)$ satisfont l'équation Yule Walker

$$\psi_k = \varphi_1\psi_{k-1} + \varphi_2\psi_{k-2}$$

avec valeurs initiales $\psi_0 = 1, \psi_1 = \varphi_1$.

La solution générale de cette récurrence d'ordre 2, pour valeurs initiales arbitraires x_0, x_1 et avec des racines distinctes λ_1, λ_2 , est :

$$\psi(k) = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} x_1 - \frac{(\lambda_1^k \lambda_2 - \lambda_2^k \lambda_1)}{\lambda_1 - \lambda_2} x_0$$

(et avec des racines non-distinctes $\lambda_1 = \lambda_2 = \lambda$ nous avons $\psi(k) = k\lambda^{k-1}x_1 + (1 - k\lambda^k)x_0$).

Ici, avec $\lambda_2 = 1$, on obtient par cette formule, (avec $x_0 = 1, x_1 = \varphi_1$) $\psi(k) = \frac{\lambda^{k+1}-1}{\lambda-1}$ (et avec des racines non-distinctes nous avons $\psi(k) = (k+1)\lambda^k$).

b) Le processus n'est pas stationnaire causal car la racine 1 du polynôme caractéristique n'est pas à l'extérieur du disque unité.

(c) Pour un processus AR stationnaire X_t , la fonction d'autocorrelation satisfait l'équation Yule Walker

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2}$$

Pour $k = 1$ ça donne :

$$\rho_1 = \frac{\varphi_1}{1 - \varphi_2} = \frac{1 + \lambda}{1 + \lambda} = 1$$

et ensuite $\rho_k = 1$. Le seul processus stationnaire Gaussien satisfaisant cette récurrence est $X_t = X_{t-1} = \dots$, quand $\sigma = 0$. Pour $\sigma \neq 0$, il n'existe pas de processus stationnaire satisfaisant notre récurrence.

d) Avec la même récurrence, on arrive à :

$$X_t(k) = \frac{\lambda^{k+1} - 1}{\lambda - 1} X_t + \frac{\lambda^{k+1} - \lambda}{\lambda - 1} X_{t-1}$$

13.1 Examen d'entraînement 3

2. On considère le processus aléatoire suivant :

$$X_t = 20 - .4 X_{t-1} + .12 X_{t-2} + \epsilon_t$$

ou ϵ_t est $BB(0, \sigma^2 = 1)$.

- (a) Calculer l'espérance de X_t , en supposant que le processus est stationnaire.
 - (b) Est-ce que le processus est stationnaire causal ?
 - (c) Donner les équations de Yule-Walker du processus, calculer la variance, ainsi que les 3 premières valeurs des autocorrélations.
 - (d) Pour un modèle non-stationnaire, comment on pourrait détecter la manque de stationnarité à partir des expressions numériques fournies par les équations de Yule-Walker pour les autocorrélations ?
 - (e) Calculer les 3 premières autocorrélations partielles.
3. Soit le processus :

$$(1 - .8B + .16B^2)Y_t = (1 + \theta B)\epsilon_t$$

- (a) Est-ce que ce processus est inversible ? Calculez la formule générale des coefficients $\pi(k)$ de la "représentation π " du bruit ϵ_t en termes de la série. Quels sont les premiers quatre coefficients De quel problème on s'aperçoit si le processus n'est pas inversible ?
 - (b) Donnez la prévision un pas en avant $\hat{Y}_t(1)$ qui utilise les trois valeurs $Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3}$.
 - (c) Est-ce que ce processus est stationnaire causal ? Si oui, trouvez la fonction d'autocorrélation de Y_t .
4. Considérons le processus :

$$(1 - B)Y_t = (1 + \beta_1 B + \beta_2 B^2)\epsilon_t$$

Soit $Y_t(k)$ la prévision de Y_{t+k} au temps t .

- (a) Trouvez des expressions pour $Y_t(1), Y_t(2)$, comme fonctions de $Y_t, Y_{t-1}, \epsilon_t, \epsilon_{t-1}$. Donnez une formule de récurrence pour $Y_t(k)$ pour $k \geq 3$, et écrivez ensuite $Y_t(k)$ comme une fonction de $Y_t, Y_{t-1}, \epsilon_t, \epsilon_{t-1}$.
 - (b) Trouvez des expressions pour $Y_t(1), Y_t(2)$, comme fonctions de $Y_t, Y_{t-1}, Y_{t-2}, \dots$ en utilisant la représentation π du bruit.
5. (a) Donnez les formules des coefficients de corrélation ρ_1, ρ_2 pour le processus $MA(1)$.
- (b) Trouvez les valeurs maximales et minimales de ρ_1 et les valeurs de θ pour les quelles ces valeurs sont atteintes.

Solutions :

1. (c) Le processus peut s'écrire $Y_t = (1 + \tau_1 B + \tau_2 B^2)\epsilon_t = (1 + .4B)^2\epsilon_t$. Il est inversible car la racine $-5/2$ est à l'extérieur du disque unité. Par identification des coefficients on trouve que $\pi_1 = \theta_1, \pi_2 = \theta_2 - \theta_1^2, \pi_3 = \theta_1^3 - 2\theta_2\theta_1, \pi_4 = -\theta_1^4 + 3\theta_2\theta_1^2 - \theta_2^2, \dots$ et alors

$$\hat{Y}_t(1) = Y_t + \sum_{i=1} \pi_i Y_{t-i}$$

avec $\pi_1 = .8, \dots$

3. (a) Ce processus AR(2) pas centré peut s'écrire $(1 - .2B - .35B^2)X_t = (1 - .7B)(1 + .5B)X_t = 40 + \epsilon_t$. En prenant l'espérance on trouve $\mathbb{E}(X_t) = \frac{40}{1 - .2 - .35} = \frac{40}{.45}$.
- (b) Le processus est stationnaire causal car les racines du polynôme sont à l'extérieur du disque unité.
- (c) Soit $Y_t = X_t - \mathbb{E}X_t$. Alors Y_t est un processus AR(2) centré satisfaisant $(1 - .2B - .35B^2)Y_t = (1 - .7B)(1 + .5B)X_t = \epsilon_t$
- La fonction d'autocovariance de Y_t (qui est la même que celle de X_t) est obtenue de l'équation : $\mathbb{E}(Y_t Y_{t-h}) = 0.2\mathbb{E}(Y_{t-1} Y_{t-h}) + 0.35\mathbb{E}(Y_{t-2} Y_{t-h}) + \mathbb{E}(\epsilon_t Y_{t-h})$, ce qui donne le système de Yule Walker :

$$\begin{aligned}\rho_1 &= 0.2 + 0.35\rho_1 \\ \rho_2 &= 0.2\rho_1 + 0.35\end{aligned}$$

La première équation donne $\rho_1 = \frac{2}{6.5} = .31$, et la seconde donne $\rho_2 = 0.2\rho_1 + .35 = .41$. Finalement, $\gamma_0 = \frac{1}{1 - \sum \varphi_i \rho_i} = 2.52$.

(d) Les autocorrélations partielles $\hat{\rho}_i, i = 1, 2, 3$ se calculent à l'aide des déterminants des matrices d'autocorrélations et sont $.31, .35, \approx 0$. La troisième autocorrélation est en effet une erreur d'arrondissement, car les autocorrélations partielles sont nulles au delà du rang p .

13.2 Examen d'entraînement 4

- Soit Y_t un processus ARMA(1,1) vérifiant l'équation $Y_t = -0.5Y_{t-1} + \epsilon_t - 0.8\epsilon_{t-1}$, avec ϵ_t un bruit blanc.
 - Précisez si le processus est stationnaire, causal et inversible, et calculez sa fonction d'autocovariance.
 - Trouvez les coefficients ψ_j de sa représentation comme processus $MA(\infty)$ et les coefficients π_j de sa représentation comme processus $AR(\infty)$ et précisez si ces représentations sont convergentes.
- Considérons le processus : ARIMA(1,1,2) :

$$(1 - \alpha B)(1 - B)Y_t = (1 + \theta_1 B + \theta_2 B^2)Z_t$$

où $-1 < \alpha < 1$. Soit $Y_t(k)$ la prévision de Y_{t+k} au temps t .

- Montrez que $Y_t(1) = (1 + \alpha)Y_t - \alpha Y_{t-1} + \theta_1 Z_t + \theta_2 Z_{t-1}$ et trouvez une expression correspondante pour $Y_t(2)$. Donnez une formule de récurrence pour $Y_t(k)$ pour $k \geq 3$.
 - Montrez que la fonction de prévision $Y_t(k)$ peut s'exprimer sous la forme $Y_t(k) = a_t + b_t \alpha^k$, et donnez les formules de a_t, b_t comme fonctions de $Y_t, Y_{t-1}, Z_t, Z_{t-1}$.
- On considère le processus aléatoire AR(2) suivant :

$$X_t = 10 + (-.3) X_{t-1} + .01 X_{t-2} + \epsilon_t$$

ou ϵ_t est $BB(0, \sigma^2 = 1)$.

- (a) Calculer l'espérance de X_t , en supposant que le processus est stationnaire.
- (b) Est-ce que le processus est stationnaire causal ?
- (c) Donner les équations de Yule-Walker du processus, calculer la variance, ainsi que les 3 premières valeurs des autocorrélations.
- (d) Pour un modèle non-stationnaire, comment on pourrait détecter la manque de stationnarité à partir des expressions numériques fournies par les équations de Yule-Walker pour les autocorrélations ?
- (e) Calculer les 3 premières autocorrélations partielles.

4. *Processus ARMA(2, 1)* Soit le processus :

$$(1 - B + .25B^2)Y_t = (1 + \theta B)Z_t$$

- (a) Est-ce que ce processus est stationnaire causal? Si oui, trouvez la fonction d'autocorrélation de Y_t .
- (b) Est-ce que ce processus est inversible? Calculez les premiers cinq coefficients de la "représentation π " du bruit Z_t en termes de la série. De quel problème on s'aperçoit si le processus Y_t n'est pas inversible ?
- (c) Donnez, si possible, une méthode de prévision pour un processus (i) causal et noninversible (ii) noncausal et inversible.
- (d) Donnez la prévision un pas en avant $\hat{Y}_t(1)$ en fonction des valeurs passées Y_t, Y_{t-1}, \dots

5. Considérons le processus : ARIMA(1,1,2) :

$$(1 - \alpha B)(1 - B)Y_t = (1 + \theta_1 B + \theta_2 B^2)Z_t$$

où $-1 < \alpha < 1$. Soit $Y_t(k)$ la prévision de Y_{t+k} au temps t .

- (a) Montrez que $Y_t(1) = (1 + \alpha)Y_t - \alpha Y_{t-1} + \theta_1 Z_t + \theta_2 Z_{t-1}$ et trouvez une expression correspondante pour $Y_t(2)$. Donnez une formule de récurrence pour $Y_t(k)$ pour $k \geq 3$.
- (b) Montrez que la fonction de prévision $Y_t(k)$ peut s'exprimer sous la forme $Y_t(k) = a_t + b_t \alpha^k$, et donnez les formules de a_t, b_t comme fonctions de $Y_t, Y_{t-1}, Z_t, Z_{t-1}$.

6. On considère la série suivante :

t_i	1	2	3	4	5	6	7	8	9	10
y_i	8	4	13	15	18	15	19	19	20	28

- a) Représenter graphiquement cette série.
- b) On se propose d'ajuster une droite de régression

$$f(t) = at + b$$

de moindre carrées. Déterminer les coefficients a et b .

- c) Quel est le point le plus "éloigné" de $f(t)$?
- c) Calculez une droite de régression $f(t)$ en imposant $b \geq 6$.

7. On considère une série $(y_i, 1 \leq i \leq n)$ périodique, de période p . On suppose que le nombre d'observations n est un multiple de p : $n = Lp$. Calculez les corrélations empiriques :

$$\rho(p), ; \rho(2p), ; \dots ; \rho(jp) \dots$$

en utilisant la définition des corrélations empiriques :

$$\rho(k) = \frac{\frac{1}{n-k} \sum_{i=1}^{n-k} (y_i - \bar{y})(y_{i+k} - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n-k} (y_i - \bar{y})(y_{i+k} - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

où $\bar{y} = (\sum_{i=1}^n y_i) / n$ est la moyenne de l'ensemble de la série $(y_i, 1 \leq i \leq n)$.

Commentez sur le résultat.

8. Quelle est la limite $\lim_{k \rightarrow \infty} X_t(k)$ de la prévision linéaire d'un processus ARMA (p,q) ?

Solutions :

1. (c) Le processus peut s'écrire $Y_t = (1 + \tau_1 B + \tau_2 B^2) Z_t = (1 + .4B)^2 Z_t$. Il est inversible car la racine $-5/2$ est à l'extérieur du disque unité. Par identification des coefficients on trouve que $\pi_1 = \theta_1, \pi_2 = \theta_2 - \theta_1^2, \pi_3 = \theta_1^3 - 2\theta_2\theta_1, \pi_4 = -\theta_1^4 + 3\theta_2\theta_1^2 - \theta_2^2, \dots$ et alors

$$\hat{Y}_t(1) = Y_t + \sum_{i=1} \pi_i Y_{t-i}$$

avec $\pi_1 = .8, \dots$

13.3 Examen d'entraînement 2

1. Soit le processus :

$$(1 - B + .25B^2)Y_t = (1 + \theta B)Z_t$$

- (a) Est-ce que ce processus est stationnaire causal? Si oui, trouvez la fonction d'auto-corrélation de Y_t .
- (b) Est-ce que ce processus est inversible? Calculez les premiers cinq coefficients de la "représentation π " du bruit Z_t en termes de la série. De quel problème on s'aperçoit si le processus Y_t n'est pas inversible ?
- (c) Donnez, si possible, une méthode de prévision pour un processus (i) causal et noninversible (ii) noncausal et inversible.
- (d) Donnez la prévision un pas en avant $\hat{Y}_t(1)$ en fonction des valeurs passées Y_t, Y_{t-1}, \dots
2. Soit Y_t un processus ARMA(2,1) vérifiant l'équation $Y_t = 0.7Y_{t-1} - 0.1Y_{t-2} + \epsilon_t - 2\epsilon_{t-1}$, avec ϵ_t un bruit blanc.
- (a) Précisez si le processus est stationnaire, causal et inversible, et calculez sa fonction d'autocovariance.
- (b) Trouvez les coefficients ψ_j de sa représentation comme processus $MA(\infty)$ et les coefficients π_j de sa représentation comme processus $AR(\infty)$ et précisez si ces représentations sont convergentes.
3. *Prévision du modèle ARIMA(1,2,0)* $(1 - \varphi B)(1 - B)^2 X_t = \epsilon_t$.

- (a) Trouvez la forme de la prévision Box-Jenkins $X_t(k)$ de X_t et les poids des points pivotales X_t, X_{t-1}, X_{t-2} .
 - (b) Trouvez la limite $\lim_{k \rightarrow \infty} X_t(k)$.
 - (c) Calculez les coefficients ψ_k de l'expansion causale de ce processus.
4. Calculer la fonction d'autocovariance $\Gamma_k = \mathbb{E}Y_{n+k}Y_n^t$ du processus à valeurs en \mathbb{R}^2 :

$$Y_n = \begin{pmatrix} a_0 \varepsilon_n + a_1 \varepsilon_{n-1} \\ b_1 \varepsilon_{n-1} + b_2 \varepsilon_{n-2} \end{pmatrix}$$

où ε_n est un bruit blanc standard.

5. *Stationarité des processus*

Investiguer si les processus suivants sont stationnaires de second ordre, en calculant leurs caractéristiques statistiques de premier et second ordre (moyennes, variances, covariances). On convient de noter par $\{\varepsilon_t\}$ le bruit blanc.

- (a) $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$ (MA(1))
- (b) $Y_t = Y_{t-1} + c + \varepsilon_t$ (marche aléatoire avec tendance)
- (c) $Y_t = \varphi Y_{t-1} + c + \varepsilon_t$ ($\varphi \neq 1$) (AR(1) avec tendance)

6. *Processus AR(2)* On considère le processus aléatoire AR(2) suivant :

$$X_t = 10 + (-.3) X_{t-1} + .01 X_{t-2} + \varepsilon_t$$

ou ε_t est $BB(0, \sigma^2 = 1)$.

- (a) Calculer l'espérance de X_t , en supposant que le processus est stationnaire.
- (b) Est-ce que le processus est stationnaire causal ?
- (c) Donner la recursion de Yule-Walker pour les prévisions $\hat{X}(t+k|t)$ du processus, et résolvez la en termes des valeurs pivotales X_t, X_{t-1} .

Mêmes questions pour

$$(1 - B + .25B^2)X_t = \varepsilon_t$$

7. *Processus ARMA(2,1)*

Soit le processus :

$$(1 - .6B + .08B^2)Y_t = (1 + \theta B)\varepsilon_t$$

- (a) Est-ce que ce processus est stationnaire causal ? Si oui, obtenez la "représentation ψ " du Y_t par rapport au bruit Z_t .
- (b) Sous quelles conditions est ce processus inversible ? Obtenez la "représentation π " du bruit Z_t en termes de la série. De quel problème on s'aperçoit si le processus n'est pas inversible ?
- (c) Donnez la prévision k pas en avant $\hat{Y}_t(k)$, en utilisant les valeurs $Y_t, Y_{t-1}, \varepsilon_t$.

8. Montrez que pour ARIMA(1,2,0),

$$\psi_k = \{\varphi^{k+2} + (1 - 2\varphi) + k(1 - \varphi)\}(1 - \varphi)^{-2}, k \geq 0$$

14 TP Series Temporelles

Ce chapitre est basé sur le livre *Introductory Time Series with R*, by Paul S.P. Cowpertwait and Andrew V. Metcalfe voir <http://ebookey.org/go/?u=http://www.massey.ac.nz/pscowper/ts>

Jeux de données disponibles : `data()`

Commandes : `window,decompose,summary,layout,acf,pacf,ar`

Aide : `?decompose`

Directoire courant : `getwd()`

Exemples :

```
## example taken from Kendall/Stuart
x <- c(-50, 175, 149, 214, 247, 237, 225, 329, 729, 809, 530, 489, 540, 457, 195, 176, 337,
x <- ts(x, start = c(1951, 1), end = c(1958, 4), frequency = 4)
m <- decompose(x)
plot(m)

require(graphics)
m <- decompose(co2)
names(m)
plot(m)
mr<-window(m$random,start=c(1960,1),end=c(1996,4))
summary(mr)
layout(1:2)
acf(mr)
pacf(mr)
mrs<-as.ts(mr)
mr.ar <- ar(mrs, method = "mle")
mr.ar
```

14.1 Tendance et composante périodique

Commandes : `aggregate, boxplot`

```
data(AirPassengers)
AP<-AirPassengers
AP
class(AP)
plot(AP)
frequency(AP)
layout(1:2)
plot(aggregate(AP)) ##applique filtre "somme"
boxplot(AP~cycle(AP))
layout(1:2)
plot(AP)
plot(log(AP))
AP.decom <- decompose(AP, "multiplicative")
plot(ts(AP.decom$random[7:138]))
layout(1:2)
```

```

acf(AP.decom$random[7:138])
pacf(AP.decom$random[7:138])

www<-"http://www.massey.ac.nz/~pscowper/ts/Maine.dat"
Maine.month<-read.table(www,header=TRUE)
attach(Maine.month)
class(Maine.month)
Mts<-ts(unemploy,start=c(1996,1),freq=12)
Mtsa<-aggregate(Mts)/12 ##filtre "moy. arithm"
layout(1:2)
plot(Mts)
plot(Mtsa)
M.feb<-window(Mts,start=c(1996,2),freq=TRUE)
Febratio<-mean(M.feb)/mean(Mts)
Febratio

```

14.2 Arima, Regression

Basé sur http://www.statoek.wiso.uni-goettingen.de/veranstaltungen/zeitreihen/sommer03/ts.r_intro.pdf
 Commandes : lm, resid, factor, confint, arima

```

data(LakeHuron)
LH<-arima(LakeHuron,order=c(1,0,1))
tsdiag(LH)
Box.test(LH$residuals,lag=1)
LH.pred<-predict(fit,n.ahead=8)
plot(LakeHuron,xlim=c(1875,1980),ylim=c(575,584))
LH.pred<-predict(fit,n.ahead=8)
lines(LH.pred$pred,col="red")
lines(LH.pred$pred+2*LH.pred$se,col="red",lty=3)
lines(LH.pred$pred-2*LH.pred$se,col="red",lty=3)

www <-"http://www.massey.ac.nz/~pscowper/ts/global.dat"
Glob <- scan(www)
Glob.ts <- ts(Global, st = c(1856, 1), end = c(2005, 12),fr = 12)
plot(Glob.ts)
Gl.an<- aggregate(Glob.ts,FUN= mean)##"moy.arithm"
plot(Gl.an)
temp <- window(Glob.ts, start = 1970)
td<-decompose(temp)
tdr<-ts(td$random,start=c(1970,7),end=c(2004,12),fr=12)
acf(tdr)
pacf(tdr)
tds.ar <- ar(as.ts(tdr), method = "mle")
tdrr<-tds.ar$resid
tempc<-temp-tdrr
m<-decompose(tempc)

```

```

plot(m)
temp.lm <- lm(temp ~ time(temp))
coef(temp.lm)
summary(temp.lm) ## SE=.12, R^2=.67
acf(resid(temp.lm))
pacf(resid(temp.lm))
Seas <- cycle(temp)
fs<-factor(Seas)
Time <- time(temp)
temp.lm <- lm(temp ~ 0 + Time + factor(Seas))
coef(temp.lm)
confint(temp.lm)
new.t <- seq(2006, len = 2 * 12, by = 1/12)
new.dat<-data.frame(Time= new.t,Seas =rep(1:12,2))
pr<-predict(temp.lm, new.dat)[1:24]
plot(pr)

```

14.3 Projets

1. Effectuez une analyse **Box-Jenkins** d'un jeu de données : par exemple "WWWusage" en R (qui représente le nombre d'utilisateurs connectés à un serveur Internet chaque minute), en utilisant les commandes "acf(x)", "pacf(x)" (ou "eda.ts(x)", si disponible) "arima()", "pnorm()" et "eda.ts(x\$res)". L'analyse devrait aboutir dans un modèle avec résidus tel qu'au plus 1 sur 20 des coefficients acf et pacf sortent de l'intervalle de confiance autour de 0, et avec p-valeurs des coefficients inférieures à .05 (rappel : p-val $\approx \mathbb{P}\{|t_{val}| \geq 2\}$, où-val sont les valeurs "standardisées", i.e. divisées par l'erreur standard (s.e.)).
2. Répétez, avec un jeu de données de votre choix (à trouver à partir de la liste "data(package=NULL)" ou "data(package="ts)").
3. Ecrivez des programmes qui **simulent** (sans utiliser la commande arima.sim) des processus :
 - a) MA(2) avec $\theta_0 = 1$, les autres coefficients à choisir, et à bruit petit, b) AR(2) à bruit petit, c) ARMA(2,2) (en passant la sortie du premier program au deuxième), et d) ARIMA(2,1,2) (en appliquant cumsum à la sortie du program antérieur).

Enoncez les théorèmes satisfaits par l'acf et le pacf des premiers deux cas, et vérifiez ensuite que vos programmes produisent des résultats adéquats.

Pour le troisième et quatrième cas, **estimez** le modèle par la commande arima, avec les ordres simulés, et aussi avec des ordres plus grands. Est-ce que la commande retrouve les coefficients que vous aviez choisi quand le bruit est très petit (en supposant que l'analyse est bonne, et donc que l'acf et pacf des résidus indiquent un bruit blanc) ?

4. Interpretation de l'acf

- (a) Soit x un vecteur de nombres consecutifs. Simulez une série a) linéaire $y = ax$
 - b) quadratique $y = ax^2 + bx$
 - c) périodique $y = \sin(ax)$
 - d) "presque périodique" $y = \sin(ax) + \sin(bx)$
- en donnant deux exemples de a, b pour chaque problème (donc 12 exemples). Obtenez l'acf de toutes les séries et indiquez vos observations sur la dépendance de a, b .

- (b) Pour les séries avec acf non zero, indiquez quels filtres/transformations pourront nous amener aux résidus bruit blanc.
- (c) Démontrez sur trois des exercices antérieurs l'effet sur l'acf de l'ajout du bruit blanc d'écart type $\sigma = R/4$, $\sigma = R$ et $\sigma = 3R$, où $R = \max y_i - \min y_i$ est "l'écart du signal déterministe".

Inclure au moins un exemple qui a besoin du filtrage, et étudier encore une fois l'effet du même filtrage.