

Séries temporelles, avec R

Florin Avram

Objectif : La régression et l'interpolation déterministe sont parmi les méthodes les plus importantes en statistique et dans les mathématiques appliquées.

Leur but est d'estimer la valeur d'un "signal" $g(x)$ en un point x quelconque, en connaissant des observations Y_i bruitées du signal, observées dans un ensemble fini des points x_i :

$$Y_i = g(x_i) + \epsilon_i, \quad i = 1, \dots, I.$$

On veut à la fois : a) diminuer le bruit et b) "extrapoler" du connu à l'inconnu.

Domaines d'application :

- Estimation des densités (il suffit d'appliquer le modèle ci dessus aux observations d'une fonction de survie empirique, en ajoutant la contrainte que $g(x)$ est non-croissante).
- Séries temporelles, appliquées en économétrie, économie, finances, météo, médecine
- Traitement du signal
- Statistiques spatiales appliquées à la prospection et exploitation pétrolière et minière
- Imagerie médicale
- Sciences environnementales : océanographie, météorologie, géologie, ...

Références :

1. Yves Aragon, Séries temporelles avec R - Méthodes et cas
2. Notes de cours/TD distribuées en classe
3. Notes WEB de Sylvain Rubenthaler <http://math.unice.fr/rubenthala/cours.html> avec source .tex fournie et utilisée
4. Nonparametric regression by P.J. Green and B.W. Silverman
5. Spline Models for Observational Data, by Wahba, G., (1990), SIAM.
6. Chiles and Delfiner, 1999. Geostatistics, Modeling spatial uncertainty. Wiley séries on probability and statistics.
7. Notes WEB de Diggle et Ribeiro sur geoR, Nychka sur fields.R et notes sur R : <http://pbil.univ-lyon1.fr/R/tdr.html>
8. Notes WEB de X. Guyon http://www.math.u-psud.fr/stafav/IMG/pdf/Statistique_Spatiale.pdf
9. Philipp Pluch : Some Theory for the Analysis of Random Fields - With Applications to Geostatistics <http://front.math.ucdavis.edu/0701.5323>

Table des matières

1	Introduction	5
2	Décomposition des séries temporelles	8
2.1	Les composantes d'une série temporelle	8
2.2	Le bruit blanc	9
2.3	Le modèle AR(1)	10
2.4	Opérateurs de retard et d'avance	12
2.5	Quelques types de décomposition	13
2.5.1	Le modèle additif	13
2.5.2	Le modèle multiplicatif	13
2.5.3	Modèles mixtes	13
2.6	Exercices	14
3	Séries temporelles en R	18
3.1	decompose, residuals, acf, Box.test, stl	18
3.2	Le package caschrono	19
3.3	Lac Huron : lm, arima, armax	21
3.4	Simulation et verification du bruit blanc : rnorm, rlnorm, rt, etc	23
3.5	Simulation et analyse arima de quelques modèles de Aragon, Ch.4	24
3.6	Importation des données, auto.arima, forecast, spectrum . . .	26
3.7	AirPassengers : auto.arima du modèle multiplicatif	26
3.8	Exercices	28
3.9	Un peu de programmation	31
4	Modélisation stochastique des séries temporelles	36
4.1	Introduction	36
4.2	Processus stochastiques stationnaires	37
4.3	Exemples des processus stationnaires	38
4.3.1	Les moyennes mobiles MA(q) : l'exemple le plus simple de processus a composantes dépendantes	39
4.3.2	Les moyennes mobiles MA(∞)	40
4.3.3	Les modèles autorégressifs AR(p) : des processus faciles à prédire	41
4.4	Fonctions de transfert et applications	41
4.4.1	Réurrences infinies et la causalité des modèles AR(p) .	42

4.4.2	La récurrence de Yule-Walker pour les coefficients des fonctions de transfert causales	43
4.5	Inversibilité des processus MA(q)	45
4.6	Équations de Yule-Walker pour les covariances et corrélations d'un processus AR(p) causal	46
4.7	L'estimation des modèles AR(p) par le système de Yule-Walker	49
5	Les modèles ARMA(p,q)	52
5.1	Causalité et inversibilité des modèles ARMA(p,q)	53
5.1.1	Équations de Yule-Walker pour les covariances/corrélations des processus ARMA(p,q)	56
5.2	Les modèles ARIMA(p,d,q)	57
6	Filtres/moyennes mobiles	58
6.1	Filtres qui éliminent/laisser invariants des composantes polynômiales, saisonnières ou périodiques	59
6.2	Moyennes mobiles symétriques	61
6.3	La réduction du bruit des observations	62
6.4	A quoi servent les filtres ?	63
6.5	Récurrences et équations différentielles linéaires	64
6.5.1	L'équation de récurrence linéaire à coefficients constants	64
6.5.2	La méthode des fonctions génératrices(*)	67
6.6	Exercices	69
7	Revison	71
7.1	Exercices	71
7.2	Contrôle continu d'entraînement	74
7.3	Contrôle continu	76
8	Prévision linéaire des modèles autorégressifs ARIMA(p,d,q)	79
8.1	Prévision des processus AR(p) par la récurrence de Yule-Walker	80
8.2	Prévision des processus MA(1)	82
8.3	Prévision des processus ARMA(p,1)	83
8.4	Prévision des processus ARIMA(p,d,0)	84
8.5	Prévision des processus ARIMA(p,d,q)	85
8.6	Intervales de confiance pour la prédiction	86
8.7	HoltWinters, forecast	87
9	Sommaire des quelques définitions et résultats dans les séries temporelles	88
9.1	Filtres	88
9.2	Équations Yule Walker pour AR(p)	88
9.3	Causalité et inversibilité des modèles ARMA(p,q)	89
9.4	Prévision linéaire	90

10 Méthodes et cas, Yves Aragon	91
10.1 Trafic mensuel de l'aéroport de Toulouse-Blagnac	91
10.2 Consommation d'électricité, cf. Méthodes et cas, Yves Aragon	94
10.3 Température mensuelle moyenne à Nottingham, cf. Yves Aragon	96
10.4 Exercices	96
10.5 Modèles d'espace-état	98
11 Examens	100
11.1 Examen 1	100
11.2 Examen 2	102

Chapitre 1

Introduction

Nous allons étudier la prédiction des champs des variables aléatoires $Y(x)$, où $x \in \mathbb{R}^d$ dans la géostatistique, et $x \in \mathbb{N}, \mathbb{Z}$ ou \mathbb{R}_+ pour les séries temporelles/chronologiques. Le cas le plus intéressant est quand $x \in \mathbb{R}^d \times \mathbb{R}_+$, où une des dimensions de x représente le temps (et les autres spatiales), dans quel cas on parle d'interpolation/régression **spatio-temporelle**.

En utilisant des valeurs observées

$$Y_i = g(x_i) + \epsilon_i, \quad i = 1, \dots, I \quad (1.1)$$

et leurs coordonnées x_i , $i = 1, \dots, I$, on se propose d'estimer la fonction $g(x)$ et le champ $Y(x)$ dans un point x arbitraire. La fonction "signal" $g(x)$ devrait appartenir à un espace vectoriel de dimension finie, et la première question est comment choisir cette espace (un des choix favoris pour modéliser le signal $g(x)$ au cas des séries irrégulières, sont les splines). Nous allons commencer notre sujet par le cas le plus simple des observations unidimensionnelles, où on parle des séries temporelles.

Définition 1.0.1 *Une série temporelle (ou chronologique) est une succession des valeurs $(Y_1, Y_2, \dots, Y_n, \dots)$ observées aux moments $(T_1 < T_2 < \dots < T_n, \dots)$ (on a donc une suite des paires $(T_1, Y_1), (T_2, Y_2), \dots, (T_n, Y_n), \dots$).*

Quand les T_i sont en progression arithmétique, on dit que la série est régulière.

Exemple 1.0.1 *La figure 1.1 représente le total mondial des passagers aériens par mois entre 1949 et 1960. Noter que les points sont reliés par des traits (qui sont là pour faire joli et n'ont pas de signification particulière). Les données (**AirPassengers**) sont disponibles dans **R**.*

L'objectif de l'étude des séries temporelles est de faire des prédictions sur l'évolution de la série. Voici une liste non-exhaustive des modèles mathématiques que l'on pourra utiliser :

- Régression. On suppose que x_t est polynomial en t , par exemple $x_t = \alpha_2 t^2 + \alpha_1 t + \alpha_0 + \epsilon_t$ (avec ϵ_t un bruit aléatoire). On estime les coefficients par $\hat{\alpha}_2, \hat{\alpha}_1, \hat{\alpha}_0$ (à partir des valeurs x_1, \dots, x_n). Ainsi, avec la donnée de x_1, \dots, x_n , on fera la prédiction $\hat{x}_{n+1} = \hat{\alpha}_2(n+1)^2 + \hat{\alpha}_1(n+1) + \hat{\alpha}_0$ de la valeur x_{n+1} .

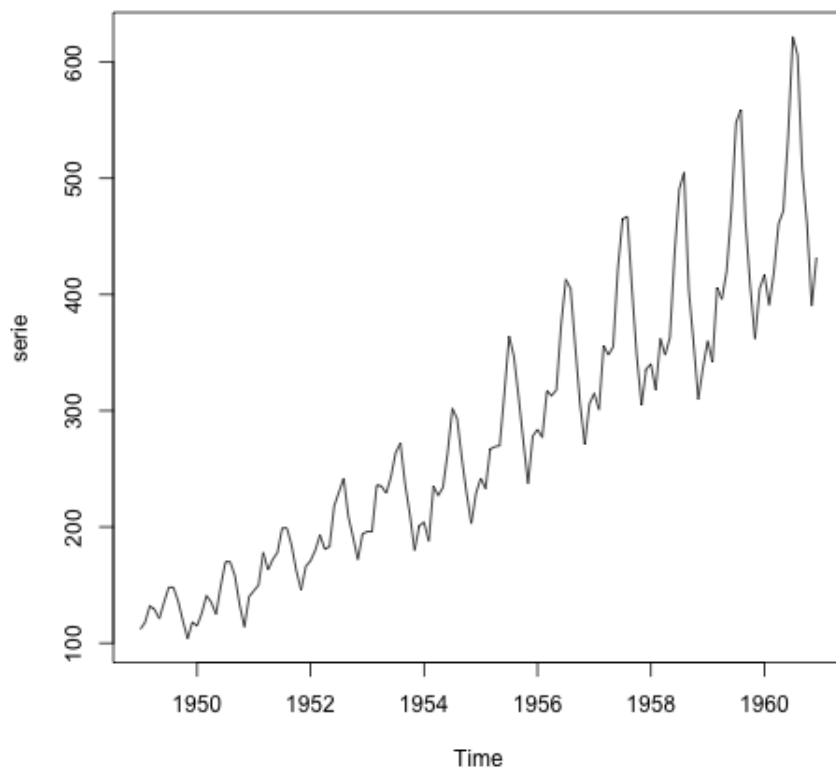


FIGURE 1.1 – AirPassengers

- Lissages exponentiels.
- Modèles ARMA. Ces modèles sont plus lourds numériquement, mais plus performants.

Les défis à relever (dans l'ordre) :

- Définir un modèle avec un nombre fini de paramètres.
- Estimer les paramètres du modèle.
- Vérifier la qualité de l'ajustement du modèle, comparer différents modèles (on pourra découper les données en un échantillon d'apprentissage et un échantillon de test).
- Effectuer des prédictions.

Par rapport aux autres types de données statistiques, la particularité des séries chronologiques tient à la présence d'une relation d'antériorité entre les temps, 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 supposer $T_i \in \mathbb{N}$. **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 âge, entre 55 et 104 (c'est le premier exemple d'utilisation

de splines, par Whittaker (1923)).

La figure 1.2 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.

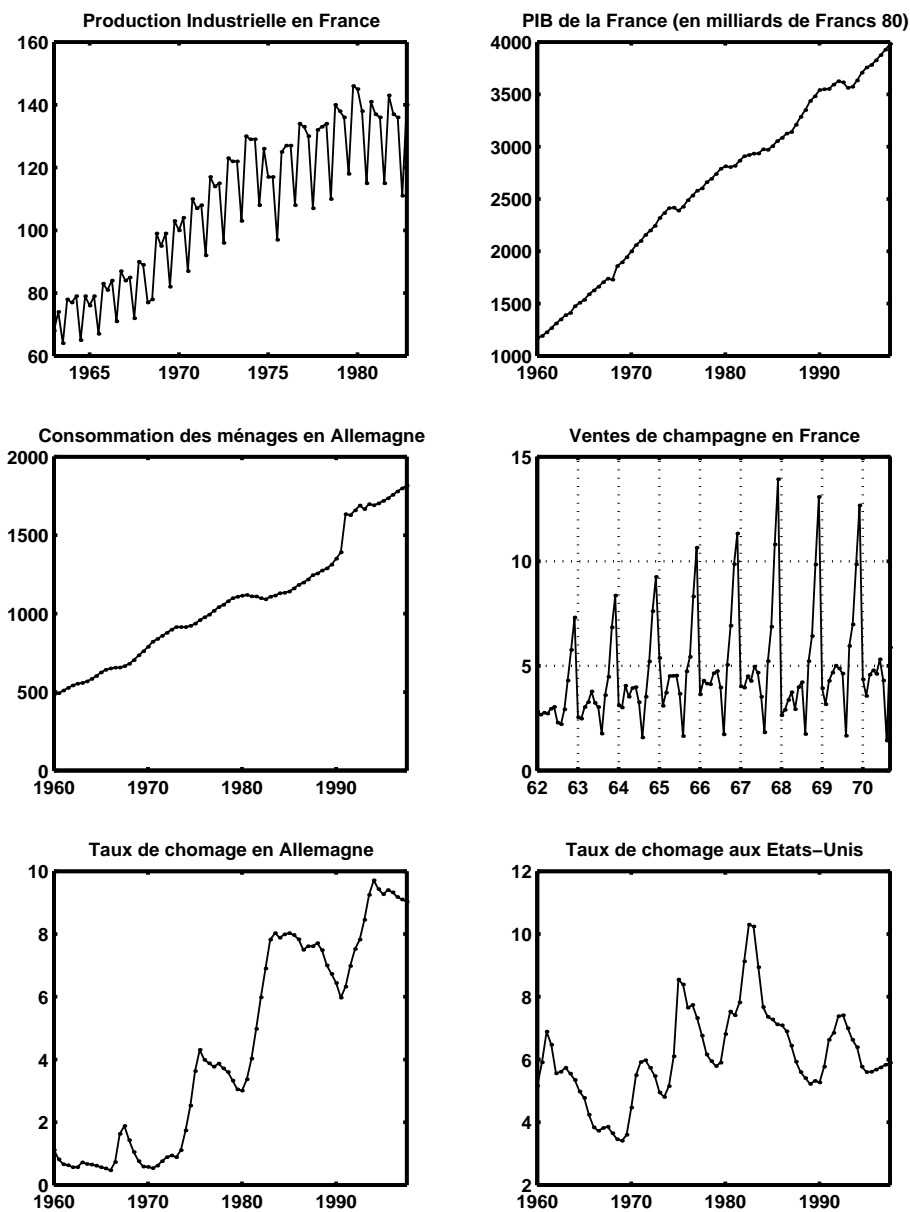


FIGURE 1.2 – *Quelques exemples de séries chronologiques*

Pour une collection extensive des données, voir <https://datamarket.com/data/list/?q=provider:tadlne>
<http://pegasus.cc.ucf.edu/~lni/sta4852/>

Chapitre 2

Décomposition des séries temporelles

2.1 Les composantes d'une série temporelle

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, qui peut contenir par exemple des composantes polynomiales ou exponentielles.
2. *Les variations périodiques* ($s_i, 1 \leq i \leq n$) 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, il s'agit des fonctions périodiques, c'est-à-dire qu'il existe un entier p , appelé *période*, tel que

$$s_i = s_{i+p}, \forall i \geq 1 \quad (2.1)$$

Evidemment, cette composante est entièrement déterminée par ses p premières valeurs s_1, s_2, \dots, s_p , i.e.

$$s_i = \sum_{k=1}^p s_k 1_{i \cong k \pmod{p}}, i = 1, 2, \dots$$

et on a aussi (comme la composante périodique satisfait une récurrence à coefficients constants)

$$s_i = \sum_{k=1}^p A_k \xi_k^i, i = 1, 2, \dots$$

où $\xi_k^p = 1$ sont les racines de l'unité.

Ces deux formules impliquent que les composantes périodiques constituent un espace vectoriel de dimension p . La première formule (mais pas la deuxième) identifie une base de décomposition réelle.

Pour ne pas incorporer une composante constante dans la tendance et aussi dans la composante périodique (i.e. pour des raisons d'unicité

dans la décomposition proposée), on impose que la somme des facteurs saisonniers soit nulle :

$$\sum_{j=1}^p s_j = 0. \quad (2.2)$$

3. *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" ?
4. *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).
5. *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.

Les premières deux composantes constituent le signal, et les deux suivantes constituent le bruit. Dans les analyses à suivre, nous nous concentrerons sur les premières 3 composantes : ($f_i, 1 \leq i \leq n$) la tendance, ($s_j, 1 \leq j \leq p$) les coefficients saisonniers, ($e_i, 1 \leq i \leq n$) les fluctuations irrégulières/résidues.

Les premières 2 composantes relèvent de la bien connue méthode de régression, le but de laquelle est d'obtenir des résidus non corrélés, appelés "bruit blanc". Dans le cas où ça ne réussit pas, une méthode qui peut améliorer les prédictions est de faire une deuxième régression en utilisant les résidus du passé pour prédire les derniers résidus; on appelle cela une "auto-régression", et l'exemple le plus simple est le modèle AR(1).

2.2 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 2.2.1 *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 variable aléatoire ϵ_t est Gaussienne.*

Un bruit blanc a la covariance

$$\gamma(s, t) = \mathbb{E}[\epsilon_s \epsilon_t] = 0, \forall s \neq t \quad \text{et donc le coefficient de corrélation} \quad (2.3)$$

$$\rho(s, t) = \frac{\gamma(s, t)}{\sigma_s \sigma_t} = \delta(s - t) \quad (2.4)$$

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 "idéaliste probabiliste moins structuré" : le "bruit blanc de second ordre" défini par les deux dernières formules équivalentes (2.3), (2.4).

Définition 2.2.2 *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 aleatoire.
2. Le bruit blanc stationnaire est une idéalisation du processus des résidus de la régression linéaire, qu'on aimerait "independents". 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 de Box 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, c.-à-d. quand on a du **bruit corrélé**, la régression classique doit être remplacé par une analyse plus fine, par exemple ARIMA.

2.3 Le modèle AR(1)

L'autorégression la plus simple est

$$Y_t = \varphi Y_{t-1} + b + \epsilon_t, t = 1, 2, \dots, T.$$

Les paramètres φ, b peuvent être estimés par une régression linéaire des points

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

Soit a la moyenne de Y_t . On rappelle que $b = a(1 - \varphi)$, $\varphi = \frac{\text{cov}(Y_t, Y_{t-1})}{\text{Var}(Y_t)}$, et qu'en enlevant la moyenne on arrive à un modèle plus simple pour $Y_t - a$

$$(Y_t - a) = \varphi(Y_{t-1} - a),$$

avec moyenne 0 et $b = 0$. Le fait d'avoir enlevé la moyenne ramène à une droite passant par l'origine $y = \varphi x$. Pour simplifier, on supposera souvent qu'on a déjà enlevé la moyenne de Y_t .

Si on suppose que ϵ_t sont un BB, on arrive au modèle exhibant des corrélations le plus simple : le processus AR(1)

$$Y_t = \varphi Y_{t-1} + b + \epsilon_t, \forall t \in \mathbb{Z}$$

Ce modèle suggère une prévision

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

Exercice 2.3.1 Processus AR(1) stationnaire.

1. Montrez que l'équation :

$$Y_t = \varphi Y_{t-1} + \epsilon_t, \forall t \in \mathbb{Z} \tag{2.5}$$

avec valeur initiale Y_0 a la solution unique

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

2. Rederiver ce résultat par le calcul formel de Heaviside.

3. Montrez que l'équation :

$$Y_t = \varphi Y_{t-1} + \epsilon_t, \forall t \in \mathbb{Z}$$

avec valeur initiale $Y_k, k < t$ connue a la solution unique

$$Y_t = \sum_{i=0}^{t-k-1} \varphi^i \epsilon_{t-i} + \varphi^{t-k} Y_k.$$

4. Montrez que si $|\varphi| < 1$, alors l'équation (2.5) a une solution unique **causale** (qui dépend seulement du **bruit présent et passé**) et stationnaire.

5. Montrez que si $|\varphi| > 1$, alors l'équation (2.5) a une solution unique **noncausale**, qui dépend seulement du **bruit futur**) stationnaire.

Indication : 1. Calculer la solution de (2.5) par des substitutions répétées.

2. Montrez que la série obtenue quand $k \rightarrow -\infty$ est convergente dans le sens L_2 . Pour vérifier la stationnarité, examiner la formule de Y_{t+j} . 3. Si : $|\varphi| > 1$, iterez vers le futur.

Remarque 2.3.1 Dans le cas $|\varphi| = 1$, les solutions de l'équation : (2.5) (appelée marches aléatoires) ne peuvent pas être stationnaires. Cette situation plus compliquée sera considérée dans le chapitre sur les processus $ARIMA(p,d,q)$.

En conclusion :

Théorème 2.3.1 Le processus (de Markov) $AR(1)$ $Y_t = \varphi Y_{t-1} + \epsilon_t$ a une solution stationnaire ssi $|\varphi| \neq 1$ et cette solution est causale ssi $|\varphi| < 1$. Aussi

1. pour $|\varphi| < 1$, on vérifie que ϵ_t est un **bruit d'innovation**, c.-à-d.

$$E[\epsilon_t Y_{t-k}] = 0, \forall k \geq 1.$$

2. pour $|\varphi| > 1$, l'équation : (2.5) a une solution stationnaire unique, qui dépend seulement du **bruit futur** (et ϵ_t n'est pas un bruit d'innovation).

2.4 Opérateurs de retard et d'avance

Pour manipuler les récurrences comme (2.5), (2.1) et (2.2), il est convenable d'introduire un **opérateur de retard** B (ou encore de "rétro-décalage") défini sur l'ensemble des suites par

$$BX_t := X_{t-1} \quad \text{donc } B^i X_t = X_{t-i},$$

où on prend $X_t = 0$ si on travaille avec des suites X_0, X_1, X_2, \dots qui commencent au temps 0, et si $t < 0$.

Alternativement, on peut travailler avec $F = B^{-1}$. En utilisant l'opérateur de retard, les équations (2.1), (2.2) deviennent

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

Définition 2.4.1 a) Une série s_t sera appelée **périodique** de période p ssi

$$s_{t+p} = s_t \iff (1 - B^p)s_t = 0, \quad \forall t \quad (2.6)$$

b) Une série s_t sera appelée **saisonnnière** de période p ssi

$$\sum_{i=1}^p s_{t+i} = 0 \iff \left(\sum_{i=0}^{p-1} B^i\right)s_t = 0, \quad \forall t \quad (2.7)$$

Outrement dit, les opérateurs $1 - B^p, \sum_{i=0}^{p-1} B^i$ élimine les séries périodiques/saisonnnières de degré p .

Remarque 2.4.1 Dans la notation d'opérateurs de retard, on voit immédiatement que la deuxième équation (2.7) implique la première (2.6) (en multipliant par $(I - B)$, et donc les séries saisonnières (périodiques de somme nulle) sont périodiques.

Remarque 2.4.2 Si un opérateur de récurrence annule une suite x_t , alors x_t est une combinaison des puissances des racines du polynôme caractéristique.

2.5 Quelques types de décomposition

Après avoir détecté graphiquement quelles sont les composantes présentes, on propose un modèle de décomposition :

2.5.1 Le modèle additif

$$y_i = f_i + s_i + e_i, \quad 1 \leq i \leq n. \quad (2.8)$$

2.5.2 Le modèle multiplicatif

revient à un modèle additif pour le log des données :

$$y_i = e^{f_i} e^{s_i} e^{e_i} \iff \text{Log}(y_i) = f_i + s_i + e_i, \quad 1 \leq i \leq n. \quad (2.9)$$

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

Remarque 2.5.1 On choisit ce modèle si la courbe qui unit les maximums des observations s'éloigne ou s'approche de celle qui unit les minimums des observations, i.e. si les amplitudes des fluctuations dépendent du niveau.

Remarque 2.5.2 Le modèle multiplicatif est souvent utilisé pour des données de type économique.

2.5.3 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 :

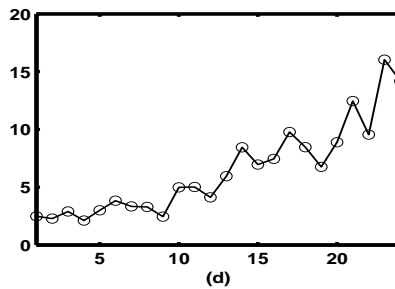
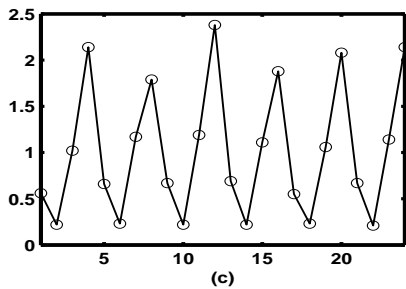
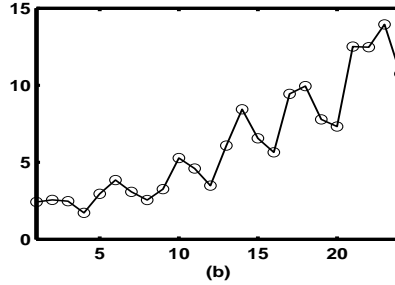
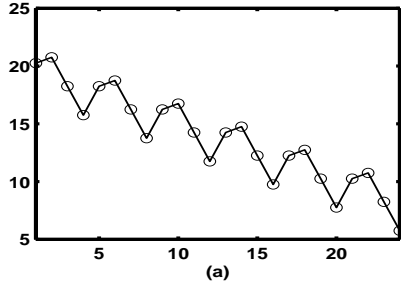
$$y_i = f_i(1 + s_i) + e_i = f_i + f_i s_i + e_i, \quad 1 \leq i \leq n. \quad (2.10)$$

Remarque 2.5.3 Une fois un modèle est obtenu, il peut être utilisé pour la prédiction des valeurs futures.

Toutes les décompositions peuvent être implémentées en effectuant une régression avec `lm` sur des fonctions de base appropriées.

2.6 Exercices

Exercice 2.6.1 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 inconnus, et donnez la formule de la régression qui permet à déterminer ces paramètres.

Exercice 2.6.2 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 moyennes saisonniers ($t_j, 1 \leq j \leq p$), les facteurs saisonniers ($s_j = t_j - m, 1 \leq j \leq p$), où m est la moyenne, et la somme $\sum_{j=1}^p s_j$.
- En notant ($e_i, 1 \leq i \leq n$) la série des fluctuations irrégulières, calculer e_1, e_2, \dots, e_n .
- Proposer une méthode pour l'estimation des paramètres, en supposant une tendance linéaire $m_t = at + b$, en plus de la composante saisonnière, en utilisant `lm`, `cycle`, `factor`.
- Est-ce que a est significatif ? Si non, refaire la régression et comparer avec t_i . Quelle est la matrice de régression utilisée par `lm` ? Comparer les résidues de `lm` avec celles obtenues en d).

Réponse:

```

y=c(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)
t=1:15
plot(t,y,type="b")
#Calculs des facteurs saisonnières:
(t1=0.2*(y[1]+y[4]+y[7]+y[10]+y[13]))
(t2=0.2*(y[2]+y[5]+y[8]+y[11]+y[14]))
(t3=0.2*(y[3]+y[6]+y[9]+y[12]+y[15]))
s1=t1-mean(y)
s2=t2-mean(y)
s3=t3-mean(y)
s1+s2+s3
#Calcul bruit
e=y-mean(y)-rep(c(s1,s2,s3),5)
e
mean(e)
#Régression
z=ts(y, frequency = 3)
cz<-cycle(z)
fz<-factor(cz)
zlm<-lm(z~ time(z)+ fz)
summary(zlm)# Comme l'intervalle de confiance pour le coefficient de "time" inclu le $0$
n'est pas plus grande que $3$ en valeur absolue, ce coefficient n'est pas significatif
(z1lm<-lm(z~ fz)) #(Intercept)7.54 (t1) fz2-3.46 (t2-t1) fz3 -4.70(t3-t1)

```

Exercice 2.6.3 Soit $p = 4$.

1. Trouvez une base complexe **orthogonale** de l'espace vectoriel des séries périodiques d'ordre p , et de longueur $n = kp, k \in \mathbb{N}$. Quelle est la dimension de cet espace ?
2. Trouvez une base réelle de l'espace vectoriel des séries périodiques réelles d'ordre p , et de longueur $n = kp, k \in \mathbb{N}$. Donnez une formule explicite décomposant une série arbitraire appartenant à cet espace vectoriel par rapport à cette base.
3. Même questions pour une série arbitraire appartenant à l'espace vectoriel des séries saisonnières.
4. Montrez que chaque série périodique est la somme d'une série saisonnière et d'une série constante.
5. Formulez en forme matricielle le problème de régression des moindres carrés qui permet de décomposer une série comme une somme d'une tendance linéaire, une partie saisonnière de période 4, et des résidus.
6. Pour $p = 3$, trouvez une base réelle **orthogonale** de l'espace vectoriel des séries saisonnières d'ordre p , et ensuite une base de l'espace des séries périodiques, qui la contient.

Remarque 2.6.1 Après le choix d'une forme appropriée pour la tendance et une période spécifique (selon le graphe), on peut déterminer en même temps

les coefficients de la tendance et de la partie périodique par une régression linéaire.

Par 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. Comme les données sont mensuelles, la période est de un an, et donc $p = 12$. 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 $\sum_i e_i^2$, c.-à-d. par régression.

Exercice 2.6.4 Simuler une série chronologique $Y_t, t = 1, \dots, 100$ suivant le modèle $Y_t = .01t + 1 + 2\sin(2\pi t/5) + \epsilon_t$, où ϵ_t est un bruit blanc gaussien de variance $1/100$. 1. Déterminer la tendance par une régression linéaire de moindres carrés. 2. Éliminer la tendance, puis estimer la saisonnalité de la série chronologique. 3. Comparer les estimateurs avec les vrais coefficients. 4. Analyser les résidus. Les représenter. 5. Comparer avec les résidues obtenues avec la fonction `decompose`.

Exercice 2.6.5 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

- Représenter graphiquement cette série.
- On se propose d'ajuster une tendance f de la forme $f(t) = \frac{1}{a + bt}$. Justifier ce choix.
- 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.
- représenter les 2 tendances ainsi obtenues sur le graphique précédent et comparer les résultats. Est-ce que les résidus ont une allure irrégulière ?

Résumé : Après avoir enlevé des données exceptionnelles et rompu la série en morceaux correspondant aux changements nets, la modélisation des séries chronologiques comporte les pas suivants :

- Tracer la série et examiner les caractéristiques du graphique (par la commande `decompose` par exemple).
Est-ce que les résidues sont bruit blanc, ou plutôt ARMA(p, q) (tester en examinant les fonctions d'autocorrélations et d'autorrelations partielles initiales) ?

2. *Si ce n'est pas le cas, tester si les résidues des séries obtenues par différenciation (commande `diff` ou `filter`), ou en appliquant une transformation, sont "plus *i.i.d.*".*
3. *Typiquement, on aura besoin des modèles ARIMA plus complexes pour les résidues. Une première possibilité est d'utiliser `auto.arima` et en suite de considérer des petites perturbations du modèle proposé. Retester si ces modèles bien produisent des résidues *i.i.d.* (ou au moins réduisent la variance du bruit, pour une série initiale stationnaire).*

Chapitre 3

Séries temporelles en R

```
getwd()#Directoire courant
?ts
data() #tous les jeux de données disponibles en R
library() #nous allons utiliser souvent caschrono
data(package = .packages(all.available = TRUE))#jeux de données groupés par package
data(package = "MASS")
```

3.1 decompose, residuals, acf, Box.test, stl

La commande "decompose" décompose une ts avec frequency en trois composantes : tendance, saisonnier, résidues.

```
is.ts(co2)
frequency(co2)
require(graphics)
m <- decompose(co2)
plot(m) #m$figure
```

Testons les résidues :

```
re=m$random #? decompose
acf(re,na.action=na.pass)
```

Le plot des corrélations montre déjà que decompose n'a pas produit du bruit blanc ! decompose est un bon outil pédagogique, mais peut-être pas idéal pour produire des vrais résidues bruit blanc noncorrélés.

```
Box.test(re)
```

La probabilité que ce soit du bruit blanc est $\approx 2.2e-16$.

Une décomposition alternative est stl, mais elle échoue aussi, avec des résultats semblables.

```
no=stl(co2, "per")
plot(no)
nx=no$time.s\'eries[,"remainder"]
Box.test(nx)
acf(nx,na.action=na.pass)
```

Exercice 3.1.1 Repeter cette analyse pour les deux séries suivantes.

```
sun<-sunspots
is.ts(sun)
plot(sun)

?BJsales
bj<-BJsales
is.ts(bj)
plot(bj)
```

Ces deux séries semblent contenir des composantes périodiques, et la deuxième aussi une tendance.

Exercice 3.1.2 Analyser le jeux des données suivant.

```
## exemple taken from Kendall/Stuart
x <- c(-50, 175, 149, 214, 247, 237, 225, 329, 729, 809, 530, 489, 540, 457,
  195, 176, 337, 239, 128, 102, 232, 429, 3, 98, 43, -141, -77,
  -13, 125, 361, -45, 184)
x <- ts(x, frequency = 4, start = c(1951, 1), end = c(1958, 4))
m <- decompose(x)
plot(m)
m$figure/10 ## seasonal figure: 6.25, 8.62, -8.84, -6.03
```

3.2 Le package caschrono

Nous allons utiliser souvent les ressources à <http://www.seriestemporelles.com/> ou <http://elisabetharagon.pagesperso-orange.fr/SiteWebYves/codes.html>

```
# Exemple 1.1 (Populations)
require(caschrono)
data(popfr)
plot.ts(popfr,xlab="année", ylab="population",main ="Population française, 1846-1951")

plot.ts(uspop,xlab="année", ylab="population",
  main ="Population des Etats-Unis, 1790-1970")
```

```

#par(mfrow=c(2,2) # On decoupe la fenetre graphique en 4
op=par(mfrow=c(2,1),mar=c(2.5,2.5,2,2),mgp=c(1.5,.5,0),
  oma=c(0,0,0,0),cex.main=.8,cex.lab=.7,cex.axis=.7)
plot.ts(popfr,xlab="année", ylab="population",
  main ="Population française, 1846-1951")
plot.ts(uspop,xlab="année", ylab="population",
  main ="Population des Etats-Unis, 1790-1970")
par(op)

# Exemple 1.3 (Champagne)
aa=scan(system.file("/import/champagne_2001.txt",package="caschrono"))
is.ts(aa)
champa = aa/1000; # en milliers de bouteilles
ytr=cbind(champa,log(champa))
colnames(ytr)=c("champagne","log(champagne)")
ytr.ts = ts(ytr, frequency=12, start=c(2001,1))
plot.ts(ytr.ts,xlab="temps")

?cbind

require(graphics)
ts(1:10, frequency = 4, start = c(1959, 2)) # 2nd Quarter of 1959
print( ts(1:20, frequency = 12, start = c(12, 2)), calendar = TRUE)

# Exemple 1.2 (Morts par accident)
data(m30)
op = par(mfrow=c(2, 1),mar=c(2,3,2,2),oma=c(1,1,0,0),
  mgp=c(2,.4,0),cex.main=.75,cex.axis=.8,cex=.8,cex.lab=.8)
plot.ts(m30, xlab="année", ylab="nombre de morts",las=1,
  main ="Accidents de la route en France, mortalité mensuelle")
polygon(c(1995,2005,2005,1995),c(340,340,910,910),lty=2)
debut = c(1995,1); fin = c(2004,12) # zoom
plot.ts(window(m30,start=debut,end=fin),
  xlab="année",ylab="nombre de morts", las=1,
  main ="Années 1995 à 2004")
par(op)

```

3.3 Lac Huron : lm, arima, armax

```
# Exemple 1.5 (Lac Huron):Regression,decomposition, correlations des residues
temps = time(LakeHuron)
reglin = lm(LakeHuron~temps)
summary(reglin)
resi = residuals(reglin)
v12=c(var(LakeHuron),var(resi))
v12
plot(reglin)
```

Le QQ plot montre un bon fit a la loi de Gauss. La parabole observée en Residuals vs fitted suggère une regression quadratique ?

```
plot(LakeHuron)
abline(reglin)
```

Même si une regression quadratique pourrait ameliorer, cela est deja pourvu par auto.arima, donc attendons. Regardons quelques graphiques du resultat.

```
op=par(mar=c(4,3,2,2),mgp=c(2,.4,0),oma=c(0,0,0,0))
plot.ts(LakeHuron, las=1, ylab="niveau", xlab="année")
abline(coef= coef(reglin))
s = c(12,48,59,80);
ychap = fitted(reglin)
segments(temps[s],ychap[s],temps[s],LakeHuron[s],lty=1,lwd=2)
y.gra=1.0009*par()$usr[3]
text(temps[1],y.gra,labels="obs.",cex=.8)
text(temps[s],rep(y.gra,length(s)),labels=s,cex=.8)
par(op)
```

Remarque 3.3.1 *lines()* ajoute des elements a un plot antérieur; *abline(a,b)* ajoute une ligne au plot antérieur (et fait rien dans l'absence d'un plot antérieur).

```
op=par(mfrow=c(2,1),mar=c(3.5,3.5,2,1),mgp=c(2,.5,0),
  oma=c(0,0,0,0),cex.main=.8,cex.axis=.8,cex=.8,cex.lab=.8)
plot(as.vector(temps),resi,type="l",xlab="année",ylab="résidu")
abline(h=0)
zero = rep(0, length(s))
segments(temps[s], zero, temps[s], resi[s], lty=3,lwd=1 )
y.gra=0.9*par()$usr[3]
text(temps[1],y.gra,labels="obs.",cex=.8)
text(temps[s],rep(y.gra,length(s)),labels=s,cex=.8)
n =length(resi)
plot(resi[-n],resi[-1],xlab="résidu en t-1",asp=1,
  ylab="résidu en t")
par(op)
```

```
acf(resi)
pacf(resi)
```

Ces diagnostics suggèrent respectivement des modèles MA(3) et AR(2), et nous en rajouterons une regression linéaire (ARMAX). Mais, comment choisir entre MA(3) et AR(2) ?

```
ma3=arima(LakeHuron,order=c(0,0,3),xreg=temps,method="ML")
ma3
#sigma^2 estimated as 0.461: log likelihood = -101.68, aic = 215.37
```

Tous les coefficients semblent significatifs, sauf peut-être ma3.

```
rem3 = residuals(ma3)
acf(rem3)
pacf(rem3)
```

Résidues non-corrélés cette fois !

```
library(caschrono)
ret= c(3, 6, 9, 12,15)
(a1=Box.test.2(rem3, nlag=ret,type="Ljung-Box",decim=2,fitdf=5))
```

et ils passent le Box.test !

```
ar2=arima(LakeHuron,order=c(2,0,0),xreg=temps,method="ML")
ar2
rea2 = residuals(ar2)
(a2=Box.test.2(rea2, nlag=ret,type="Ljung-Box",decim=2,fitdf=4))
#sigma^2 estimated as 0.4566: log likelihood = -101.2, aic = 212.4
```

On préfère ce model, car le AIC est plus petit, et le Box.test mieux.

```
ychap = fitted(ar2)
op= par(mfrow=c(2,1),mar=c(3.5,3.5,2,1),mgp=c(2,.5,0),
oma=c(0,0,0,0),cex.main=.8,cex.axis=.8,cex=.8,cex.lab=.8)
plot.ts(cbind(LakeHuron,ychap),lty=1:2,type="l",las=1,
plot.type="single",ylab="niveau",xlab="année")
leg.txt = c("Série observée","Prédiction (MCG + erreur AR(2))")
legend(x=1876, y= 577, leg.txt, lty=1:2 , cex=.8)
n=length(rea2)
plot(rea2[-n],rea2[-1],xlab="résidu en t-1",asp=1,
ylab="résidu en t") #Lag plot
title("Lag plot des résidus ARMAX")
par(op)
```

Comparer avec

```
lag.plot(rea2, set.lags = 1:1)
```

On peut automatiser la recherche d'un bon modèle ARIMA

```
require(forecast)
aa=auto.arima(LakeHuron,xreg=temps)
summary(aa)
reaa = residuals(aa)
(a3=Box.test.2(reaa, nlag=ret,type="Ljung-Box",decim=2,fitdf=4))
#ARIMA(1,0,1) + xreg;
#sigma^2 estimated as 0.476: log likelihood=-101.2 AIC=212.4 AICc=213.05 BIC=225.32
```

même log likelihood et AIC, très bon Box.test, mais cette procédure a aussi compare le BIC, donc on la préfère.

3.4 Simulation et verification du bruit blanc : rnorm, rlnorm, rt, etc

```
eps=rnorm(100) #Simuler un bruit blanc gaussien de variance 1 et de longueur 100
Box.test (eps, lag = 1)
x <- rnorm (10)
Box.test (x, lag = 1)
x <- rt (10)
Box.test (x, lag = 1)
```

Remarque 3.4.1 *The Ljung-Box test statistic (X -squared) gets larger as the sample auto-correlations of the residuals get larger (see its definition); its p -value is the probability of getting a value as large as or larger than that observed under the null hypothesis that the residuals are independent. Therefore a small p -value rejects independence.*

Hyndman- Athanasopoulos recommend using $h=10$ for non-seasonal data and $h=2\times$ per for seasonal data.

Verifions l'adéquation des séries simulées avec leurs lois.

```
hist(eps) #Tracer l'histogramme
qqnorm(eps) #Tracer eps en fonction des quantiles d'une loi N (0, 1)
qqline(eps) #Tracer la droite de régression sur les données
abline(0,1) #Tracer la première bissectrice et commenter.
help(qqnorm) #Pour avoir un descriptif de la fonction
? qqline
```

```
epl=rlnorm(100) #Simuler un bruit blanc lognormale (meanlog = 0, sdlog = 1)
hist(epl) #Tracer l'histogramme
qqnorm(epl) #Tracer epl en fonction des quantiles d'une loi N (0, 1)
qqline(epl) #Tracer la droite de régression sur les données
abline(0,1) #Tracer la première bissectrice et commenter.
```



```
require(graphics)
y <- rt(200, df = 5)
qqnorm(y); qqline(y, col = 2)
qqplot(y, rt(300, df = 5))
```

Exercice 3.4.1 a) Comment générer un vecteur gaussien de moyenne m et de variance σ^2 à partir de eps ? Et avec la commande `rnorm`? b) Recommencer cette simulation mais avec un bruit blanc de longueur 1000 puis 10000. Vérifier que la droite de régression sur les données est de plus en plus proche de la première bissectrice. Pourquoi?

3.5 Simulation et analyse arima de quelques modèles de Aragon, Ch.4

Quelques commandes utiles : `arima.sim`, `Box.test.2`, `nlag`, `fitdf`, `col(umn)names`, `t(ranspose)`, `arima`, `order`, `fixed`, `t_stat`

```
# ARMA simulation
as=arima.sim(n = 63, list(ar = c(0.8897, -0.4858), ma = c(-0.2279, 0.2488)),
sd = sqrt(0.18))
ts.plot(as)
```

Ca a une allure periodique! On a envie d'estimer la fréquence, on fera ca plus tard avec le periodogramme

```
# ARIMA simulation
as1 <- arima.sim(list(order = c(1,1,0), ar = 0.7), n = 200)
ts.plot(as1)

#avec residues de loi t
as2=arima.sim(n = 63, list(ar = c(0.8897, -0.4858), ma = c(-0.2279, 0.2488)),
rand.gen = function(n,...) sqrt(0.1796) * rt(n, df = 5))
ts.plot(as2)
```

```
#avec "residues" 1:n
as3=arima.sim(n = 63, list(ar = c(0.8897, -0.4858), ma = c(-0.2279, 0.2488)),
rand.gen = function(n,...) 1:n)
ts.plot(as3)
```

Testons si quelques séries simulées (AR(1) et SAR(1)₁₂) satisfont le test de Ljung-Box de manque de corrélation aux retards (3,6,9,12).

```
require(caschrono)
set.seed(123)
y1 = arima.sim(n=100,list(ar=-.7), sd=sqrt(4))
pacf(y1)
ret= c(3, 6, 9, 12,15)
(a1=Box.test.2(y1, nlag=ret,type="Ljung-Box",decim=2,fitdf=1))
```

Pour ce AR(1), la probabilité qu'il s'agit d'un BB est 0

```
y2 = arima.sim(n=100,list(ar=c(rep(0,11),-.7)), sd=sqrt(4))
pacf(y2)
acf(y2)
(a2=Box.test.2(y2, nlag=ret, type="Ljung-Box",decim=2,fitdf=12))
(a12 = cbind(a1,a2[,2]))
```

Pour ce SAR(1)₁₂, le test de Ljung-Box rejette l'hypothèse de BB seulement au retard 12.

```
colnames(a12)

colnames(a12)= c("Retard","p-val. y1","p-val. y2")
a12
  Retard p-val. y1 p-val. y2
[1,]      3         0      0.49
[2,]      6         0      0.54
[3,]      9         0      0.50
[4,]     12         0      0.00
[5,]     15         0      0.00
```

Remarque 3.5.1 Les resultats pour cette simulation varie beaucoup ; parfois la non BB est depiste par le premier test.

```
y21 = arima.sim(n=100,list(ar=c(rep(0,11),-.7)), sd=sqrt(4))
(a21=Box.test.2(y21, nlag=ret, type="Ljung-Box",decim=2,fitdf=12))
(a22 = cbind(a2,a21[,2]))
colnames(a22)= c("Retard","p-val. y2","p-val. y21")
a22
  Retard p-val. y2 p-val. y21
[1,]      3      0.49      0.42
[2,]      6      0.54      0.59
[3,]      9      0.50      0.66
[4,]     12      0.00      0.00
[5,]     15      0.00      0.00
```

Verifions que arima peut calculer les paramètre $\theta_1 = -.7, \sigma^2 = 4$ du AR(1) en sachant le modèle.

```
(my1 = arima(y1, order=c(1,0,0),include.mean = FALSE))
```

A peu près. Au moins, Ljung-Box confirme maintenant que les residues sont BB.

```
ret=c(3,6,9,12)
aa=Box.test.2(residuals(my1),nlag=ret,type="Ljung-Box",decim=4,fitdf=1)
aa
t(aa)#transpose
```

```
t_stat(my1)#coef. AR a probabilité 0 d'être due aux fluctuations
```

Modèle SAR

```
(my2a = arima(y2, order=c(12,0,0),include.mean=FALSE))
```

```
(my2a = arima(y2, order=c(12,0,0),include.mean=FALSE,  
fixed=c(rep(0,11),NA)))
```

```
(my2a=arima(y2,include.mean=FALSE,  
seasonal=list(order=c(1,0,0),period=12)))  
#avec spécification des param\`etres connus
```

3.6 Importation des données, auto.arima, forecast, spectrum

La modélisation et prédiction d'une série sans périodicité est assez simple

```
w=WWWusage #nombre d'utilisateurs Internet chaque minute  
plot(w)  
library(forecast)  
fit <- auto.arima(w)  
plot(forecast(fit,h=20))
```

Pour les séries avec une fréquence connue, la modélisation et prédiction est encore simple

Exercice 3.6.1 Analyser la série *co2* :

```
?co2  
plot(co2)  
frequency(co2)  
spectrum(co2)  
#remarquer les sommets 1,2,3, ..6 confirmant la p\`eriodicit\`e  
abline(v=1:6, lty=3)  
(ca=auto.arima(co2,seasonal=TRUE))  
cr=ca$residuals  
require(caschnono)  
lag=c(3,6,12,24)  
Box.test.2(cr,nlag=lag)
```

3.7 AirPassengers : auto.arima du modèle multiplicatif

```
plot(AirPassengers)  
lg<-log10(AirPassengers)  
plot(lg)
```

```

library(forecast)
tm=time(lg)
(mod1=auto.arima(lg)) #SARIMA(0,1,1)(0,1,1)[12]
#sigma^2 = 0.0002586,loglik=353.96,AIC=-701.92 BIC=-693.29
plot(time(lg),forecast(mod1),h=20)# a arranger ...

```

Remarque 3.7.1 Le critère d'information d'Akaike, (en anglais Akaike information criterion ou AIC) est une mesure de la qualité d'un modèle statistique proposée par Hirotugu Akaike en 1973.

Lorsque l'on estime un modèle statistique, il est possible d'augmenter la vraisemblance du modèle en ajoutant un paramètre. Le critère d'information d'Akaike, tout comme le critère d'information bayésien, permet de pénaliser les modèles en fonction du nombre de paramètres afin de satisfaire le critère de parcimonie. On choisit alors le modèle avec le critère d'information d'Akaike le plus faible. Le critère d'information d'Akaike s'écrit comme suit :

$$AIC = 2(k - \ln(L)),$$

où k est le nombre de paramètres à estimer du modèle et L est le maximum de la fonction de vraisemblance du modèle. Dans l'exemple ci-dessus par exemple, $AIC/2 = 2 - 353.96 = -351.96$.

```

library(caschrono)
t_stat(mod1)
ret=c(3,6,9,12,15)
re1=residuals(mod1)
Box.test.2(re1,nlag=ret,fitdf=2) #autour de .5

(fit112 <- arima(log10(AirPassengers), c(0, 1, 1),
                 seasonal = list(order = c(1, 1, 2), period = 12)))
t_stat(fit112)
re2=residuals(fit112)
Box.test.2(re2,nlag=ret,fitdf=4) # AIC augmente un peu,
# mais les p-val sont plus acceptables

(fit111 <- arima(log10(AirPassengers), c(0, 1, 1),
                 seasonal = list(order = c(1, 1, 1), period = 12)))
t_stat(fit111)
re1=residuals(fit111)
Box.test.2(re1,nlag=ret,fitdf=3) # AIC entre les deux, mais sar1 pas fiable
#et les p-val revient autour de .5; on retient le modele precedent ou le premier

## Revenons au premier modele, recommande par R
(fit <- arima(log10(AirPassengers), c(0, 1, 1),
              seasonal = list(order = c(0, 1, 1), period = 12)))
plot(forecast(fit),h=20)
(Box.test.2(residuals(fit),ret,type="Ljung-Box",fitdf=2,decim=4))

```

```

update(fit, method = "CSS")
update(fit, x = window(log10(AirPassengers), start = 1954))
op=par(oma=rep(0,4))
xy.acfb(residuals(fit),numer=FALSE)
par(op)

pred <- predict(fit, n.ahead = 24)
tl <- pred$pred - 1.96 * pred$se
tu <- pred$pred + 1.96 * pred$se
ts.plot(AirPassengers, 10^tl, 10^tu, log = "y", lty = c(1, 2, 2))

```

Avec le modele plus sophistiqué, donnons aussi un plot plus sophistiqué

```

#Bande de prédiction pour fit112
ec80 = fit112$sigma2^.5 * qnorm(0.90)
vajust = fitted(fit112)
matri=as.ts(cbind(lg,vajust-ec80,vajust+ec80),
  frequency=12)

op=par(oma=rep(0,4))
plot(matri, plot.type='single', lty=c(1,2,2), xlab="temps",
  ylab='trafic', main="", cex.main=0.8 )
legend( par("usr")[1], par("usr")[4], c("Val. observée","Bande de prédiction"),
  lwd=1, lty=c(1,2) )
par(op)

indi = (lg - (vajust-ec80)) > 0 & ( vajust+ec80 - lg) > 0
(prop = 100*sum(indi)/length(indi))
#proportion dans la bande .82 approx .80 ?

(Box.test.2(residuals(fit112),ret,type="Ljung-Box",fitdf=4,decim=4))
op=par(oma=rep(0,4))
xy.acfb(residuals(fit112),numer=FALSE)
par(op)

```

3.8 Exercices

Exercice 3.8.1 Sauvegarder et analyser les données `tor=scan("http://www.dms.umontreal.ca/morales/docs/tornadoes2010.txt")` `fire=scan("http://www.dms.umontreal.ca/morales/docs/fires2010.txt")`

du projet de la page `http://www.dms.umontreal.ca/morales/risk.php` a)

Examiner les données (plot). b) Transformer par `ts=as.ts(tor)` c) Analyser : `(ta=auto.arima(ts,seasonal=TRUE)) plot(forecast(ta,h=20))`

Testons si `ts` est un bruit blanc :

```

lag=c(3,6,9,12)
Box.test.2(ts,nlag=lag)

```

Oui, ca semble probable. Pour tornadoes, on pourrait reessayer pour les logs :

```
tts=as.ts(log(tor))
(ta=auto.arima(tts,seasonal=TRUE))
plot(forecast(ta,h=20))
```

Testons si tts est un bruit blanc :

```
require(caschrono)
lag=c(3,6,9,12)
Box.test.2(tts,nlag=lag)
```

Oui, ca semble probable !

Exercice 3.8.2 Repeter pour fires :

```
fts=as.ts(fire)
fa=auto.arima(fts,seasonal=TRUE)
plot(forecast(fa,h=20))
Box.test.2(fts,nlag=lag)
```

Dans ce cas, l'hypothèse du BB semble moins probable.

Exercice 3.8.3 Les analyses prochains devraient aboutir dans des modèles avec résidus satisfaisant le Box.test, et 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 .15 (rappel : $p\text{-val} \approx \mathbb{P}\{|t_{val}| \geq 2\}$, où $t\text{-val}$ sont les valeurs "standardisées", i.e. divisées par l'erreur standard (s.e.).

```
?BJsales
bj<-BJsales
plot(bj)
frequency(bj)
spectrum(bj, xlim=c(0,0.1)) #pas des composante periodique
tm=time(bj)
(bj1=lm(bj~tm))
tm2=tm^2
(bj12=lm(bj~tm+ tm2))
(bj12n=lm(bj~tm+ I(tm^2)))
plot(bj)
lines(fitted(bj12),col="red")
summary(bj12)# l'intercept est le plus significatif, et la pente est deja fournie par
auto.arima
require(forecast)
abj=auto.arima(bj)
summary(abj)#ARIMA(1,1,1) with drift
confint(abj)
```

```

require(lmtest) #test p-val
coefstest(abj)
(1-pnorm(abs(abj$coef)/sqrt(diag(abj$var.coef))))*2 #test p-val
require(caschrono)
abjr=abj$residuals
lag=c(6,12,24)
Box.test.2(abjr,nlag=lag)
  acf(abjr)
  pacf(abjr)
plot(forecast(abj))

```

b) uspop

Exercice 3.8.4 sun<-sunspots

```

is.ts(sun)
plot(sun)
(sa=auto.arima(sun,seasonal=TRUE))
sr=sa$residuals
require(caschrono)
lag=c(3,6,12,24)
Box.test.2(sr,nlag=lag)

```

Hypothèse BB rejetée. Voila une serie plus courte, montrant la periodicité plus clairement

plot(sunspot.year)

```
spectrum(sunspots)
```

Pour l'instant, on ne voit rien : on demande à R de lisser un peu le périodogramme.

spectrum(sunspots, spans=10)

On voit maintenant un sommet proche de zéro. On peut estimer sa valeur à l'aide de la fonction locator, comme tout-à-l'heure, puis grossir le dessin.

```

spectrum(sunspots, spans=10, xlim=c(0,1))
abline(v=1:3/11, lty=3) # une seule harmonique

```

```

x <- as.vector(time(sun))
sun <- as.vector(sun)
r1 <- lm( sun ~ sin(2*pi*x/11) + cos(2*pi*x/11) )
r2 <- lm( sun ~ sin(2*pi*x/11) + cos(2*pi*x/11)
          + sin(2* 2*pi*x/11) + cos(2* 2*pi*x/11)
          + sin(3* 2*pi*x/11) + cos(3* 2*pi*x/11)
          + sin(4* 2*pi*x/11) + cos(4* 2*pi*x/11)
          )
plot(sun~x, type='l')
lines(predict(r1)~x, col='red')
lines(predict(r2)~x, col='green', lty=3, lwd=3)

```

Les termes supplémentaires n'apportent pas grand-chose.

```
summary(r2)
```

Pour autant l'analyse n'est pas terminée : les résidus ne ressemblent pas du tout à du bruit blanc.

```
z <- sun - predict(r2)
op <- par(mfrow=c(2,2))
hist(z)
qqnorm(z)
acf(z)
pacf(z)
par(op)
```

On pourra utiliser un modèle ARMA pour les résidus.

Exercice 3.8.5 Analyser lh :

```
plot(lh) #luteinizing hormone in blood samples at 10 mins intervals from a human female
frequency(lh)
spectrum(lh)
```

3.9 Un peu de programmation

Quelques commandes utiles : *filter*, *tapply*

On peut décomposer avec "lm", en spécifiant les fonctions de base du modèle linéaire désiré, et aussi, plus pédagogiquement, un pas après l'autre, comme ci-dessous :

```
X=co2
plot(X)
#Estimation de la tendance
filt=rep(1/12,11) #Définition des coefficients
filt=c(1/24,filt,1/24) #Définition de la moyenne mobile
Z=filter(X,filter=filt,sides=2) #Calcul de la série filtrée par moyenne arithmétique
d'ordre 13 modifiée aux extrémités
Z=ts(Z,start=1963,frequency=12)
ts.plot(X,Z) #Superposition de la série brute et de la série filtrée
```

Estimation du saisonnier, en éliminant la tendance

```
S=X-Z
cycle(S) #indicateurs des colonnes
?tapply #appliquer fonction a chaque groupe
s=tapply(S,cycle(S),mean,na.rm=T) #Estimation des moyennes de chaque periode
sd=s-mean(s) #Estimation des composantes saisonnieres
mean(sd)
```


Série corrigée des variations saisonnières CVS

```
CVS=matrix(1,18,12)
for (i in 1 :18) {for (j in 1 :12) {CVS[i,j]=t(matrix(X,12,18))[i,j]-s[j]}}
#Calcul de CVS = X -s; t donne la transposé
CVS=as.vector(t(CVS))
CVS=as.ts(CVS)
CVS=ts(CVS,start=1963,frequency=12)
ts.plot(X,CVS,col = c("blue", "red"), lwd = c(1, 2))
```

Regression polynômiale de la tendance

```
y=time(CVS)
z=time(CVS)^2
CVS.lm=lm(CVS ~ y+z) #Estimation de la tendance : f(t) = a + bt + ct2
v<-CVS.lm$coefficients#Valeurs des coefficients a,b,c
ts.plot(time(CVS),CVS.lm$fitted.values)
```

Analyse des résidus

```
res=CVS-CVS.lm$fitted.values #Définition des résidus
res=res/sqrt(var(res)) #Définition des résidus réduits
acf(res) #Corrélogramme des résidus
```

Prévision

```
X1=rep(1,12)
for (i in 1 :12)
{X1[i]=v[1]+v[2]*(1981+(i-1)/12)+v[3]*(1981+(i-1)/12)^2+s[i]} #Prévoir trafic en 1981
X2=c(as.vector(X),X1)#Juxtaposition de la série
#chronologique (jusqu'en 1980)et de la prévision pour #1981
X2=as.ts(X2) #Transformation en SC
X2=ts(X2,start=1963,frequency=12) #Renouvellement des temps
ts.plot(X2)
```

Exercice 3.9.1 Decomposition avec nls

```
?BOD # biochem oxygen demand versus time in evaluation of water quality
is.ts(BOD)
as.ts(BOD)
plot.ts(BOD)
require(stats)
# simplest form of fitting a first-order model to these data
fm1 <- nls(demand ~ A*(1-exp(-exp(lrc)*Time)), data = BOD,
  start = c(A = 20, lrc = log(.35)))
coef(fm1)
fm1
# using the plinear algorithm
```

```
fm2 <- nls(demand ~ (1-exp(-exp(lrc)*Time)), data = BOD,
  start = c(lrc = log(.35)), algorithm = "plinear", trace = TRUE)
# using a self-starting model
fm3 <- nls(demand ~ SSasympOrig(Time, A, lrc), data = BOD)
summary(fm3)
```

Modèle ARMA saisonnier (modèle SARMA). 4.5.2, Exemple 4.7

```
require(polynom)
set.seed(7392)
(autopol = polynomial(c(1,0.8))*polynomial(c(1,0,0,0,-0.7)))
yd=arima.sim(n=200,list(ar=-autopol[-1],ma=c(0,0.6)),sd=sqrt(1.5))
yd=yd+4
acf.th=ARMAacf(ar=-autopol[-1],ma=c(0,0.6),lag.max=20,pacf=FALSE)
pacf.th=ARMAacf(ar=-autopol[-1],ma=c(0,0.6),lag.max=20,pacf=TRUE)
plotacfthemp(yd,ar=-autopol[-1], ma=c(0,0.6), lag.max=20)

op=par(oma=rep(0,4))
lag.plot(rev(yd),9,layout=c(3,3),ask=NA,do.lines=FALSE,main="SARMA yd",diag.col="red")
par(op)

(msarma=Arima(yd,order=c(1,0,2),
seasonal=list(order=c(1,0,0),period=4)))

(msarma=Arima(yd,order=c(1,0,2),
  seasonal=list(order=c(1,0,0),period=4),
  fixed=c(NA,0,NA,NA,NA)))
```

Exercice 3.9.2 *tui* est un jeu Excel de Zucchini et Nenadic, contenant stock data for the TUI AG from Jan., 3rd 2000 to May, 14th 2002, namely date (1st column), opening values (2nd column), highest and lowest values (3rd and 4th column), closing values (5th column) and trading volumes (6th column). A convenient way of preparing the data is to clean up the table within Excel, so that only the data itself and one row containing the column-names remain.

Cela est fait sauvegardant dans le repertoire courant comme *tui.csv*

```
getwd() #verifier la presence de tui.csv
tui <- read.csv("tui.csv", header=T, dec=".", sep=",");
```

header=T indique that the first row contains column-names (in case one does not have a naming row, one would use *header=F* instead). The option *dec* sets the decimal separator used in case it differs from a point

```
plot(tui[,5], type="l", lwd=2, col="red", xlab="time",
  ylab="closing values", main="TUI AG", ylim=c(0,60)); #closing values
```

In order to access particular elements of objects, square brackets ([]) are used. Columns of matrices can be accessed with object[,column]. type="l" sets the plot-type to "lines". lwd=2 (line width) is used to control the thickness

Comparison de la hist des differences avec la normale :

```
hist(diff(tui[,4]),prob=T,ylim=c(0,0.6),xlim=c(-5,5),col="red")
lines(density(diff(tui[,4])),lwd=2);
mu<-mean(diff(tui[,4]));
sigma<-sd(diff(tui[,4]));
x<-seq(-4,4,length=100);
y<-dnorm(x,mu,sigma);
lines(x,y,lwd=2,col="blue");
```

the hist() – command is used for creating a histogram of the differences. The option prob=T causes the histogram to be displayed based on relative frequencies. The other options (ylim, xlim and col) are solely for enhancing the display. A nonparametric estimate of the density is added using the function density() together with lines().

Another option for comparison with the normal distribution :

```
qqnorm(diff(tui[,4]));
abline(0,1);
```

Une analyse alternative

```
plot(diff(log(tui[,5])),type="l");
```

Shapiro normality test :

```
x<-diff(log(tui[,5]));
shapiro.test(x);
```

L'effet des filtrages :

```
plot(tui[,5],type="l");
tui.1 <- filter(tui[,5],filter=rep(1/5,5));
tui.2 <- filter(tui[,5],filter=rep(1/25,25));
tui.3 <- filter(tui[,5],filter=rep(1/81,81));
lines(tui.1,col="red");
lines(tui.2,col="purple");
lines(tui.3,col="blue");
```

Decomposition avec stl (ou decompose)

```
beer<-read.csv("beer.csv",header=T,dec=".",sep=",");
beer<-ts(beer[,1],start=1956,freq=12);
plot(stl(log(beer),s.window="periodic"));
```

Regression quadratique avec lm ou lsfit :

```
lbeer<-log(beer);  
plot(lbeer);  
t<-seq(1956,1995.2,length=length(beer));  
t2<-t^2;  
lm1<- lm(lbeer~t+t2);  
lines(t, lm1$fit,lwd=2, col=2);  
summary(lm1);
```

Ajoutons des composantes Fourier :

```
sin.t<-sin(2*pi*t)  
cos.t<-cos(2*pi*t)  
lm2<- lm(lbeer~t+t2+sin.t+cos.t);  
lines(t, lm2$fit,col=4);  
summary(lm2);
```

Chapitre 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", qui sera la composante la plus importante dans la prévision. Intuitivement, c'est un "mouvement lisse à long terme", et mathématiquement c'est une espérance conditionnelle.
- $\epsilon_t = Y_t - m_t$ 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 serviront 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 modélisations se sont avérées 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 *predicteurs exogènes*, il est assez naturel d'adopter une *modélisation autoregressive* pour la *tendance*. Sous certaines conditions de *regularité*, ça ramenera à des *prévisions autoregressives un pas en avant* :

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

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 "stationarité", i.e une certaine uniformité de structure par rapport au temps. ça nous ramene a considerer les processus stochastiques stationnaires, les distributions jointes de quelles ne varient pas avec le temps.

4.2 Processus stochastiques stationnaires

Définition 4.2.1 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'instantants $t_1 \dots t_r \in T$ 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 4.2.2 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, c.-à-d. 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 $k = t - s$, c.-à-d. il existe une fonction d'une variable $\gamma(k)$, paire, telle que :

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

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.

Comme la plupart de séries 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 4.2.1 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 fausse. Une suite Y de variable aléatoire 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.

4.3 Exemples des processus stationnaires

L'idéalisation probabiliste de la propriété que les résidus sont "complètement irréguliers", 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 $AR(p)$.

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

Définition 4.3.1 On appelle processus $MA(q)$ un processus linéaire $Z_t, t \in \mathbb{Z}$ vérifiant une relation :

$$Z_t = \sum_{i=0}^q \theta_i \epsilon_{t-i}, \forall t \in \mathbb{Z} \quad (4.1)$$

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

La notation des polynômes de retard ramène (4.1) à la forme :

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

Exercice 4.3.1 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 $MA(q)$.

4.3.2 Les moyennes mobiles MA(∞)

Définition 4.3.2 Un processus Y_t sera appelé MA(∞) ou linéaire en ϵ_t s'il peut être représenté dans la forme :

$$Y_t = \sum_{i=-\infty}^{\infty} \psi_i \epsilon_{t-i} \quad \text{avec} \quad \sum \psi_i^2 < \infty \quad (4.2)$$

où ϵ_t est un bruit blanc.

Théorème 4.3.1 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 (c.-à-d. $\text{Var } Y_t < \infty$), b) à variance constante stationnaire $\text{Var } Y_t = \sigma_\epsilon^2 \sum_{i=-\infty}^{\infty} \psi_i^2$ c) à autocovariance donnée par :

$$\gamma(t, t+k) = \sigma_\epsilon^2 \sum_{i=-\infty}^{\infty} \psi_i \psi_{i+k} < \infty \quad (4.3)$$

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 .

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

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \quad (4.4)$$

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

Exercice 4.3.2 a) Calculer la fonction génératrice des covariances

$$g(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k$$

du modèle MA(1) et donner une factorisation de cette fonction.

b) Calculer la fonction génératrice des covariances

$$g(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k$$

du modèle causal MA(∞) et donner une factorisation de cette fonction.

4.3.3 Les modèles autorégressifs AR(p) : des processus faciles à prédire

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

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots) + \epsilon_t \quad (4.5)$$

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 4.3.4 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 $\phi_i, i = 1, \dots, p$ tels qu'une relation de récurrence :

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + \epsilon_t, \forall t \in \mathbb{Z} \quad (4.6)$$

est vérifiée.

La notation des polynômes de retard ramène (4.6) à la forme :

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

Définition 4.3.5 Le polynôme

$$\phi(B) = 1 - \sum_{i=1}^p \phi_i B^i$$

sera appelé **symbole** du modèle (4.6). Le polynôme avec toutes les puissances renversées sera appelé **polynôme caractéristique**.

Rq : Les processus autorégressifs sont définis par une équation, qui à priori, peut ne pas avoir des solutions; comme "solution" de l'équation (4.6) nous aimerions avoir une représentation du processus Y_t par rapport au processus ϵ_t .

4.4 Fonctions de transfert et applications

Définition 4.4.1 La fonction $\psi(z) = \frac{1}{\phi(z)}$ est appelée fonction de transfert du modèle AR(p).

Nous venons de voir que le modèle $AR(1)$ a des solutions stationnaires et causales ssi ϕ est à l'intérieur du cercle unitaire (ou quand la singularité $\lambda = \phi^{-1}$ de la fonction de transfert $\psi(z) = (1 - \phi z)^{-1}$ est à l'extérieur du cercle unitaire).

Cela est lié à l'existence des plusieurs développements limités possibles pour la fonction de transfert $\phi(z)^{-1} = \frac{1}{1-z\phi}$.

Exercice 4.4.1 Donner tous les développements en série des puissances Laurent de la fonction de transfert du modèle $AR(1)$.

Réponse:

$$\frac{1}{1-z\phi} = \begin{cases} \sum_{n=0}^{\infty} \phi^n z^n & \text{à l'intérieur du cercle } |z| < \phi^{-1}, \text{ mais} \\ -\sum_{n=-\infty}^{-1} \phi^n z^n & \text{à l'extérieur du cercle, } |z| > \phi^{-1} \\ \text{pas de développement} & |z| = \phi^{-1} \end{cases}$$

Nous verrons qu'on a une situation pareille pour tous les modèles $AR(p)$: la causalité est équivalente au manque des singularités de $\psi(z)$ dans le cercle unitaire.

4.4.1 Récurrences infinies et la causalité des modèles $AR(p)$

Remarque 4.4.1 La fonction de transfert intervient quand on essaie de résoudre formellement le modèle $AR(p)$, "en multipliant" l'équation

$$\phi(B)Y_t = \epsilon_t$$

par $\phi(B)^{-1}$. On arrive à la représentation

$$Y_t = \psi(B)\epsilon_t.$$

Mais quelle sense donner à la multiplication par $\phi(B)^{-1}$? En fait, cela n'est pas toujours possible.

Considérons le problème plus simple du développement en série des puissances Laurent

$$\psi(z) = \phi(z)^{-1} = \sum_{k=-\infty}^{\infty} \psi_k z^k.$$

Il existent plusieurs, mais un seule qui est causal, et un seul qui satisfait $\sum_k \psi_k^2 < \infty$, et ils coincident ssi in n'y a aucune singularité de $\psi(z)$ dans le cercle unitaire.

Cela nous suggère d'accepter seulement ces modèles $ARMA$. Dans ce cas, la solution

$$Y_t = \psi(B)\epsilon_t = \sum_{k=0}^{\infty} \psi_k B^k \epsilon_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k} \quad (4.7)$$

est convergente et causale.

Théorème 4.4.1 a) Un processus $AR(p)$ est causal, c.-à-d. il peut être représenté sous la forme :

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \quad (4.8)$$

où $\sum \psi_i^2 < \infty$ ssi toutes les racines du symbole $\phi(z) = 0$ sont à l'extérieur du cercle unitaire (ou ssi toutes les racines du polynôme caractéristique $z^p \phi(1/z)$ sont à l'intérieur du cercle unitaire).

Dans ce cas, les coefficients ψ_k de l'expansion de Taylor de la fonction de transfert peuvent être calculés par identification des coefficients de z^k dans $\psi(z)\phi(z) = 1$.

Demo : Començons par le cas le plus simple, avec toutes les racines de $\phi(z) = 0$ dehors le cercle unitaire. Alors, on peut vérifier que les coefficients ψ_k satisfont une récurrence avec polynôme caractéristique $z^p \phi(1/z)$, qui a toutes les racines à l'intérieur du cercle unitaire. Dans ce cas, le développement de Laurent devient une série de Taylor $\psi(z) = \phi(z)^{-1} = \sum_{k=0}^{\infty} \psi_k z^k, \forall |z| \leq 1$ convergente. Finalement, cela assure la convergence en L_2 de (4.7), on a

$$Y_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k},$$

et notre processus $AR(p)$ a une représentation $MA(\infty)$ causale.

Dans le cas $|z_i| \neq 1$, mais avec certaines racines du polynôme caractéristique à l'extérieur du cercle unitaire, on peut encore choisir un développement de Laurent tel que l'expansion $MA(\infty)$ fournisse une solution stationnaire, mais non causale de l'équation $\phi(B)Y_t = \epsilon_t$. Ce cas n'a pas d'application en statistiques.

En conclusion, on acceptera désormais seulement les $AR(p)$ causales, avec toutes les racines du polynôme caractéristique à l'intérieur du cercle unitaire, dans quel cas la fonction de transfert a un développement de Taylor valable dans le cercle unitaire, y compris autour de $z = 1$.

Exercice 4.4.2 Donnez une solution stationnaire de l'équation $(1-2B)(1-B/2)Y_t = \epsilon_t$.

4.4.2 La récurrence de Yule-Walker pour les coefficients des fonctions de transfert causales

En appliquant la représentation des modèles $AR(p)$ comme des modèles $MA(\infty)$, nous découvrirons maintenant une "récurrence fondamentale" satisfaite par les coefficients $\psi(k)$ de la fonction de transfert.

Exercice 4.4.3 Soit $\sum_{i=0}^{\infty} \psi_i z^i$ le développement limité autour de 0 de $\frac{1}{1-\phi_1 z - \phi_2 z^2}$.

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

$$\begin{aligned}\psi_0 &= 1, & \psi_1 &= \phi_1 \\ \psi_i &= \phi_1\psi_{i-1} + \phi_2\psi_{i-2} \iff \phi(B)\psi_i = 0, & i &\geq 2\end{aligned}$$

ou

$$\begin{aligned}\psi_{-1} &= 0, \psi_0 = 1, \\ \psi_i &= \phi_1\psi_{i-1} + \phi_2\psi_{i-2} \iff \phi(B)\psi_i = 0, & i &\geq 1\end{aligned}$$

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

3. Généraliser pour le modèle $AR(p)$.

Sol : 2. $1, \phi_1, \phi_1^2 + \phi_2, \phi_1^3 + 2\phi_2\phi_1, \dots$

3. Les solutions explicites de ψ_i , obtenues par le développement limité autour de 0 de $\frac{1}{1-\phi_1z-\phi_2z^2-\dots-\phi_pz^p}$ se compliquent, mais la récurrence

$$\boxed{\phi(B)\psi_i = 0, \quad i \geq 1, \psi_0 = 1, \psi_{-1} = 0, \dots, \psi_{-p+1} = 0} \quad (4.9)$$

reste vraie.

Remarque 4.4.2 Il suit que si un processus $AR(p)$ $Y_t = \phi_1Y_{t-1} + \phi_2Y_{t-2} + \dots + \epsilon_t$ 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 $\phi(z) = (1 - \lambda_1z)(1 - \lambda_2z)\dots(1 - \lambda_pz)$ a ses racines $\lambda_1^{-1}, \dots, \lambda_p^{-1}$ dehors le cercle unitaire), alors les coefficients ψ_i satisfont la **récurrence Yule-Walker** (4.9).

Exercice 4.4.4 a) Montrez que les coefficients $\psi_k = A_1\lambda_1^k + A_2\lambda_2^k$ du développement limité de la fonction de transfert d'un processus $AR(2)$ causal (avec λ_i dans le cercle unitaire) :

$$\frac{1}{(1 - \lambda_1z)(1 - \lambda_2z)} = \sum_{k=0}^{\infty} \psi_k z^k$$

sont des différences divisées de Newton d'ordre 2 de la fonction λ^{k+1} :

$$\psi_k = (\lambda^{k+1})_{[\lambda_1, \lambda_2]} = \begin{cases} \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2}, & \lambda_1 \neq \lambda_2 \\ (k+1)\lambda^k, & \lambda_1 = \lambda_2 = \lambda. \end{cases}$$

Vérifier que $\sum_i \psi_i^2 \leq \infty$.

b) (*) Démontrer le développement limité (due à Sylvester)

$$\frac{1}{(1 - \lambda_1z)(1 - \lambda_2z)\dots(1 - \lambda_nz)} = \sum_{k=0}^{\infty} z^k \psi(k), \quad |z| < \min_i |\lambda_i|$$

où

$$\psi(k) = (\lambda^{k+n-1})_{[\lambda_1, \lambda_2, \dots, \lambda_n]} := \sum_{i=1}^n \frac{\lambda_i^{k+n-1}}{\prod_{j \neq i} (\lambda_i - \lambda_j)}$$

sont des différences divisées.

Corollaire 4.4.1 Pour un processus $AR(p)$ causal Y_t on a $\epsilon_{t+k} \perp Y_t$ si $k \geq 0$. Si le processus est centré, il suit que $E[\epsilon_{t+k}Y_t] = 0$.

Exercice 4.4.5 (*) Trouvez les solutions stationnaires des processus $AR(2)$ avec au moins une des racines caractéristiques λ_1, λ_2 strictement dehors le cercle unitaire $|\lambda| < 1$.

4.5 Inversibilité des processus $MA(q)$

Définition 4.5.1 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 :

$$\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i} \quad (4.10)$$

où $\sum \pi_i^2 < \infty$

Exemple 4.5.1 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 4.5.1 On considère le processus aléatoire $MA(2)$ suivant :

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

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 4.5.1 Un processus $MA(q)$ avec les racines du symbole $\theta(z)$ à l'extérieur du cercle unitaire est inversible, c.-à-d. le bruit peut être représenté sous la forme : $\epsilon_t = \sum_{i=0}^{\infty} \pi_i Y_{t-i}$ où $\sum |\pi_i| < \infty$ (alternativement, les racines du polynôme caractéristique sont à l'intérieur du cercle unitaire). Dans ce cas :

1. Les coefficients π_i sont les coefficients de la série Taylor de $\pi(z) = \frac{1}{\theta(z)}$.
2. ϵ_t appartient à l'espace linéaire engendré par le passé du signal observé

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

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

4.6 Équations de Yule-Walker pour les covariances et corrélations d'un processus AR(p) causal

Remarque 4.6.1 Les covariances et les corrélations d'un processus sont des fonctions paires, i.e. $\gamma_k = \gamma_{-k}$, $\rho_k = \rho_{-k}$.

Exercice 4.6.1 1. Donner des équations satisfaites par les corrélations ρ_i , $i = 1, 2$ et les covariances γ_k , $k = 0, 1, 2$ d'un processus AR(2) causal, en multipliant la formule autoregressive de Y_{t+k} , $k = 0, 1, 2$ par Y_t , en prenant espérance et en utilisant le Corollaire 4.4.1.

2. Calculer les ϕ_i à partir de ρ_i .

3. Calculer la corrélation ρ_1 et γ_0 à partir de ϕ_i , $i = 1, 2$.

Réponse: 1. Les premières deux équations pour les corrélations ρ_k , $k = 1, 2$, sont

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_{-1} = \phi_1 + \phi_2\rho_1 \\ \rho_2 &= \phi_1\rho_1 + \phi_2.\end{aligned}$$

2. Pour trouver les ϕ_i à partir de ρ_i , on résout le système :

$$\begin{cases} \phi_1 + \phi_2\rho_1 = \rho_1 \\ \phi_1\rho_1 + \phi_2 = \rho_2 \end{cases} \iff \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}$$

3. La première équation donne $\rho_1 = \frac{\phi_1}{1-\phi_2}$, et

$$\gamma_0 = \frac{\sigma^2}{1 - \sum_{i=1}^2 \phi_i\rho_i} = \sigma^2 \left(1 - \begin{pmatrix} \phi_1 & \phi_2 \end{pmatrix} \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \right)^{-1}.$$

Théorème 4.6.1 Les covariances et les corrélations d'un processus AR(p) causal sont liées aux coefficients $\phi = (\phi_1, \dots, \phi_p)$ par les équations de Yule-Walker.

1. Pour les covariances, en multipliant la formule autoregressive de Y_{t+k} par Y_t , en prenant espérance et en utilisant le Corollaire 4.4.1, on obtient :

$$\gamma_k = \sum_{i=1}^p \phi_i \gamma_{k-i} \quad \text{pour } k \geq 1 \quad (4.11)$$

$$\gamma_0 = \sum_{i=1}^p \phi_i \gamma_i + \mathbb{E}Y_t \epsilon_t = \gamma_0 \sum_{i=1}^p \phi_i \rho_i + \sigma^2 \quad \text{pour } k = 0$$

2. Pour les corrélations, on remarque d'abord, en divisant par γ_0 , qu'elles satisfont aussi la récurrence (4.11) :

$$\rho_k = \sum_{i=1}^p \phi_i \rho_{k-i} \quad \text{pour } k \geq 1 \quad (4.12)$$

Par conséquent, il suffit de déterminer les premières p corrélations. En appliquant la récurrence (4.12) pour $k = 1, \dots, p$, et en tenant compte aussi de la symétrie de ρ_k , on arrive au système Yule-Walker :

$$\boxed{\mathbf{R}_p \boldsymbol{\phi} = \boldsymbol{\rho}} \quad (4.13)$$

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

$$\mathbf{R}_p = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-1} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix}$$

Ces équations permettent de calculer les coefficients $\boldsymbol{\phi}$ à partir des premières p corrélations et viceversa (en suite, on peut calculer aussi les autres corrélations en utilisant la récurrence).

3. En prenant γ_0 comme facteur commun dans la deuxième equation en (4.11), on trouve γ_0 en fonction des corrélations ρ_i :

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

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

Remarque 4.6.2 Les équations (4.11) sont un cas particulier des "équations normales" de la régression.

Exemple 4.6.1 Dans le cas $p = 3$, les premières trois équations pour $k = 1, 2, 3$ sont

$$\begin{cases} \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 = \rho_1 \\ \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1 = \rho_2 \\ \phi_1 \rho_2 + \phi_2 \rho_1 + \phi_3 = \rho_3 \end{cases} \iff \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix}$$

Remarque 4.6.3 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 resolvant le système de Yule-Walker
2. en représentant d'abord le processus $AR(p)$ comme un processus $MA(\infty)$, par le developpement limité $\psi(z) = 1/\phi(z)$, et en utilisant en suite la formule (4.3) du corrélogramme des processus $MA(\infty)$

3. par le developpement limité de la **fonction génératrice des covariances**

$$\gamma(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k = \frac{1}{\phi(z)\phi(z^{-1})}.$$

Exercice 4.6.2 Calculer le corrélogramme de **AR(1)** avec $\sigma_\epsilon^2 = 1$ par :

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

Tracez la corrélogramme $\{\rho_k\}$ pour (a) $\phi = 0.5$, (b) $\phi = 0$, and (c) $\phi = -0.5$.

Réponse: 1. Les équations Yule-Walker sont :

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

et $\gamma_0 = \frac{1}{1-\phi^2}$. 2. Décomposant en fractions simples :

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

Théorème 4.6.2 Formule générale des corrélations pour $AR(2)$:
Pour le processus $AR(2)$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

le système de Yule-Walker (4.13) donne :

$$\rho_1 = \frac{\phi_1}{1-\phi_2}, \quad \rho_2 = \phi_2 + \rho_1 \phi_1 = \phi_2 + \frac{\phi_1^2}{1-\phi_2}$$

(et les inegalités $|\rho_i| \leq 1, \rho_1^2 \leq \frac{1+\rho_2}{2}$ implique les restrictions $|\phi_1| \leq 2, -1 \leq \phi_2 < 1 - |\phi_1|$).

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

$$\rho_k = \frac{\rho_1 - \lambda_2}{\lambda_1 - \lambda_2} \lambda_1^k + \frac{\lambda_1 - \rho_1}{\lambda_1 - \lambda_2} \lambda_2^k = \rho_1 \psi_{k-1} + \phi_2 \psi_{k-2}, k \geq 1$$

Finalement, en utilisant $\phi_1 = \lambda_1 + \lambda_2, \phi_2 = -\lambda_1 \lambda_2$, 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 4.6.3 *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) $\rho_k = \frac{17}{129}(0.2)^k + \frac{112}{129}(-0.7)^k$, $k = 0, 1, 2, \dots$ et b) $\rho_k = \frac{1}{2}i^k(0.6)^{k/2}\{1 + (-1)^k\} = (0.6)^{k/2} \cos(k\pi/2)$, $k = 0, 1, 2, \dots$*

Exercice 4.6.4 *Tracez les corrélogrammes pour les processus AR(2) avec : (i) $\phi_1 = 0.2$, $\phi_2 = 0.35$ et (ii) $\phi_1 = -0.8$, $\phi_2 = -0.16$*

4.7 L'estimation des modèles AR(p) par le système de Yule-Walker

Le système de Yule-Walker (4.13) permet aussi l'estimation, à partir des corrélations observées $\rho_i, i = 1, \dots, p$, des coefficients $\phi_i, i = 1, \dots, p$. L'idée la plus naturelle est d'adopter une approche recursive, en commençant avec $p = 1, 2, \dots$. Pour chaque p , on denotera par $\phi_{i,p}, i = 1, \dots, p$ le coefficient de Y_{t-i} dans l'équation $Y_t = \sum_{i=1}^p \phi_{i,p} Y_{t-i} + \epsilon_t$, c.-à-d. la solution du problème de moindres carrés de trouver les coefficients qui minimisent l'objectif $E[\epsilon_t^2] = E(Y_t - \sum_{i=1}^p \phi_{i,p} Y_{t-i})^2$, et qui satisfait le système de Yule-Walker.

On augmente en suite le p recursivement, en s'arrêtant quand le dernier coefficient introduit $\phi_{p,p}$ est "petit". L'inconnue essentielle est en effet p .

Définition 4.7.1 *Les coefficients*

$$\tilde{\rho}_p := \phi_{p,p}$$

sont appelés autocorrélations partielles et aussi coefficients de Verblunsky.

Exercice 4.7.1 *Soit Y_t, Y_{t-1}, \dots, Y_0 une suite des observations à allure stationnaire, qu'on aimerait modéliser comme AR(p), avec p à déterminer.*

1. On considère d'abord une modélisation :

$$Y_t = \phi Y_{t-1} + \epsilon_t, t = 1, 2, \dots, n \quad (4.14)$$

où on propose de prendre pour ϕ la valeur qui minimise la somme des erreurs carrées $SC(\phi) = \sum_{i=1}^n \epsilon_i^2$. Quelle est la valeur de ϕ qui minimise cet objectif ?

2. Supposons maintenant que Y_t est un processus stationnaire AR(1), où ϵ_t est un bruit blanc de variance $\sigma^2 = 1$. Quelle est la valeur de ϕ qui minimise l'objectif $E[\epsilon_t^2]$? Comment estimer cette valeur à partir des données observées ? Comparer avec la réponse précédente.

3. On suppose en suite que Y_t sont des observations d'un processus stationnaire $AR(2)$,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t, t = 2, 3, \dots, n \quad (4.15)$$

où ϵ_t est un bruit blanc de variance $\sigma^2 = 1$. Trouver et résoudre un système d'équations satisfaites par les valeurs ϕ_1, ϕ_2 qui minimisent l'objectif $E[\epsilon_t^2]$.

4. Reprenons maintenant la première question, i.e. essayons d'approximer notre suite par le modèle (4.15) par le principe des moindres carrés, sans des hypothèses statistiques sur le bruit.

Trouver un système d'équations satisfaites par les valeurs ϕ_1, ϕ_2 qui minimisent l'objectif $SC = \sum_{i=2}^n \epsilon_t^2$. Résolver le système et comparer avec la réponse précédente.

5. Calculer les 2 premières autocorrélations partielles $\tilde{\rho}_i = \phi_{i,i}, i = 1, 2$ du processus Y_t , en fonction de coefficients ϕ_1, ϕ_2 .

Réponse: 1) Pour $p = 1$, on trouve $\tilde{\rho}_1 = \phi_{1,1} = \rho_1$.

3) Après avoir déterminé que $\tilde{\rho}_1 = \rho_1$, passons à $p = 2$. Soit φ' les solutions du système YW :

$$\begin{aligned} \rho_1 &= \varphi'_1 + \varphi'_2 \rho_1 \\ \rho_2 &= \varphi'_1 \rho_1 + \varphi'_2 \end{aligned}$$

On trouve

$$\varphi'_1 = \phi_{2,1} = \frac{\begin{vmatrix} \rho_1 & \rho_1 \\ \rho_2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2} \quad \varphi'_2 = \phi_{2,2} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \tilde{r}_2$$

Exercice 4.7.2 On accepte finalement un modèle $AR(3)$ avec bruit blanc $\epsilon_t \sim N(0, \sigma^2 = 1)$ (avec coefficients ϕ_1, ϕ_2, ϕ_3 déterminés par les données. Donner les équations de Yule-Walker liant les autocorrélations ρ_1, ρ_2, ρ_3 d'un processus $AR(3)$ aux coefficients ϕ_1, ϕ_2, ϕ_3 . Calculer les 2 premières autocorrélations $\rho_i, i = 1, 2$ du processus Y_t , ainsi que la variance, en fonction de coefficients ϕ_1, ϕ_2, ϕ_3 .

Calculer la troisième autocorrélation partielle $\tilde{\rho}_3$ du processus Y_t , en fonction de coefficients ϕ_1, ϕ_2, ϕ_3 .

Réponse: Pour $p = 3$,

$$\phi_{3,3} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}.$$

Cette expression permet déjà de démontrer que :

Exercice 4.7.3 Pour un processus $AR(2)$, le coefficient $\phi_{3,3} = 0$, et aussi $\phi_{4,4} = \phi_{5,5} = \dots = 0$. Plus généralement, pour un processus $AR(p)$, on a $\phi_{p+k,p+k} = 0, \forall k \geq 1$.

Remarque 4.7.1 Les équations de Yule-Walker sont juste un cas particulier des équations de la regression multivariée, et ϕ'_i sont exactement les coefficients de l'autoregression de Y_t sur Y_{t-1}, Y_{t-2}, \dots

Un autre exercice intéressant est de montrer la formule recursive de Szego-Levinson-Durbin :

$$\phi_{k,k} = \frac{\rho_k - \sum_{i=1}^{k-1} \rho_{k-i} \phi_{k-1,i}}{1 - \sum_{i=1}^{k-1} \rho_i \phi_{k-1,i}}, \forall k \geq 1 \quad (4.16)$$

Exercice 4.7.4 (*) Démontrer (4.16), en commençant avec le cas $k = 3$.

Ind : Cherchez une équation pour $\phi'_3 = \phi_{3,3}$, en éliminant $\phi'_1 = \phi_{3,1}, \phi'_2 = \phi_{3,2}$ du système YW :

$$\begin{aligned} \rho_1 &= \phi'_1 + \phi'_2 \rho_1 + \phi'_3 \rho_2 \\ \rho_2 &= \phi'_1 \rho_1 + \phi'_2 + \phi'_3 \rho_1 \\ \rho_3 &= \phi'_1 \rho_2 + \phi'_2 \rho_1 + \phi'_3 \end{aligned}$$

Les deux premières équations donnent

$$\begin{pmatrix} \phi'_1 \\ \phi'_2 \end{pmatrix} = R_2^{-1} \begin{pmatrix} \rho_1 - \phi'_3 \rho_2 \\ \rho_2 - \phi'_3 \rho_1 \end{pmatrix} = R_2^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} - \phi'_3 R_2^{-1} \begin{pmatrix} \rho_2 \\ \rho_1 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} - \phi'_3 R_2^{-1} \begin{pmatrix} \rho_2 \\ \rho_1 \end{pmatrix}$$

Finalement,

$$\phi'_3 (1 - (\rho_2, \rho_1) R_2^{-1} \begin{pmatrix} \rho_2 \\ \rho_1 \end{pmatrix}) = \rho_3 - (\rho_2, \rho_1) \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \iff \phi'_3 (1 - (\rho_1, \rho_2) \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}) = \rho_3 - (\rho_2, \rho_1) \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

et

$$\phi'_3 = \frac{\rho_3 - (\rho_2 \phi_1 + \rho_1 \phi_2)}{1 - (\rho_1 \phi_1 + \rho_2 \phi_2)}$$

Les coefficients $\phi_{k,k}, \forall k \geq 1$ étaient bien connues aux mathématiciens sous le nom des coefficients de Verblunsky. En statistique, ils sont appelées coefficients d'autocorrélation partielle, due à la propriété suivante :

Exercice 4.7.5 Montrer que pour un processus stationnaire

$$\phi_{k,k} = E[(Y_k - E[Y_k/Y_{k-1}, Y_{k-2}, \dots, Y_1])(Y_0 - E[Y_0/Y_{k-1}, Y_{k-2}, \dots, Y_1])] / \gamma_0$$

Vérifiez que cette formule donne la même valeur pour $\phi_{2,2}$ comme celle de l'exercice 4.7.1.

Chapitre 5

Les modèles ARMA(p,q)

Définition 5.0.1 On appelle processus ARMA(p,q) un processus stationnaire $Y_t, t \in \mathbb{Z}$ vérifiant une relation de récurrence :

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{i=0}^q \theta_i \epsilon_{t-i}, \forall t \in \mathbb{Z} \quad (5.1)$$

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 (5.1) à la forme :

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

Nous verrons dessous que les processus ARMA(p,q) avec des polynômes caractéristiques $\phi(B), \theta(B)$ 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 (5.1), suivies par un développement de la fraction correspondante dans une série des puissances :

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

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

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

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

Théorème 5.1.1 a) Un processus ARMA(p,q) avec toutes les racines du polynôme caractéristique $\phi(z)$ à l'extérieur du cercle unitaire est causal, c.-à-d. 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 $\psi(z) = \frac{\theta(z)}{\phi(z)}$ b) Un processus ARMA(p,q) avec les racines du polynôme caractéristique $\theta(z)$ à l'extérieur du cercle unitaire est inversible, c.-à-d. 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 $\pi(z) = \frac{\phi(z)}{\theta(z)}$

Corollaire 5.1.1 Pour un processus ARMA(p,q) avec toutes les racines des polynômes caractéristiques $\phi(z), \theta(z)$ 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 $\phi(z)$ et $\theta(z)$ à l'extérieur du cercle unitaire, qui sont causales et inversibles, et on s'appuyera sur le corollaire ci-dessus.

Exercice 5.1.1 ARMA(1,1)

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

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

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(∞) et les coefficients π_j de sa représentation comme processus AR(∞) et précisez si ces représentations sont convergentes.

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

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

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

Rémarque : Dans le cas le plus simple avec les racines λ_i de l'équation $\phi(z) = 0$ distinctes, on obtient facilement des formules générales pour les coefficients ψ_n en commençant par un développement en fractions simples $\pi(z) = \frac{\phi(z)}{\theta(z)} = \sum_i K_i \frac{1}{1-z/\lambda_i}$ où λ_i sont les racines du $\theta(z)$ et donc $K_i = -\frac{\theta(\lambda_i)}{\phi'(\lambda_i)}$. 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'aborder le développement Taylor $\frac{\theta(z)}{\phi(z)} = \psi(z)$ directement, en obtenant des équations de récurrence pour ψ_k , à partir des coefficients du développement $\phi(z)\psi(z) = \theta(z)$. Cette méthode, est en effet applicable toujours :

L'exercice ci-dessus peut être généralisé :

Théorème 5.1.2 (*) a) Pour un processus ARMA(p,q) $\phi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\phi(z)$ 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

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

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

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

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

généralisé Ainsi,
$$\begin{cases} \psi_1 = \theta_1 + \phi_1 \\ \psi_2 = \theta_2 + \phi_1(\theta_1 + \phi_1) + \phi_2 \\ \psi_3 = \theta_3 + \phi_3 + 2\phi_1\phi_2 + \phi_1^3 + (\phi_2 + \phi_1^2)\theta_1 + \phi_1\theta_2 \\ \dots \end{cases} \quad b) \text{ Pour}$$

un processus ARMA(p, q) $\phi(B)Y_T = \theta(B)\epsilon_t$ 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

$$\pi_0 = 1, \quad \pi_k = -\phi_k - \sum_{i=1}^{\min[k, q]} \theta(i)\pi(k-i), \quad 1 \leq k \leq p$$

$$\pi_k = - \sum_{i=1}^{\min[k, q]} \theta(i)\pi(k-i), \quad k > p$$

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

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

Démonstration: Formellement, il suffit de projeter les relations $\phi(B)\psi(B) = \theta(B)$, $\theta(B)\pi(B) = \phi(B)$ 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

$$\hat{X}_t := E[X_t / F_{t-1}] = - \sum_{i=1}^{\infty} \pi_i X_{t-i} =$$

$$(\phi(1) + \theta(1))X_{t-1} + (\phi(2) - \theta(1)^2 - \phi(1)\theta(1) + \theta(2))X_{t-2} +$$

$$(\phi(3) + \theta(1)^3 + \phi(1)\theta(1)^2 - 2\theta(2)\theta(1) - \phi(2)\theta(1) - \theta(2)\phi(1) + \theta(3))X_{t-3} + \dots$$

La formule est plus compliquée à lire, mais, finalement, pour un logiciel symbolique, pas plus compliquée à dériver 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 5.1.3 (*) La fonction génératrice des covariances $\gamma(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k$ d'un processus ARMA est

$$\gamma(z) = \frac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})}$$

Pour $z = e^{itx}$, $x \in \mathbb{R}$ on obtient la transformée Fourier des corrélations aussi appelée densité spectrale

$$f(\theta) = \frac{1}{2\pi} \gamma(e^{itx})$$

La densité spectrale est nonnegative, ce qui est équivalent au fait que les matrices Toeplitz \mathbf{R} des corrélations sont positives définies.

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

On obtient les mêmes équations de Yule-Walker $\gamma_k = \sum_{i=1}^p \phi_i \gamma_{k-i}$ 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 autrement (par la causalité). On obtient :

$$\gamma_k = \sum_{i=1}^p \phi_i \gamma_{k-i} + \sum_{0 \leq j \leq q} \theta_j \gamma_{k-j}^{(Y,\epsilon)} = \sum_{i=1}^p \phi_i \gamma_{k-i} + \sigma^2 \sum_{k \leq j \leq q} \theta_j \psi_{k-j} \quad (5.2)$$

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 (5.2) 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

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

Exercice 5.1.2 ARMA(1,1)

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

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

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

En général, nous trouvons :

Théorème 5.1.4 (*) Les premières $p + 1$ covariances s'obtiennent du système à $p + 1$ équations et $p + 1$ inconnues :

$$\mathbf{\Gamma} \begin{pmatrix} 1 \\ -\phi_1 \\ -\phi_2 \\ \dots \\ -\phi_p \end{pmatrix} = \sigma^2 \mathbf{\Psi} \begin{pmatrix} 1 \\ \theta_1 \\ \theta_2 \\ \dots \\ \theta_q \end{pmatrix} \quad (5.3)$$

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

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

5.2 Les modèles ARIMA(p,d,q)

Définition 5.2.1 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) :

$$Y_t = \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{i=0}^q \theta_i \epsilon_{t-i}, \quad \forall t \in \mathbb{Z} \quad (5.4)$$

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 (5.4) à la forme :

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

où $\phi(B), \theta(B)$ sont des polynômes relativement premiers 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 donné par les mêmes formules que celle des processus stationnaires ARMA(p,q).

Chapitre 6

Filtres/moyennes mobiles

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

Définition 6.0.1 La série Y_t s'appelle une **moyenne mobile** de X_t ou **filtre**[§] si

$$Y_t = \sum_{i=-k_1}^{k_2} \theta_i X_{t-i} \quad (6.1)$$

où $k_1, k_2 \geq 0$. Le **degré** du filtre est $k_1 + k_2$, et son **ordre**/nombre des coefficients est $k_1 + k_2 + 1$.

Le cas des θ_i positifs 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**. Un filtre de lissage est supposé réduire le bruit, mais laisser in-variantes des tendances constantes ou polynomiales.

Remarque 6.0.1 La notation des polynômes de retard ramène (6.1) à la forme :

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

où $\theta(B)$ est un **opérateur de récurrence** (sur l'espace des series). a) Le polynôme de Laurent associé $\theta(z) = \sum_{i=-k_1}^{k_2} \theta_i z^i, : \mathbb{C}^- \rightarrow \mathbb{C}$ sera appelée le symbole de l'opérateur $\theta(B)$. b) En écrivant le filtre comme

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

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

§. 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.

Un filtre de lissage

$$Y_t = \sum_{i=1}^k \theta_i X_{t-i}, \quad \sum_i \theta_i = 1$$

peut être utilisé pour calculer la tendance et pour la prédiction de X_t .

Lemme 6.0.1 Un filtre de prédiction est "non-biaisée pour une série stationnaire", c.-à-d. :

$$\mathbb{E}Y_t = \mathbb{E} \sum_{i=1}^k \theta_i X_{t-i} = \left(\sum_{i=1}^k \theta_i \right) \mathbb{E}X_t,$$

ssi

$$\sum_{i=1}^k \theta_i = 1.$$

Remarque 6.0.2 Cette condition est aussi équivalente au fait que la série égale à $\mathbf{1}$ (et en fait chaque série constante $X_t = c$) reste invariante après l'application du filtre, c.-à-d. $\theta(B)\mathbf{1} = \mathbf{1}$.

6.1 Filtres qui éliminent/laissent invariantes des composantes polynômiales, saisonnières ou périodiques

Exercice 6.1.1 Trouvez l'élément suivant des séries y_t ci-dessous, ainsi que des équations de récurrences qu'elles satisfont et leurs solutions analytiques :

$$2, 6, 12, 20, 30, 42, \dots \quad (6.2)$$

$$4, 10, 20, 36, 62, 104, \dots \quad (6.3)$$

$$3, 2, 1, 6, 3, 2, 1, \dots \quad (6.4)$$

$$0, -1, -2, 3, 0, -1, -2, \dots \quad (6.5)$$

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).

L'opérateur de différentiation D enlève les fonctions constantes continues. Pour les séries temporelles, il faut le remplacer par l'opérateur différence.

Définition 6.1.1 L'opérateur qui fait la différence entre une suite et sa version décalée d'une unité de temps :

$$\Delta X_t := X_t - X_{t-1} = (1 - B)X_t$$

sera appelé opérateur différence.

Exercice 6.1.2 Montrer que l'opérateur différence **élimine** les constantes, et que ses puissances $\Delta^d = (1 - B)^d$ élimine les tendances polynomiales de degré $d - 1$. Réciproquement, un filtre polynôme $\theta(B)$ élimine les tendances polynomiales de degré $d - 1$ ssi son symbole $\theta(z)$ est divisible par $(1 - z)^d$, c.-à-d. admet 1 comme racine multiple de degré d .

Lemme 6.1.1 Un filtre $\theta(B)$ laisse invariante une suite x_t ssi le filtre $\tilde{\theta}(B) := I - \theta(B)$ annule la suite x_t .

Des lors, ces deux problèmes sont équivalentes.

Le théorème suivant nous donne un critère pour identifier les filtres $\theta(B)$ qui laissent invariants tous les polynômes de degré donné.

Exercice 6.1.3 Montrez qu'un filtre polynôme $\theta(z)$ qui est divisible par $1 + z + \dots + z^{p-1}$, c.-à-d. de la forme $\theta(z) = (1 + z + \dots + z^{p-1})\theta_1(z)$, élimine les composantes saisonnières de période p , c.-à-d. :

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

pour chaque série s_t satisfaisant (2.7).

En effet, la réciproque est aussi vraie :

Théorème 6.1.1 a) Un filtre polynôme $\theta(B)$ **élimine** (ou enlève/annule) les composantes saisonnières d'ordre p ssi son symbole $\theta(z)$ est divisible par $1 + z + \dots + z^{p-1}$.

b) Un filtre polynôme $\theta(B)$ annule (ou enlève) les composantes périodiques d'ordre p ssi son symbole $\theta(z)$ est divisible par $1 - z^p$.

Dem (*) : a) Soit $r_1(z)$ le reste de la division de $\theta(z)$ par $1 + z + \dots + z^{p-1}$, qui a au plus $p - 1$ coefficients (en prenant $p = 4$ par exemple, $r_1(z) = \theta_0 + \theta_1 z + \theta_2 z^2$). Alors, $r_1(z)$ annulent l'espace vectoriel de dimension $p - 1$ des séries saisonnières. Comme $r_1(z)$ a au plus $p - 1$ coefficients, il doit être identiquement 0 (prenez une base de l'espace vectoriel, appliquez $r_1(B)$, obtenez un système homogène d'ordre $p - 1$ pour les coefficients θ_i , et remarquez que son déterminant est non nul).

Exercice 6.1.4 1. Montrez que le filtre $P(B) = \frac{1}{3}(2 + B + B^2 - B^3)$ "enlève" les composantes saisonnières de période 3.

2. Trouvez le degré de la tendance polynomiale maximale conservée (laissée invariante) par ce filtre.

Sol : 1. $\theta(z) = \frac{1}{3}(2 - z)(z^2 + z + 1)$. 2. $1 - \theta(z) = \frac{1}{3}(z - 1)^2(z + 1)$, donc degré 1.

Exercice 6.1.5 Trouver un filtre $1 + \alpha B + \beta B^2 + \gamma B^3$ qui laisse passer une tendance affine sans distorsion et élimine les périodicités d'ordre 2.

Indication : Trouver un système des 2 équations avec deux inconnues, et résoudre.

Exercice 6.1.6 Trouvez un filtre symétrique à 5 termes $\theta(B)$ qui laisse invariants les polynômes de troisième degré, et enlève les composantes saisonnières d'ordre 3.

R :

$$\begin{cases} 1 - \theta(z) = bz^{-2}(1 - z)^4, \\ z^2\theta(z) = (1 + z + z^2)\theta_1(z) = z^2 - b(1 - z)^4 \end{cases} \implies$$

$$b = \frac{1}{9}, \quad \theta(B) = \frac{1}{9}(-B^2 + 4B + 3 + 4B^{-1} - B^{-2}).$$

Exercice 6.1.7 Montrez qu'un filtre élimine les composantes saisonnières d'ordre p ssi il transforme chaque fonction de période p dans une fonction constante.

6.2 Moyennes mobiles symétriques

Exercice 6.2.1 Les moyennes mobiles arithmétiques symétriques.

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

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

conserve (laisse invariants) les tendances linéaires $p_t = a + bt$, et élimine les séries saisonnières de période 3. b) Généraliser pour q quelconque. c) Étudier la première propriété pour une moyenne symétrique d'ordre 3 arbitraire.

Sol : a) $\theta(B) = \frac{1}{3}(B^{-1} + 1 + B)$ laisse invariants les tendances linéaires ssi $(1 - \theta(B))B:(1 - B)^2$.

c) $\theta(B) = (\theta_-B^{-1} + (1 - \theta_- - \theta_+) + \theta_+B)$ laisse invariants les tendances linéaires ssi $(1 - \theta(B))B:(1 - B)^2$.

Exercice 6.2.2 Moyennes mobiles symétriques, modifiés au bord. Trouver un filtre d'ordre $2k + 1$

$$\theta(z) = \theta_0(z^{-k} + z^k) + \theta_1\left(\sum_{i=1-k}^{k-1} z^i\right)$$

qui élimine les séries saisonnières de période $2k$ et laisse invariants des tendances polynômiales de degré maximal possible.

Quel est le degré maximal possible des tendances polynômiales laissées invariants par des telle moyennes mobiles modifiées au bord d'ordre 4 et 5 ?

Exercice 6.2.3 Moyennes mobiles arithmétiques polynomiales, modifiés au bord. Trouver un filtre polynomial d'ordre $2k + 1$

$$\theta(z) = \theta_0(1 + z^{2k}) + \theta_1\left(\sum_{i=1}^{k-1} z^i + z^{2k-i}\right)$$

qui élimine les séries saisonnières de période $2k$ et laisse invariants des tendances polynômiales de degré maximal possible.

Quel est le degré maximal possible des tendances polynômiales laissées invariantes par une moyenne mobile arithmétique modifiée au bord d'ordre 5 ?

Quel est le degré maximal possible des tendances polynômiales laissées invariantes par une moyenne mobile arithmétique d'ordre 4 ?

R : $\theta_1 = \frac{3k}{4k^2-1}, \theta_0 = -\frac{1+k}{2(2k-1)}.$

Remarque 6.2.1 Pour enlever les composantes saisonnières d'ordre 4, on peut utiliser donc la moyenne mobile arithmétique d'ordre 4, ou encore mieux, une moyenne mobile modifiée au bord d'ordre 5. Pour une périodicité mensuelle on peut utiliser la moyenne mobile arithmétique d'ordre 12, etc... En R, on utilise "aggregate" En suite, après éliminer la partie saisonnière, une regression obtiendra la tendance et les résidues.

6.3 La réduction du bruit des observations

Un résultat désirable de l'application des filtres de lissage : la réduction de la variance des observations.

Exercice 6.3.1 a) 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 $2q + 1$, et que ce résultat est optimal. b) **La meilleure moyenne mobile arithmétique modifiée au bord ?** Trouver un filtre polynômiale symétrique d'ordre $2k + 1$

$$\theta(z) = \theta_0(1 + z^{2k}) + \theta_1\left(\sum_{i=1}^{k-1} z^i + z^{2k-i}\right)$$

qui laisse invariants les polynômes de premier degré et minimise la variance σ^2 d'un bruit blanc.

Remarque 6.3.1 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 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 $e'_t = \frac{\epsilon_{t+q} + \dots + \epsilon_t + \dots + \epsilon_{t-q}}{2q+1}$ de variance inférieure à celle de e_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 .

6.4 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 en plus la variance (i.e. sont optimales pour enlever le bruit).

Aussi, les filtres fournissent un langage adéquat pour décrire les séries temporelles. Rappelons 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 sûr le signal, et son depistage peut se faire par l'approche classique de régression. Une fois depisté, la prediction du signal est évidente (par sa formule!). Par contre, la prediction du bruit (corrélé) demande des approches plus sophistiquées, comme la modélisation ARMA, où l'utilisation des filtres devienne essentielle.

Il est naturel d'essayer de depister le signal par un filtre de lissage causal $\theta(B)$, donc $m_t = \theta(B)Y_t$, tq ce filtre "détruit 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 $\pi(B)$ "détruit 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 d'appliquer des

tests pour décider quand une série est un bruit. Nous allons examiner plus tard des statistiques comme les corrélations, corrélations partielles, le périodogramme, etc., issues de la modélisation probabiliste des séries temporelles.

Théorème 6.4.1 (*) Soit $\theta(B) = \sum_{i=-k_1}^{k_2} \theta_i B^i = B^{-k_1} \sum_{i=-k_1}^{k_2} \theta_i B^{i+k_1} = B^{-k_1} p(B)$ un filtre (typiquement symétrique), déterminé par k_1 et par le polynôme $p(B)$. Alors, l'espace invariant des séries Z_t satisfaisant $\theta(B)Z_t = Z_t$ contient les polynômes de degré $\leq n$ ssi a) 1 est une racine d'ordre au moins $n + 1$ de l'équation $\theta(z) = 1$, c.-à-d. $\theta(1) = 1, \theta'(1) = 0, \theta''(1) = 0, \theta^{(n)}(1) = 0$, où $\theta(z)$ est le symbole. b) 1 est une racine d'ordre au moins $n + 1$ de l'équation polynomiale $p(z) - z^{k_1} = 0$, c.-à-d. $p(1) = 1, p'(1) = k_1, p''(1) = k_1(k_1 - 1), p^{(n)}(1) = (k_1)_n$, où $p(z)$ est la partie polynomiale du symbole.

Exercice 6.4.1 Demontrez le théorème pour $p = 0, 1, \dots$

6.5 Récurrences et équations différentielles linéaires

L'étude des marches aléatoires et des processus Markoviens ramène souvent à des équations différentielles ou des récurrences linéaires. Le cas des coefficients constants est assez simple, car toutes les solutions peuvent être construites à partir des solutions basiques exponentielles e^{rx} . Comme le cas des équations différentielles à coefficients constants est très bien connu, on rappelle ici seulement le cas de récurrences linéaires.

6.5.1 L'équation de récurrence linéaire à coefficients constants

Les deux équations de récurrence linéaire de deuxième ordre ci-dessous

$$au_{n+2} + bu_{n+1} + cu_n = 0, \quad (6.6)$$

$$av_{n+2} + bv_{n+1} + cv_n = d_n, \quad (6.7)$$

sont appelées **homogène** et **nonhomogène** respectivement.

L'équation homogène

Si les coefficients a, b et c sont constants, on sait qu'ils existent des solutions de la forme $u_n = x^n$ pour tout n (fonctions exponentielles). Pour trouver x on remplace x^n en (6.6) et on trouve que x doit satisfaire l'équation auxiliaire :

$$ax^2 + bx + c = 0. \quad (6.8)$$

Soient x_1 et x_2 les deux racines de l'équation de deuxième degré (6.8). On en déduit que la solution générale de (6.6) est toujours de la forme

1. Si $x_1 \neq x_2$

$$u_n = Ax_1^n + Bx_2^n,$$

2. Si $x_1 = x_2$,

$$u_n = Ax_1^n + Bnx_1^n,$$

avec des constantes A et B .

Dans les deux cas A et B doivent être déterminées à partir des conditions supplémentaires sur la frontière.

L'équation nonhomogène

La résolution du problème nonhomogène (6.7) comporte quatre pas :

1. Trouver une base pour l'espace vectoriel des solutions de l'équation auxiliaire homogène (6.6), et donc la solution générale u_n pour cette équation.
2. Déterminer une solution particulière de (6.7), par exemple en utilisant une expression «essai» \tilde{v}_n qui a la même forme générale que le membre droit d_n , mais des coefficients non déterminés. Par exemple, si d_n est un polynôme d'ordre k , on essaie un polynôme général d'ordre k .
3. Néanmoins, si votre expression d'essai a des termes qui sont inclus dans l'espace vectoriel des solutions de l'équation homogène obtenue au pas 1 (et donc qui vont être annihilés par l'opérateur des différences), il faut multiplier l'expression d'essai par n, n^2, \dots jusqu'à ce qu'il n'y a plus des termes inclus dans cet'espace .
4. Après la décision de la forme d'essai, on trouve les valeurs des coefficients de \tilde{v}_n à partir de (6.7), par la méthode des coefficients non déterminés.
5. La solution générale de (6.7) est de la forme $v_n = \tilde{v}_n + u_n$. On trouve finalement les coefficients encore non déterminés en u_n , en utilisant les conditions sur la frontière pour v_n .

Exemple 6.5.1 On considère l'ensemble E des suites $(u_n)_{n \in \mathbb{N}}$ qui vérifient la relation suivante :

$$(R) \quad \forall n \in \mathbb{N}, \quad u_{n+2} + \frac{3}{2}u_{n+1} - u_n = 0.$$

1. Rechercher les suites géométriques qui vérifient cette relation (R).
2. On note r_1 et r_2 leurs raisons et on admet que E est un espace vectoriel de dimension 2, c.-à-d. toute suite de E s'écrit sous la forme

$$u_n = \alpha r_1^n + \beta r_2^n, \quad \forall n \in \mathbb{N}.$$

Soit (a_n) la suite de E qui vérifie $a_0 = 1$ et $a_1 = 0$. Calculer a_n .

Exemple 6.5.2 On considère l'ensemble E' des suites (v_n) qui vérifient la relation :

$$(R') \quad \forall n \in \mathbb{N}, \quad v_{n+2} + \frac{3}{2}v_{n+1} - v_n = 4n + 1.$$

1. On pose $\forall n \in \mathbb{N}, u_n = an + b$. Déterminer a et b pour que (u_n) soit une solution particulière de (R') .
2. Soit (v_n) une suite de E' .
 - (a) Pour toute suite (t_n) de E' on définit la suite (u_n) par $\forall n \in \mathbb{N}, u_n = v_n - t_n$. Vérifier que $(u_n) \in E$.
 - (b) En déduire que $\forall n \in \mathbb{N}, v_n = \alpha r_1^n + \beta r_2^n + an + b$.
 - (c) Déterminer v_n pour $v_0 = -\frac{5}{9}$ et $v_1 = -\frac{26}{9}$.

Exemple 6.5.3 Obtenez les formules analytiques des suites décrites par les relations de récurrence ci-dessous, et vérifiez-les en utilisant les premiers termes de la suite t_2, t_3 .

1. $t_i = 2t_{i-1} + i - 1, t_0 = 0$
2. $t_i = 2t_{i-1} + 5 \cdot 2^i, t_0 = 0$
3. $t_i = 3t_{i-1} - 2t_{i-2} + 2, t_0 = 0, t_1 = 2$
4. $t_i = 2t_{i-1} - t_{i-2} + 2, t_0 = 0, t_1 = 2$

Solution :

1. C'est une équation nonhomogène, alors nous aurons :

$$t_i = \tilde{t}_i + A2^i, \quad \tilde{t}_i = c_1i + c_2 \quad \text{avec} \quad c_1i + c_2 = 2(c_1i | c_1 + c_2) + i - 1$$

et alors

$$\begin{aligned} c_1 &= 2c_1 + 1 & \text{et} & \quad c_1 = -1 \\ c_2 &= |2c_1 + 2c_2 - 1 & \text{et} & \quad c_2 = 2c_1 + 1 = -1 \\ \tilde{t}_i &= |i - 1 & \text{Finalement,} & \\ t_0 &= = 0 = -1 + A & \text{et} & \quad A = 1 \\ t_i &= |i - 1 + 2^i \end{aligned}$$

2. C'est une équation nonhomogène, alors :

$$t_i = \tilde{t}_i + A2^i, \quad \tilde{t}_i = ci2^i \quad \text{avec} \quad ci2^i = 2(c|i - 1|2^i/2) + 52^i$$

et alors

$$\begin{aligned} c &= 5, & t_i &= 5i2^i + A2^i & \text{et finalement,} \\ t_0 &= = 0 = A & \text{et} & \quad A = 0 \\ t_i &= 5i2^i \end{aligned}$$

3. C'est une équation de différences nonhomogène et l'équation quadratique attachée a les racines 1, 2, alors nous aurons :

$$t_i = \tilde{t}_i + A_12^i + A_2, \quad \tilde{t}_i = ci \quad \text{avec} \quad ci = 3(ci|c)|2(ci|2c) + 2$$

et alors

$$\begin{aligned} c &= -2 & \text{et} & \quad c_2 = 2c_1 + 1 = -1 \\ \tilde{t}_i &= |i - 1 & \text{Finalement,} & \\ t_0 &= = 0 = -1 + A & \text{et} & \quad A = 1 \\ t_i &= |i - 1 + 2^i \end{aligned}$$

4. C'est une équation de différences nonhomogène dont les racines de l'équation quadratique attachée sont confondues égales à 1 donc nous aurons :

$$t_i = \tilde{t}_i + A_1 + A_2 i, \quad \tilde{t}_i = c_1 i + c_2 \quad \text{avec} \quad c_1 i + c_2 = 2(c_1 i | c_1 + c_2) + i - 1$$

$$\begin{aligned} \text{et alors} \quad c_1 &= 2c_1 + 1 \quad \text{et} \quad c_1 = -1 \\ c_2 &= |2c_1 + 2c_2 - 1 \quad \text{et} \quad c_2 = 2c_1 + 1 = -1 \\ \tilde{t}_i &= |i - 1 \quad \text{Finalement,} \\ t_0 &= = 0 = -1 + A \quad \text{et} \quad A = 1 \\ t_i &= |i - 1 + 2^i \end{aligned}$$

6.5.2 La méthode des fonctions génératrices(*)

Exercice 6.5.1 Calculez, par la méthode des fonctions génératrices : a) $T(z) = \sum_{n \geq 0} T_n z^n$, où

$$T_0 = 0, \quad T_n = 2T_{n-1} + 1, n \geq 1$$

Trouvez T_n . b) Calculez $T(z) = \sum_{n \geq 0} T_n z^n$, où

$$T_0 = 1, \quad T_n = 2T_{n-1} + n - 1, n \geq 1$$

Trouvez T_n .

Sol : b) Par la méthode de décomposition des équations linéaires : $T_n = 2^{n+1} - (n+1)$. La fonction génératrice de T_n est $T(z) = 2/(1-2z) - 1/(1-z)^2 = (2z^2 - 2z + 1)/(1-2z)(1-z)^2$. En appliquant directement la méthode des fonctions génératrices à $T_n = 2T_{n-1} + n - 1$ on trouve l'équation : $T(z) = 1 + 2zT(z) + z/(1-z)^2 - (1/(1-z) - 1) = 2zT(z) + (2z^2 - 2z + 1)/(1-z)^2$, et on retrouve la même réponse. L'avantage de cette méthode plus compliquée est qu'elle réussit parfois quand la première méthode échoue.

Exercice 6.5.2 a) Obtenez une formule explicite pour la suite décrite par la relation de récurrence ci-dessous (et vérifiez-la en utilisant les premiers termes de la suite)

$$T_n = 4T_{n-1} - 3T_{n-2} + n, n \geq 1, \quad T_0 = a, T_{-1} = 0 \quad (6.9)$$

b) Calculez la fonction génératrice $T(z) = \sum_{n \geq 0} T_n z^n$ c) (*) Obtenez la fonction génératrice $T(z)$ directement, en appliquant la méthode des fonctions génératrices à l'équation (6.9).

Sol : a)

$$\begin{aligned} &(a + 3/4) 3^n - \frac{1}{4} (3 + 2n) \\ \frac{a + 3/4}{1 - 3z} - \frac{1/4}{1 - z} - \frac{1/2}{(z - 1)^2} &= \frac{a(z - 1)^2 + z}{(z - 1)^2(3z - 1)} \end{aligned}$$

Exercice 6.5.3

a) Obtenez une formule explicite pour la suite décrite par la relation de récurrence ci-dessous (et vérifiez-la en utilisant les premiers termes de la suite)

$$T_n = 2T_{n-1} - T_{n-2} + 2, n \geq 2, \quad T_0 = T_1 = 0 \quad (6.10)$$

b) Calculez la fonction génératrice $T(z) = \sum_{n \geq 0} T_n z^n$ c) Obtenez la fonction génératrice $T(z)$ directement, en appliquant la méthode des fonctions génératrices à l'équation (6.10)

Exercice 6.5.4 a) Obtenez la fonction génératrice $T(z) = \sum_{n \geq 0} T_n z^n$ pour la récurrence

$$T_n = \lambda T_{n-1} + \lambda_2 T_{n-2}, n \geq 2, \quad T_0 = 1, T_1 = \lambda T_0 \quad (6.11)$$

b) Calculez (directement, ou par développement limité) les premiers termes, et vérifiez qu'ils sont des combinaisons linéaires de puissances des racines caractéristiques.

Exercice 6.5.5 a) Obtenez une formule explicite pour la suite décrite par la relation de récurrence ci-dessous (et vérifiez-la en utilisant les premiers termes de la suite)

$$T_n = 4T_{n-1} - 3T_{n-2} + 2, n \geq 2, \quad T_0 = T_1 = 0 \quad (6.12)$$

b) Calculez la fonction génératrice $T(z) = \sum_{n \geq 0} T_n z^n$ c) (*) Obtenez la fonction génératrice $T(z)$ directement, en appliquant la méthode des fonctions génératrices à l'équation (6.13).

Solution :

1. a)

$$T_n = \frac{1}{2}(3^n - 1) - n$$

b)

$$T(z) = \frac{1}{2(1-3z)} - \frac{1}{2(1-z)} - \frac{z}{(1-z)^2} = \frac{1}{2(1-3z)} + \frac{1}{2(1-z)} - \frac{1}{(1-z)^2} = \frac{2z^2}{(1-z)^2(1-3z)}$$

La dernière expression peut être obtenue directement, à partir de la récurrence.

Exercice 6.5.6 a) Obtenez une formule explicite pour la suite décrite par la relation de récurrence ci-dessous (et vérifiez-la en utilisant les premiers termes de la suite)

$$T_n = 4T_{n-1} - 3T_{n-2} + n, n \geq 1, \quad T_0 = a, T_{-1} = 0 \quad (6.13)$$

b) Calculez la fonction génératrice $T(z) = \sum_{n \geq 0} T_n z^n$ c) (*) Obtenez la fonction génératrice $T(z)$ directement, en appliquant la méthode des fonctions génératrices à l'équation (6.13).

Solution : a)

$$(a + 3/4) 3^n - \frac{1}{4} (3 + 2n)$$

b)

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

6.6 Exercices

1. Trouvez

(a) l'élément suivant de la série y_t ci-dessous

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

(b) une équation de récurrence,

(c) un filtre qui annule la série

(d) Obtenez une formule explicite 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. Etudier les séries

$$y_t = \{1, 3, 7, 1, 3, 7, 1, 3, \dots\}$$

et

$$y_t = \{1, 3, -4, 1, 3, -4, 1, 3, \dots\}$$

3. 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 , par exemple 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 : $a(z) = \frac{1+2z}{(1+z)(1-z)^2} = \frac{3}{2(1-z)^2} - \frac{1/4}{(1-z)} - \frac{1/4}{(1+z)}$, $a(n) = (5 - (-1)^n + 6n)/4$ c) en utilisant le calcul (formel) opérationnel de différences de Arbogast et Heaviside (formel) http://en.wikipedia.org/wiki/Operational_calculus

Sol : $y_t = (1 - B)^{-2}(-1)^{t-1} = \sum_{n=0}^{\infty} (n+1)B^n(-1)^{t-1}$

4. 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 donnée par $m_t = f(B)Y_t$. **Sol** : $\frac{1+B+B^2+B^3}{4} \frac{5-3B}{2}$

5. a) Trouvez les conditions pour qu'un filtre causal $\theta(B) = \sum_{i=0}^q \theta_i B^i$ d'ordre q laisse passer une tendance linéaire sans distorsion. b) Trouvez le filtre de cette forme qui minimise la variance de $X_t = \sum_{i=0}^n \theta_i \epsilon_{t-i}$. Formuler ce problème comme un problème d'optimisation avec des contraintes, qu'on résoudra par la méthode des multiplicateurs de Lagrange. **Ind** : On rappelle que $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$. **R** : a) $\sum_i \theta_i = 1$, $\sum_i i\theta_i = 0$. b) $\text{Var } X_t = \sum_i \theta_i^2$. Elle doit être minimisée sous ces contraintes.

On trouve $\theta_i = \frac{\lambda + i\mu}{2}$ et les multiplicateurs de Lagrange satisfont

$$\lambda + \frac{n\mu}{2} = \frac{2}{n+1}$$

$$\lambda + \frac{(2n+1)\mu}{3} = 0$$

6. On considère un modèle simple où la tendance est une constante ($f(t) = a$).

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?

b) 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?

c)* Reprendre la question b) lorsque le nombre d'observations n'est pas un nombre entier de périodes : $n = Lp + m$.

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 : $n = Lp$. Montrer alors que les corrélations suivantes sont :

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

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

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

Chapitre 7

Revisison

7.1 Exercices

1. Calculer la fonction d'autocovariance 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.

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

Trouvez, pour le processus MA(1), les valeurs maximales $\rho_{1,M}$ et minimales $\rho_{1,m}$ de la corrélation $\rho_1(\theta)$ et les valeurs de θ pour les quelles ces valeurs sont atteintes.

R : $\rho_{1,M} = \rho_1(1), \rho_{1,m} = \rho_1(-1)$.

3. 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 $\theta(z) = 1 + \theta_1z + \theta_2z^2$, la condition pour racines réelles de module plus grand que 1 sont plus compliquées que les conditions (équivalentes) que le polynôme "reciproque" $\tilde{\theta}(z) = z^2\theta(1/z) = z^2 + \theta_1z + \theta_2$ ait des racines réelles z_i de module plus petit que 1. Pour ce dernier polynôme, les conditions sont :

(a) racines complexes : $|z_i|^2 = |z_1z_2| = \left|\frac{c}{a}\right| = |\theta_2| < 1$.

(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 region d'inversibilité est le triangle (θ_1, θ_2) :

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

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

$$\rho_1 = \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2} \quad \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} \quad (7.1)$$

Théorème 7.1.1 Le domaine des corrélations (ρ_1, ρ_2) pour les quelles il existent des processus MA(2) inversibles est

$$4\rho_2 \leq 1 + \sqrt{1 - 2\rho_1^2} \iff |\rho_1| \leq 2\sqrt{\rho_2(1 - 2\rho_2)} \quad (7.2)$$

$$\rho_2 + 1/2 \geq \rho_1 \quad (7.3)$$

$$\rho_2 + 1/2 \geq -\rho_1 \quad (7.4)$$

Dem : Les dernières deux inégalités sont obtenues respectivement en appliquant les inégalités $4xy \leq (x + y)^2$, $0 \leq (x + y)^2$ avec $x = \theta_1$, $y = 1 + \theta_2$ (elles représentent des courbes extrémales de la transformation (7.1)).

La première inégalité est obtenue a partir de $\theta_2 = 1$, en utilisant (7.1).

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

$$\rho_1^2 = 4\rho_2(1 - 2\rho_2) \iff 2\rho_1^2 + 16\rho_2^2 - 8\rho_2 + 1 = 1 \iff \frac{\rho_1^2}{1/2} + \frac{(\rho_2 - 1/4)^2}{1/16} = 1.$$

Il s'agit d'une ellipse des semiaxes $1/\sqrt{2}$, $1/4$, passant par les sommets $(0, 0)$, $(0, 1/2)$ et tangente aux droites $\rho_2 + 1/2 = \rho_1$, $\rho_2 + 1/2 = -\rho_1$ dans les points $(\pm 2/3, 1/6)$.

Similairement, a partir des frontières $\theta_1 \pm \theta_2 = 1$ on obtient les courbes ... qui sont intérieures au domaine (7.4).

4. a) Déterminez le 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 le corrélogramme 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) $MA(3)$: $\theta_1 = 0.8, \theta_2 = -0.4, \theta_3 = -0.3$

5. Investiguez si les deux processus $MA(2)$ sont inversibles, en examinant la récurrence obtenue par la méthode des coefficients indéterminés.
6. 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.
7. 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 : Rémarquez que pour le processus $ARMA(p,q)$, la recurrence pour $k > q$ est exactement comme pour le processus $AR(p)$; mais, les premières équations obtenues en comparant les coefficients des puissances en $\boxed{\psi(B) \phi(B) = \theta(B)}$ pour $k \leq q$ changent par l'ajout de θ_k de côté droite.
8. 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 ?
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 $AR(2)$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

(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 $\phi_2 + \phi_1 < 1, \phi_2 - \phi_1 < 1$ et dessus $\phi_2 = -1$.

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 $\theta^2\sigma^2$, $0 < |\theta| < 1$. 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.

7.2 Contrôle continu d'entraînement

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

R : $\theta(z) = (1 + z + z^2) \left(\frac{2-z}{3} + \frac{2(1-z^2)}{9} \right)$.

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 (c.-à-d. 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 $BB(0, \sigma^2 = 1)$.

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

5. La série binomiale négative

- Calculer les développements limités autour de $x = 0$ de $\frac{1}{(1-x)^2}$ et de $\frac{1}{(1-x)^3}$, en différenciant $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$.
- Quelle est le développement limité de la "fonction binomiale négative" $\frac{1}{(1-x)^3}$?
- Est-ce que le processus Y_t défini par :

$$\left(1 - \frac{3}{2}B + \frac{3B^2}{4} - \frac{B^3}{8}\right)Y_t = \epsilon_t$$

est stationnaire causal ? Si oui, donner une formule générale pour les coefficients ψ_k de sa représentation comme processus $MA(\infty)$.

Solutions :

- $1 - \frac{1}{2}z - \frac{3}{16}z^2 = (1 - \frac{3}{4}z)(1 + \frac{1}{4}z)$. $\psi_k = (\frac{3}{4})^{k+1} - (-\frac{1}{4})^{k+1}$ Oui, le processus est causal, car les racines duales $\lambda_1 = \frac{3}{4}$, $\lambda_2 = -\frac{1}{4}$ sont dedans le cercle unitaire $|\lambda| \leq 1$.
- En demandant que $\theta(B) = \frac{1+B+B^2}{3}(a_0 + a_1B + a_2B^2)$ satisfait $\theta(1) = 1$, $\theta'(1) = 0$, $\theta''(1) = 0$, on trouve : $\theta(B) = \frac{1}{9}(8 - 7B + 2B^2)(1 + B + B^2) = \frac{8}{9} + \frac{B}{9} + \frac{B^2}{3} - \frac{5B^3}{9} + \frac{2B^4}{9}$.
- a) $\text{Var } X_t = \sum_i \theta_i^2$. b) Elle est minimisée sous la contrainte $\sum_i \theta_i = 1$ par la moyenne arithmétique à coefficients égaux.
- (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 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 $(1 - .2B - .35B^2)Y_t = (1 - .7B)(1 + .5B)X_t = \epsilon_t$
- (c) Pour un processus stationnaire causal (avec les racines du polynôme $\phi(B)$ à l'extérieur du disque unité), les coefficients $\psi(k)$ satisfont les équations de Yule Walker

$$\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2}, \psi_{-1} = 0, \psi_0 = 1.$$

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 (8.2), où par (8.4), 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

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$$

Pour $k = 1, 2$ on a 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, $\gamma_0 = \frac{1}{1 - \sum \phi_i \rho_i} = 2.52$.

- (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 .

7.3 Contrôle continu

- (a) Trouvez un filtre causal $\theta(B) = \sum_{i=0}^2 \theta_i B^i$ qui conserve les polynômes de degré ≤ 1 et qui enlève les composantes saisonnières d'ordre 2.
- (b) Trouvez un filtre causal $\theta(B) = \sum_{i=0}^3 \theta_i B^i$ qui conserve les polynômes de degré ≤ 1 , qui enlève les composantes saisonnières d'ordre 2, et qui minimise la variance des moyennes mobiles $Y_t = \theta(B)\epsilon_t$.

2. Est-ce que le processus Y_t défini par :

$$(1 - B + \frac{B^2}{4})Y_t = \epsilon_t$$

est stationnaire causal ? Si oui, donner les coefficients ψ_k de sa représentation comme processus $MA(\infty)$.

3. On considère le processus aléatoire suivant :

$$X_t = 20 - \frac{2}{5} X_{t-1} + \frac{3}{25} X_{t-2} + \epsilon_t$$

ou ϵ_t est $BB(0, \sigma^2 = 1)$.

- (a) Calculer l'espérance $m = \mathbb{E}X_t$ de X_t , en supposant que le processus est stationnaire. Quelle est l'équation de récurrence satisfaite par le processus $Y_t = X_t - m$?
- (b) Donner les équations de Yule-Walker pour les autocorrélations d'un processus $AR(2)$. Calculer les 3 premières autocorrélations $\rho_i, i = 1, 2, 3$ du processus X_t , ainsi que la variance.

4. Soit Y_n, Y_{n-1}, \dots, Y_0 une suite qu'on suppose être bien approximée comme $AR(1)$

$$Y_t = \phi Y_{t-1} + \epsilon_t, t = 1, 2, \dots, n$$

- (a) On propose de prendre pour ϕ la valeur qui minimise la somme des erreurs $E = \sum_{t=1}^n \epsilon_t^2$. Quelle est la valeur de ϕ qui minimise cet objectif ?
- (b) On suppose maintenant que notre suite peut être bien approximée comme $AR(2)$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t, t = 2, 3, \dots, n$$

Trouver un système d'équations satisfaites par les valeurs ϕ_1, ϕ_2 qui minimisent l'objectif $E = \sum_{i=2}^n \epsilon_t^2$?

Solutions :

1. (a) $\phi(z) = 1 - a(1 - z)^2$ et $\phi(-1) = 0$ donnent $\phi(z) = 1 - \frac{(1-z)^2}{4} = \frac{1}{2} \left(\frac{3}{2} - \frac{z}{2} \right) (z + 1) = -\frac{z^2}{4} + \frac{z}{2} + \frac{3}{4}$

- (b) Il faut résoudre
$$\begin{cases} \min \sum_{i=0}^3 \theta_i^2 \\ \sum_{i=0}^3 \theta_i = 1 \\ \sum_{i=1}^3 i\theta_i = 0 \\ \sum_{i=0}^3 \theta_i (-1)^i = 0 \end{cases} \quad \text{Il est plus simple d'utiliser}$$

une des deux paramétrisations utilisant les factorisations connues

$$\theta(z) = (1+z)(a_0 + a_1 z + a_2 z^2) \iff \theta(-1) = 0, \quad \theta(z) = 1 + (z-1)^2(b + b_1 z),$$

surtout la deuxième. Alors, $\theta(-1) = 0 \implies b_1 = b + 1/4 \implies \theta(z) = 1 + (z-1)^2(b + (b+1/4)z) = (b + \frac{1}{4})z^3 - (b + \frac{1}{2})z^2 - (b - \frac{1}{4})z + b + 1$. Il faut minimiser $(b + \frac{1}{4})^2 + (b + \frac{1}{2})^2 + (b - \frac{1}{4})^2 + (b + 1)^2$, le minimum est quand $2b + b + \frac{1}{2} + b + 1 = 0, b = -\frac{3}{8}$, et le filtre est $\frac{1}{2}(z+1)\left(\frac{5}{4} - \frac{z^2}{4}\right) = -\frac{z^3}{8} - \frac{z^2}{8} + \frac{5z}{8} + \frac{5}{8}$ On peut vérifier que la variance $\frac{13}{16}$ a été réduite par rapport à $\frac{7}{8}$ dans la question précédente.

2. Un processus est stationnaire causal (avec les racines du polynôme $\phi(B)$ à l'extérieur du disque unité) ssi les coefficients $\psi(k)$ définies par l'équation Yule Walker

$$\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2}$$

avec valeurs initiales $\psi_{-1} = 0, \psi_0 = 1$ satisfont $\sum_k \psi^2(k) < \infty$. La solution dans le cas des racines distinctes est

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

et avec racines confondues

$$\psi(k) = (k+1)\lambda^k$$

(en resolvant la recurrence YW, ou en differenciant $\frac{1}{(1-x)^1} = \sum_{k=0}^{\infty} x^k$).

Ici, $\psi_k = \frac{k+1}{2^k}$

3. (a) Ce processus AR(2) pas centré peut s'écrire $(1 + \frac{2}{5}B - \frac{3}{25}B^2)X_t = (1 - \frac{1}{5}B)(1 + \frac{3}{5}B)X_t = 20 + \epsilon_t$ En prenant espérance on trouve $\mathbb{E}(X_t) = \frac{20}{1 + \frac{2}{5} - \frac{3}{25}} = \frac{125}{8}$. Soit $Y_t = X_t - \mathbb{E}X_t$. Alors Y_t est un processus AR(2) centré satisfaisant $(1 + \frac{2}{5}B - \frac{3}{25}B^2)Y_t = (1 - \frac{1}{5}B)(1 + \frac{3}{5}B)Y_t = \epsilon_t$
- (b) La fonction d'autocorrélation de Y_t (qui est la même que celle de X_t) satisfait l'équation Yule Walker

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$$

Pour $k = 1, 2$ on a par symmetrie 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 $\rho_1 = -\frac{5}{11} = -.45$, et la rec. YW donne $\rho_2 = \frac{83}{275} = .302, \rho_3 = -\frac{241}{1375} = -.175$. Finalement, $\gamma_0 = \frac{1}{1 - \sum \phi_i \rho_i} = 1.2788$.

Chapitre 8

Prévision linéaire des modèles autorégressifs ARIMA(p,d,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. $t + k$ 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 linéaire, 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 8.0.1 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 8.0.1 Si un processus Y_t défini par une autorégression (4.5) $Y_t = f(Y_{t-1}, Y_{t-2}, \dots) + \epsilon_t$ 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,

$$\mathbb{E}[\epsilon_t | [Y_{t-1}, Y_{t-2}, \dots]] = 0 \quad (8.1)$$

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 (4.6) 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 :

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

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

8.1 Prédiction des processus AR(p) par la récurrence de Yule-Walker

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

Théorème 8.1.1 Pour un modèle AR(p)

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

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

$$\hat{X}_t(k) := \mathbb{E}[X_{t+k} | \{X_t, X_{t-1}, \dots\}] = \sum_{i=1}^p \phi_i \hat{X}_t(k-i), \quad \hat{X}_t(k) = X_{t+k}, \quad \text{si } k \in \{0, -1, -2, \dots\} \quad (8.2)$$

En particulier,

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

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

Exemple 8.1.1 La prévision linéaire $X_t(k)$ pour un processus $AR(1)$ à moyenne 0 satisfait la recursion Yule Walker

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

étant donc simplement

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

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

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

Théorème 8.1.2 Soit

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

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

$$X_t(k) = A_1(t)\lambda_1^k + A_2(t)\lambda_2^k = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} X_t - \lambda_1 \lambda_2 \frac{(\lambda_1^k - \lambda_2^k)}{\lambda_1 - \lambda_2} X_{t-1} \quad (8.3)$$

où encore

$$X_t(k) = \psi_k X_t - \lambda_1 \lambda_2 \psi_{k-1} X_{t-1} \iff X_t(k) = \psi_k X_t + \varphi_2 \psi_{k-1} X_{t-1}, \quad k = 0, 1, \dots \quad (8.4)$$

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

$$\hat{X}_t(k) - \phi_1 \hat{X}_t(k-1) - \phi_2 \hat{X}_t(k-2) = \hat{X}_t(k) - (\lambda_1 + \lambda_2) \hat{X}_t(k-1) + \lambda_1 \lambda_2 \hat{X}_t(k-2) = 0 \quad (8.5)$$

et donc

$$\hat{X}_t(k) = \phi_1 \hat{X}_t(k-1) + \phi_2 \hat{X}_t(k-2),$$

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 $\psi(k) = \phi_1 \psi(k-1) + \phi_2 \psi(k-2)$, 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 8.1.1 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 $\phi(z) = 1$, ou des dernières valeurs observées $X_t, X_{t-1}, \dots, X_{t-p+1}$.

Théorème 8.1.3 La fonction génératrice $X(z) := \sum_0^\infty z^k \hat{X}_t(k)$ des prévisions d'un processus $AR(p)$ est

$$X(z) = \frac{X_t + z(\phi_2 X_{t-1} + \phi_3 X_{t-2} + \dots) + z^2(\phi_3 X_{t-1} + \phi_4 X_{t-2} + \dots) + \dots + z^{p-1} \phi_p X_{t-1}}{1 - \phi(z)}$$

Exemple 8.1.2 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 $\hat{X}_t(1) = \phi_1 X_t + \phi_2 X_{t-1}$, $\hat{X}_t(2) = (\phi_1^2 + \phi_2)X_t + \phi_1 \phi_2 X_{t-1}, \dots$ et en général is s'agit des coefficients du développement limité du théorème 8.1.3

$$\frac{X_t + z\phi_2 X_{t-1}}{1 - \phi(z)} = X_t + z(\phi_1 X_t + \phi_2 X_{t-1}) + z^2((\phi_1^2 + \phi_2)X_t + \phi_1 \phi_2 X_{t-1}) + \dots$$

On peut aussi résoudre les équations YW, ou utiliser la formule (8.4).

8.2 Prévision des processus MA(1)

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

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

Exemple 8.2.1 Pour le processus $MA(1)$, on vérifie facilement que $X_t(k) = 0$ pour $k = 2, 3, \dots$, (pour une généralisation, voir le théorème 8.6.1). 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 :

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

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, c.-à-d. $F_t = \{X_t, X_{t-1}, \dots, X_1\}$. 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 $q \geq 1$.

8.3 Prédiction des processus ARMA(p,1)

Exercice 8.3.1 On considère le processus ARMA(1,1) à moyenne 0 $(1 - \phi B)Y_t = (1 + \theta B)\epsilon_t$ où $-1 < \phi < 1$ et $-1 < \theta < 1$.

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

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

- On utilise ce modèle pour ajuster une série et on obtient comme estimations des paramètres $\phi = 0.8$, $\theta = 0.3$ 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 8.3.2 Est-ce que les résultats ci-dessus restent vrais si $\phi = 1$, 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).

1. Montrez qu'on peut le représenter :

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

2. 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).

8.4 Prévision des processus $ARIMA(p,d,0)$

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 8.4.1 La prévision linéaire $X_t(k)$ pour le processus $ARIMA(0,1,0)$ à moyenne c satisfait la recursion 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 8.4.2 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 $\phi(z)$, et on verra que ça reste vrai pour les processus ARIMA(p, d, q), pour $k > q$.

Proposition 8.4.1 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 recursion $\phi(B)X_t(k)$, 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 8.4.3 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 8.4.1 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 recursion de Yule Waker devient valable) s'appellent les **valeurs pivots**.

Il suit clairement que :

Proposition 8.4.2 La prévision "eventuelle" de Box-Jenkins pour les processus ARIMA(p, d, q) est la fonction dans l'espace linéaire des solutions de la recursion $\phi(B)X_t(k)$ qui passe par les valeurs pivots.

Corollaire 8.4.1 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 derniers points.

8.5 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 $\epsilon_{-1}, \epsilon_{-2}, \dots$, c.-à-d. 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 précédant le début d'observations 0 ; l'approximation la plus simple dans l'absence des moyennes est $\epsilon_k = Y_k = 0$ pour $k < 0$.

Théorème 8.5.1 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 $\phi(B)(1-B)^d$ (dans le cas d'un modèle ARMA(p, q) $\tilde{\varphi}_i = \phi_i$). Pour $k > q$, cette formule est exactement la récurrence homogène Yule-Walker $\phi(B)\hat{X}_t(k) =$, 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.

Projet : Il serait intéressant de déterminer analytiquement "la tendance asymptotique", c.-à-d. le polynôme d'ordre $d - 1$ vers le quel les prévisions convergent asymptotiquement pour les processus ARIMA(p, d, q). Considérons par exemple ARIMA($p, 2, 0$); ce modèle inclue une tendance linéaire, pour laquelle le premier candidat est la droite par les deux derniers points pivots (comme dans le cas $p = 0$, quand les pivots coïncident 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 précèdent les 2 derniers points auront aussi une influence sur la "droite de tendance".

8.6 Intervalles de confiance pour la prédiction

Théorème 8.6.1 (*) 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 donné 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 , c.-à-d. $X_t = \psi(B)\epsilon_t$ 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 $ARIMA(p, d, q)$.

8.7 HoltWinters, forecast

Passons maintenant à une décomposition (régression par rapport à un certain modèle) + prédiction très basique, de *HoltWinters*.

```
rain <- scan("http://robjhyndman.com/tsdldata/hurst/precip1.dat", skip=1)
rts <- ts(rain)
plot(rts)
rsH <- HoltWinters(rts, beta=FALSE, gamma=FALSE)
rsH
plot(rsH)
rsH$SSE
```

Pour rajouter des prédictions, on a besoin de *library(forecast)*

```
library(forecast)#install.packages("forecast", repos="http://R-Forge.R-project.org")
rs2 <- forecast.HoltWinters(rsH, h=14)
plot.forecast(rs2)
#HoltWinters ne sauve pas les résidues, mais on peut les récupérer avec
#res=rsH-rsH$fitted"
```

Nouveau essai, avec β :

```
skirts <- scan("http://robjhyndman.com/tsdldata/roberts/skirts.dat", skip=5)
skirtsseries <- ts(skirts, start=c(1866))
plot.ts(skirtsseries)
ssH <- HoltWinters(skirtsseries, gamma=FALSE)
plot(ssH)

s2 <- forecast.HoltWinters(ssH, h=19)
plot.forecast(s2)

rsforecasts <- HoltWinters(rts, gamma=FALSE)
rs2 <- forecast.HoltWinters(rsforecasts, h=14)
plot.forecast(rs2)
```


Chapitre 9

Sommaire des quelques définitions et résultats dans les séries temporelles

9.1 Filtres

Définition 9.1.1 a) Une série s_t sera appelée saisonnière de période p ssi

$$\sum_{i=1}^p s_{t+i} = 0 \iff \left(\sum_{i=0}^{p-1} B^i \right) s_t = 0, \quad \forall t \quad (9.1)$$

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

$$s_{t+p} = s_t \iff (1 - B^p) s_t = 0, \quad \forall t \quad (9.2)$$

Théorème 9.1.1 a) Un filtre $\psi(B)$ annule (ou "enlève") les composantes périodiques d'ordre p ssi $\psi(z)$ est divisible par $1 - z^p$.

b) 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 9.1.2 Un filtre $\psi(B)$ laisse invariants les polynômes de degré $\leq p$ ssi $1 - \psi(B)$ 1 les polynômes de degré $\leq p$, et ssi 1 est une racine d'ordre au moins $p + 1$ de l'équation $\psi(z) = 1$, i.e. $\psi(1) = 1, \psi'(1) = 0, \psi''(1) = 0, \psi^{(p)}(1) = 0$.

9.2 Équations Yule Walker pour AR(p)

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

$$\mathbf{R} \boldsymbol{\phi} = \boldsymbol{\rho} \quad (9.3)$$

ou R est la matrice Toeplitz symétrique :

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

2. En suite, pour $k > p$ on utilise la récurrence : $\rho_k = \sum_{i=1}^p \phi_i \rho_{k-i}$
3. La variance est $\gamma_0 = \frac{\sigma^2}{1 - \sum_i \phi_i \rho_i}$ (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)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

avec racines λ_1, λ_2 de $0 = \lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 \phi(\lambda^{-1})$ dedans le cercle unitaire (pour assurer la causalité), on obtient, en résolvant (9.3) :

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 \lambda_2}, \quad \rho_2 = \phi_2 + \frac{\phi_1^2}{1 - \phi_2}, \dots$$

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

$$\rho_k = \rho_1 \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} - \lambda_1 \lambda_2 \frac{\lambda_1^{k-1} - \lambda_2^{k-1}}{\lambda_1 - \lambda_2} = \boxed{\rho_1 \psi_{k-1} + \phi_2 \psi_{k-2}}, k \geq 1, \quad \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

Finalement, en utilisant $\phi_1 = \lambda_1 + \lambda_2, \phi_2 = -\lambda_1 \lambda_2$, 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 :

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

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

Théorème 9.3.1 a) Un processus ARMA(p) $\phi(B)Y_t = \theta(B)\epsilon_t$ 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 $\phi(z)$ ($\theta(z)$) sont à l'extérieur du cercle unitaire. Les coefficients ψ_i ($\pi(i)$) sont dans ce cas les coefficients de la série Taylor de $\psi(z) = \frac{\theta(z)}{\phi(z)}$ ($\pi(z) = \frac{\phi(z)}{\theta(z)}$)

Théorème 9.3.2 (*) a) Pour un processus ARMA(p, q) $\phi(B)Y_T = \theta(B)\epsilon_t$ avec toutes les racines du polynôme caractéristique $\phi(z)$ 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

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

Note : a) Cette dernière équation est appelée **recursion de Yule-Walker**.
b) Pour un processus ARMA(p, q) $\phi(B)Y_T = \theta(B)\epsilon_t$ 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

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

Exemple 9.3.1 Ainsi, $\psi_1 = \phi_1 + \theta_1, \psi_2 = \phi_2 + \phi_1^2 + \theta_1\phi_1 + \theta_2, \psi_3 = \phi_3 + 2\phi_1\phi_2 + \phi_1^3 + (\phi_2 + \phi_1^2)\theta_1 + \phi_1\theta_2 + \theta_3, \dots$

9.4 Prévision 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 $\phi(B)(1-B)^d$, et dans le cas d'un modèle ARMA(p, q) $\tilde{\varphi}_i = \phi_i$. 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 $\phi(B)\hat{Y}_t(k) = \theta(B)\epsilon_{t+k}$ - voir Thm 5.1.2.

Chapitre 10

Méthodes et cas, Yves Aragon

10.1 Trafic mensuel de l'aéroport de Toulouse-Blagnac

Quelques commandes utiles : *str(ucture)*, *summary*, *aggregate*, *window*, *monthplot*, *lag.plot*, *acf*, *pacf*, *xy.acfb*, *diff*, *arima*, *Box.test.2*, *auto.arima*

```
# 8.1 Préparation des données
require(caschrono)
aa=read.table(file= system.file("/import/trafquoti.txt",package="caschrono"),
header=FALSE,quote="",sep="", colClasses=c('numeric','character'),
col.names =c('trafic','date'))
str(aa)
summary(aa)

# agrégation par mois
an = substr(aa$date,1,4)
mois =substr(aa$date,6,7)
mois.an= as.numeric(paste(an,mois,sep=""))
trafmens=aggregate(aa$traf, list(Mois.An = mois.an), sum)
str(trafmens)
trafmensu=ts(trafmens$x/1000,start= c(1993,1),frequency= 12)

# 8.2.1 Décomposition de la série en tendance, saisonnalité et erreur
dec.m = decompose(trafmensu)
op=par(oma=rep(0,4))
plot(dec.m)
abline(v=2001.75)
par(op)

# 8.2.2 Month plot
op=par(oma=rep(0,4))
monthplot(trafmensu, main="Month plot du trafic")
par(op)
```

```

# trafic avant septembre 2001
trafav=window(trafmensu, start=c(1996,1), end=c(2001,8))
summary(trafav)
temps = time(trafav)
traf.apr01=window(trafmensu, start=c(2001,10), end=c(2007,10))
time.apr01=time(traf.apr01)

# 8.3 Modélisation MCO avant septembre 2001
(mod1 = lm(trafav~temps))
resid= ts(residuals(mod1),start=c(1996,1),frequency=12)

op=par(mfrow=c(3,1),oma=rep(0,4))
par(mar=c(0,4,1,0))
plot(trafav,xlab="", ylab="trafic",xaxt="n")
par(mar=c(4,4,0,0))
plot(resid, xlab="temps",ylab="résidu MCO")
abline(h=0)
par(mar=c(4,4,2,0))
acf(resid, lag.max=50, xlab="retard",main="")
#correlations 12, 24, ... suggerent periode 12
par(op)

# 8.2.3 Lag plot
op=par(oma=rep(0,4))
lag.plot(rev(trafav),set.lags=1:12,main="Avant le 9/11",
  asp=1,diag =TRUE, diag.col="red",type="p",do.lines=FALSE)
par(op)

op=par(oma=rep(0,4))
xy.acfb(diff(trafav,12),40,numer=FALSE)
par(op)
# pacf differentie D=12 suggere SARIMA(1,1,0)

# 8.3.1 Modélisation Arima manuelle

(mod2=Arima(trafav,seasonal=list(order=c(1,1,0),period=12),method='ML',
  include.drift=TRUE))
#autre alternative possible a analyser: SARIMA(2,0,0)
op=par(oma=rep(0,4))
xy.acfb(mod2$residuals,numer=FALSE)
par(op)
#drift=1.9 confirme que la variation des fluctuations n'est pas autour du 0
#l'autocor principale qui reste est au lag 2
ret= seq(6,30,6) #test blancheur sur somme des corr. rejette
(Box.test.2(mod2$residuals,ret,type="Ljung-Box",fitdf=2,decim=4))

```

```

#modele plus complexe
(mod3=Arima(trafav,order=c(2,0,0),seasonal=list(order=c(1,1,0),period=12),
  method='ML',include.drift=TRUE))
#premier coef pourrait etre 0
t_stat(mod3)
#coefficient 1 douteux

op=par(oma=rep(0,4))
xy.acfb(residuals(mod3),numer=FALSE)
par(op)

ret= seq(6,30,6) #test blancheur accepte
(Box.test.2(residuals(mod3),ret,type="Ljung-Box",fitdf=4,decim=4))

#fixons premier coef to 0
(mod4=Arima(trafav,order=c(2,0,0),
  seasonal=list(order= c(1,1,0),period=12),
  fixed=c(0,NA,NA,NA),method='ML',include.drift=TRUE))
#summary(mod4)

op=par(oma=rep(0,4))
xy.acfb(residuals(mod4),numer=FALSE)
par(op)

(Box.test.2(residuals(mod4),ret6,type="Ljung-Box",fitdf=3,decim=4))
#test blancheur pire que pour mod3, meme AIC, BIC ameliore

t_stat(mod4)
#tous les coefficients significatifs

# 8.3.2 Modélisation automatique
best.av = auto.arima(trafav, D = 1)
summary(best.av) # ARIMA(0,0,0)(0,1,1)[12] with drift
#livre obtient ARIMA(1,0,1)(0,1,1)[12] with drift
t_stat(best.av)
#coefficient 1 super douteux

(Box.test.2(residuals(best.av),ret6,type="Ljung-Box",decim=4,fitdf =2))

#Bande de prédiction pour mod4
ec80 = mod4$sigma2^.5 * qnorm(0.90)
vajust = fitted(mod4)
matri=as.ts(cbind(trafav,vajust-ec80,vajust+ec80),start=c(1996,1),
  frequency=12)

op=par(oma=rep(0,4))

```

```

plot(matri, plot.type='single', lty=c(1,2,2), xlab="temps",
     ylab='trafic', main="", cex.main=0.8 )
legend( par("usr")[1], par("usr")[4], c("Val. observée","Bande de prédiction"),
       lwd=1, lty=c(1,2) )
par(op)

indi = (trafav - (vajust-ec80)) > 0 & ( vajust+ec80 - trafav) > 0
(prop = 100*sum(indi)/length(indi))
#proportion dans la bande .88 >.80 modele surparametre?

#Dickey-Fuller Unit Root Test rejette non-station avec p-val 1% ?
require(urca)
ur1 = ur.df(diff(trafav,12),lags=1,type="drift")
summary(ur1)

#KPSS Test ne rejette pas stationarite avec p-val 10% ?
require(tseries)
kpss.test(diff(trafav,12))

# 8.4.1 Prévision ponctuelle
prev2002 = forecast(mod4, h=16, level=80)

apredire=window(trafmensu,start=c(2001,9),end=c(2002,12))
matri2 = cbind(apredire, prev2002$lower, prev2002$upper,prev2002$mean)

plot(matri2,plot.type='single',lty=c(1,2,2,3),
     xlab="temps",xaxt='n',ylab='trafic')
axis(1,at=c(2001.7,2002,2002.9),
     labels=c("sept-2001","jan-2002","dec-2002"),
     cex.axis=.8)
legend("topleft",c("Valobs. ","Bandepréd. 80%","Préd. moy"),lwd=1, lty=c(1,2,3),cex=.8 )

```

10.2 Consommation d'électricité, cf. Méthodes et cas, Yves Aragon

```

require(caschrono)
require(fBasics)
require(xtable)
require(dse)
require(polynom)
data(khct)
??khct
# htdd (heating days) is the sum over the month of the
# daily difference between 65 F. and the average daily temperature,
# if lower than 65 F. degrees =\sum (65-t_i)_+

```

```

# cldd (cooling degree days) =\sum (t_i-65)_+
# source: Pankratz:Forecasting with dynamic regression models,1991
plot.ts(khct,xlab='temps',main="",cex.lab=.9,cex.axis=.8,
  oma.multi=c(4.5,3,.2,0),mar.multi=c(0, 4, 0, .5),las=0)
# Remarquer la complementarite des deux regresseurs

ytr= cbind(khct[,"kwh"], log(khct[,"kwh"]), (khct[,"kwh"])^0.5)
colnames(ytr)= c("kwh", "log(kwh)","kwh^0.5")
my.panel<- function(x, y,..., pf = parent.frame()) {
  lines(x, y, ...)
  abline(lm(y~x),lty=2)
}
# On diminue la heteroscedasticite avec deux transformations
# Mais log induit une courbure
plot(zoo(ytr),ylab=c("kwh","log(kwh)",expression(sqrt(kwh))),
  main="", xlab='temps',
  panel=my.panel,cex=1.2)

khct.df<-as.data.frame(window(cbind(khct,time(khct)),
  end=c(1983,12)))
colnames(khct.df) <- c("kwh","htdd","cldd","t1")
mod1 = lm(sqrt(kwh) ~ htdd +cldd+ t1, data=khct.df)
# Examen des résidus du modèle lineaire montre une legere parabole
# avec sommet en 1977. On ajoute donc un predicteur quadratique
khct.df$t1.2 = (khct.df$t1 - 1977)^2
mod2 = lm(sqrt(kwh) ~ htdd + cldd + t1 + t1.2, data=khct.df)

resi.ts=ts(cbind(residuals(mod1),residuals(mod2)),start=c(1970,1),frequency=12)
my.panel2<- function(x, y,..., pf = parent.frame()) {
  lines(x, y, ...)
  abline(h=0,lty=3,lwd=2)
}

plot(zoo(resi.ts),ylab=c("Ajust. linéaire",
  "Ajust. quadratique"), main="Examen des résidus",
  xlab='temps', panel=my.panel2,cex.lab=1.4,cex.main=1.4,cex=1.4)
# les residues montrent toujours correlation positive, suggerant
# une modelisation ARIMA

u = ts(residuals(mod2), start=c(1970,1), frequency=12)
acf2y(u,numer=FALSE)

```

Le résultat suggère un modèle autoregressif aux lags 1 et 12.

```

(mod.auto=auto.arima(u,max.p=4,max.q=4,max.P=1,approximation=FALSE))
llag = seq(6,30,6)
(Box.test.2(residuals(mod.auto),llag,type="Ljung-Box",decim=2,fitdf=4))

```


Le résultat sar1=0.983, avec s.e. = 0.017, proche de non-stationnarité, annonce des problèmes possibles.

```
kwh1rc = window(sqrt(khct[, "kwh"]), end=c(1983,12))  
xreg1 = khct.df[ ,c("htdd", "cldd", "t1", "t1.2")]
```

```
mdarx1=Arima(kwh1rc, order=c(1,0,1), seasonal=list(order=c(1,0,1)),  
xreg=xreg1)
```

En effet, l'estimation simultanée ARMAX passe trop proche de non-stationnarité

```
(mdarx2=Arima(kwh1rc, order=c(1,0,0), seasonal=list(order=c(1,1,1)),  
xreg=xreg1))
```

```
(Box.test.2(residuals(mdarx2), seq(6,30,6), type="Ljung-Box",  
decim=2, fitdf=8))
```

```
t_stat(mdarx2)
```

NAN pour s.e. de t1 indique un modèle mal spécifié

```
xreg1b = xreg1[,-3]  
(mx2b=Arima(kwh1rc, order=c(1,0,0), seasonal=list(order=c(1,1,1)),  
xreg = xreg1b, include.drift= TRUE))  
(Box.test.2(residuals(mx2b), seq(6,30,6), type = "Ljung-Box", decim=2, fitdf=7))
```

```
t_stat(mx2b)
```

Bon test de blancheur, mais sar1 non-significatif

```
(mx2b=Arima(kwh1rc, order=c(1,0,0), seasonal=list(order=c(0,1,1)),  
xreg = xreg1b, include.drift= TRUE))  
(Box.test.2(residuals(mx2b), seq(6,30,6), type = "Ljung-Box", decim=2, fitdf=7))
```

```
t_stat(mx2b)
```

Essentiellement même tests de blancheur, et tous les coefficients significatifs

10.3 Température mensuelle moyenne à Nottingham, cf. Yves Aragon

10.4 Exercices

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) $(1 - \phi B)(1 - B)Y_t = (1 + \theta B)\epsilon_t$, avec $-1 < \phi < 1$ et $-1 < \theta < 1$.

(a) Montrez que $Y_t(1) = (1 + \phi)Y_t - \phi Y_{t-1} + \theta \epsilon_t$ et $Y_t(k) = (1 + \phi)Y_t(k-1) - \phi Y_t(k-2)$, $k \geq 2$.

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

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

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

(c) Montrez que $Y_t(1) = -\theta Y_{t-1}(1) + (1 + \phi + \theta)Y_t - \phi Y_{t-1}$ 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 $\sigma_\epsilon = 2$, calculez les prévisions $Y_{101|100}, Y_{102|100}$ de Y_{101} et Y_{102} et donnez 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.

10.5 Modèles d'espace-état

La modélisation stochastique distingue en général entre trois types des processus :

1. observations Y_t
2. variables "latentes" inobservables directement X_t , soumis parfois aux diverses contraintes, et qui doivent être estimées.
3. des bruits inobservables ϵ_t

Un exemple de modélisation très générale est fournie par les **modèles espace-état** :

$$X_{t+1} = A_t X_t + b_t + \epsilon_X(t) \quad \text{équation d'évolution} \quad (10.1)$$

$$Y_t = C_t X_t + d_t + \epsilon_Y(t) \quad \text{équation d'observation} \quad (10.2)$$

Exemple 10.5.1 Le modèle additif cherche une décomposition de la forme :

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

Donc, ce modèle contient seulement l'équation d'observation. m_t peut être traité aussi comme une variable latente, et avec l'ajout des informations sur son évolution, une deuxième équation d'évolution pourrait-être ajoutée.

Une telle décomposition devient utile ssi on peut assurer que les résidus aient une structure statistique simple comme "bruit blanc Gaussien" ou "bruit blanc de second ordre", moyennes mobiles, etc.

Chapitre 11

Examens

11.1 Examen 1

1. Trouvez un filtre causal $\theta(B) = \sum_{i=0}^3 \theta_i B^i$ qui conserve les polynômes de degré ≤ 2 et qui enlève les composantes saisonnières d'ordre 2.
2. Soit Y_t, Y_{t-1}, \dots, Y_0 une suite d'observations à allure stationnaire, qu'on aimerait modéliser comme $AR(p)$, avec p à déterminer.
 - (a) On considère d'abord une modélisation :

$$Y_t = \varphi Y_{t-1} + \epsilon_t, t = 1, 2, \dots, n \quad (11.1)$$

où on propose de prendre pour φ la valeur $\hat{\varphi}(Y_0, \dots, Y_n)$ qui minimise la somme des erreurs au carré $SC(\varphi) = \sum_{i=1}^n \epsilon_i^2$. Quelle est la valeur $\hat{\varphi}(Y_0, \dots, Y_n)$?

- (b) Supposons maintenant que Y_t sont en effet des observations d'un processus stationnaire $AR(1)$ satisfaisant (11.3), où ϵ_t est un bruit blanc (centré, non-corrélé) de variance $\sigma^2 = 1$. Quelle est la valeur de φ qui minimise l'objectif $E[\epsilon_t^2]$? Comment estimer cette valeur à partir des données observées ? Comparer avec la réponse précédente.
- (c) On suppose en suite que Y_t sont des observations d'un processus stationnaire $AR(2)$,

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \epsilon_t, t = 2, 3, \dots, n$$

où ϵ_t est un bruit blanc de variance $\sigma^2 = 1$. Trouver et résoudre le système d'équations satisfaites par les valeurs φ_1, φ_2 qui minimisent l'objectif $E[\epsilon_t^2]$.

- (d) On accepte finalement un modèle $AR(3)$ avec bruit blanc ϵ_t ($\sigma^2 = 1$) (avec coefficients $\varphi_1, \varphi_2, \varphi_3$ déterminés par les données). Donner les équations liant les autocorrélations ρ_1, ρ_2, ρ_3 d'un processus $AR(3)$ aux coefficients $\varphi_1, \varphi_2, \varphi_3$ (on ne demande pas de justifier). Calculer les 2 premières autocorrélations $\rho_i, i = 1, 2$ du processus Y_t , en fonction des coefficients $\varphi_1, \varphi_2, \varphi_3$.

(e) Calculer les autocorrélations partielles $\tilde{\rho}_i = \varphi_{(i,i)}$, $i = 3, 4$ du processus Y_t , en fonction des coefficients $\varphi_1, \varphi_2, \varphi_3$.

3. On considère des processus X_t, Y_t, Z_t satisfaisant respectivement :

$$(1 - B/2)X_t = \epsilon_t, \quad (1 - B)Y_t = \epsilon_t, \quad (1 - 2B)Z_t = \epsilon_t, \quad t = T, T - 1, \dots, 1, \quad (11.2)$$

où ϵ_t est un bruit blanc de loi $N(0, \sigma^2 = 1)$.

(a) Donner l'expression de X_T , de Y_T , et de Z_T en fonction de ϵ_t , $t = 1, 2, \dots, T$, et des valeurs au temps $t = 0$ (denotées par X_0, Y_0, Z_0).

(b) Est-ce que l'un de ces trois processus, en commençant de 0 (c'est-à-dire avec $X_0 = 0, Y_0 = 0, Z_0 = 0$) est stationnaire ?

(c) Est-ce que l'une des équations (11.2) admet une solution stationnaire ? Le cas échéant, donner cette solution.

Solutions :

1. $\varphi(z) = 1 - a(1 - z)^3$ et $\varphi(-1) = 0$ donnent $a = \frac{1}{8}$,

$$\varphi(z) = 1 - \frac{(1 - z)^3}{8} = \frac{z^3}{8} - \frac{3z^2}{8} + \frac{3z}{8} + \frac{7}{8} = \frac{1}{8}(z + 1)(z^2 - 4z + 7)$$

2. (a) $\hat{\varphi}(Y_0, \dots, Y_n) = \frac{\sum_{t=1}^n Y_t Y_{t-1}}{\sum_{t=0}^{n-1} Y_t^2}$

(b) $\hat{\varphi} = \frac{\gamma_1}{\gamma_0}$

(c) Pour $p = 2$, on arrive au système YW :

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2$$

avec solution

$$\varphi_{2,1} = \frac{\begin{vmatrix} \rho_1 & \rho_1 \\ \rho_2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2} = \frac{\gamma_0 \gamma_1 - \gamma_1 \gamma_2}{\gamma_0^2 - \gamma_1^2} \quad \varphi_{2,2} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\gamma_0 \gamma_2 - \gamma_1^2}{\gamma_0^2 - \gamma_1^2}$$

(remarquer les déterminants de Hankel dans les expressions par rapport au γ_i). On s'aperçoit en résolvant cet exercice que les $\varphi_i, i = 1$, sont exactement les coefficients de la régression de Y_t sur Y_{t-1}, Y_{t-2} (d'où le nom autoregression), et que pour $p = 1$, y a coïncidence entre les valeurs du coefficient de régression φ et de la corrélation ρ_1 , mais pour $p \geq 2$, la situation se complique.

(d) Pour $p = 3$, le système YW est :

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1 + \varphi_3 \rho_2$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2 + \varphi_3 \rho_1$$

$$\rho_3 = \varphi_1 \rho_2 + \varphi_2 \rho_1 + \varphi_3$$

Les deux premières équations

$$\begin{pmatrix} (1 - \varphi_2) & -\varphi_3 \\ -(\varphi_1 + \varphi_3) & 1 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$$

donnent

$$\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \frac{1}{1 - \varphi_2 - \varphi_3(\varphi_1 + \varphi_3)} \begin{pmatrix} 1 & \varphi_3 \\ \varphi_1 + \varphi_3 & 1 - \varphi_2 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \\ = \left\{ \begin{array}{l} \frac{\varphi(1) + \varphi(2)\varphi(3)}{1 - (\varphi(2) + \varphi(3)(\varphi(1) + \varphi(3)))}, \quad \frac{\varphi(1)(\varphi(3) + \varphi(1)) - \varphi(2)^2 + \varphi(2)}{1 - (\varphi(2) + \varphi(3)(\varphi(1) + \varphi(3)))} \end{array} \right\}$$

(e) On trouve que le troisième coefficient d'autocorrélation partielle est

$$\varphi_{3,3} = \frac{\rho(1)^3 - \rho(3)\rho(1)^2 + \rho(2)^2\rho(1) - 2\rho(2)\rho(1) + \rho(3)}{(1 + \rho(2) - 2\rho(1)^2)(1 - \rho(2))},$$

et une formule encore plus compliquée pour $\varphi_{4,4}$. Mais, en supposant un modèle $AR(3)$, ça simplifie à

$$\varphi_{3,3} = \varphi_3, \varphi_{4,4} = 0$$

3. (a) Résolvant les récurrences donne

(b) Non

(c) $X_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t-k}$, $Z_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t+k}$, $t = 0, 1, 2, \dots$ sont des processus stationnaires.

11.2 Examen 2

1. Est-ce que le processus $AR(2)$:

$$Y_t = -\frac{1}{2}Y_{t-1} + \frac{7}{50}Y_{t-2} + \epsilon_t, \quad \epsilon_t \text{ bruit blanc} (\sigma^2 = 1)$$

est stationnaire causal ? Si oui, calculer ses corrélations.

2. a) Calculer le développement limité autour de $x = 0$ de $\frac{1}{(1-x)^3}$, en différenciant $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$. b) Est-ce que le processus Y_t défini par :

$$\left(1 - \frac{3}{2}B + \frac{3B^2}{4} - \frac{B^3}{8}\right)Y_t = \epsilon_t$$

est stationnaire causal ? Si oui, donner une formule générale pour les coefficients ψ_k de sa représentation comme processus $MA(\infty)$.

3. On cherche un filtre causal $\theta(B) = \sum_{i=0}^n \theta_i B^i$ d'ordre n qui laisse passer une tendance linéaire sans distortion et qui minimise la variance de moyennes mobiles

$$X_t = \sum_{i=0}^n \theta_i \epsilon_{t-i}, \quad \epsilon_t \text{ bruit blanc } (\sigma^2 = 1).$$

Formuler ce problème comme un problème d'optimization avec des contraintes, qu'on résoudra par la méthode des multiplicateurs de Lagrange. **Ind** : On rappelle que $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$.

4. Soit Y_t, Y_{t-1}, \dots, Y_0 une suite d'observations à allure stationnaire, qu'on aimerait modéliser comme $AR(p)$, avec p à déterminer. a) On considère d'abord une modélisation :

$$Y_t = \varphi Y_{t-1} + \epsilon_t, \quad t = 1, 2, \dots, n \quad (11.3)$$

où on propose de prendre pour φ la valeur $\hat{\varphi}(Y_0, \dots, Y_n)$ qui minimise la somme des erreurs au carré $SC(\varphi) = \sum_{i=1}^n \epsilon_t^2$. Quelle est la valeur $\hat{\varphi}(Y_0, \dots, Y_n)$? b) Supposons maintenant que Y_t sont en effet des observations d'un processus stationnaire $AR(1)$ satisfaisant (11.3), où ϵ_t est un bruit blanc (centré, non-corrélé) de variance $\sigma^2 = 1$. Quelle est la valeur de φ qui minimise l'objectif $E[\epsilon_t^2]$? Comment estimer cette valeur à partir des données observées ? Comparer avec la réponse précédente. c) On suppose en suite que Y_t sont des observations d'un processus stationnaire $AR(2)$,

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \epsilon_t, \quad t = 2, 3, \dots, n$$

où ϵ_t est un bruit blanc de variance $\sigma^2 = 1$. Trouver et résoudre le système d'équations satisfaites par les valeurs φ_1, φ_2 qui minimisent l'objectif $E[\epsilon_t^2]$. d) On accepte finalement un modèle $AR(3)$ avec bruit

blanc ϵ_t ($\sigma^2 = 1$) (avec coefficients $\varphi_1, \varphi_2, \varphi_3$ déterminés par les données). Donner les équations liant les autocorrélations ρ_1, ρ_2, ρ_3 d'un processus $AR(3)$ aux coefficients $\varphi_1, \varphi_2, \varphi_3$ (on ne demande pas de justifier). Calculer les 2 premières autocorrélations $\rho_i, i = 1, 2$ du processus Y_t , en fonction des coefficients $\varphi_1, \varphi_2, \varphi_3$.

5. Soit $(\varepsilon_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{N}(0, \sigma^2)$ où $\sigma^2 > 0$. Soit $a \in]0, 1[$ et $(X_n)_{n \geq 0}$ une suite de variables aléatoires telle que $X_0 = x_0$ presque sûrement et $X_n = aX_{n-1} + \varepsilon_n$ pour $n \geq 1$.

(a) Montrer que pour $n \geq 1$

$$X_n = a^n x_0 + \sum_{k=0}^{n-1} a^k \varepsilon_{n-k}.$$

Solution : Pour $n = 1$ on a $X_1 = ax_0 + \varepsilon_1 = aX_0 + \varepsilon_1$. Supposons la relation vraie au rang $n \geq 1$ alors $X_{n+1} = aX_n + \varepsilon_{n+1} = a(a^n x_0 + \sum_{k=0}^{n-1} a^k \varepsilon_{n-k}) + \varepsilon_{n+1} = a^{n+1} x_0 + \sum_{k=0}^{n-1} a^{k+1} \varepsilon_{n+1-(k+1)} + \varepsilon_{n+1} = a^{n+1} x_0 + \sum_{k'=1}^n a^{k'} \varepsilon_{n+1-k'} + \varepsilon_{n+1} = a^{n+1} x_0 + \sum_{k'=0}^n a^{k'} \varepsilon_{n+1-k'}$. Donc relation établie par récurrence.

- (b) Déterminer la loi de X_n et montrer que $X_n \xrightarrow{\mathcal{L}} X$ où $X \sim \mathcal{N}(0, \sigma^2/(1-a^2))$.

Solution : X_n étant une combinaison linéaire de variables aléatoires normales indépendantes, elle est normale de moyenne et variance : $\mathbb{E}(X_n) = a^n x_0$ et $\text{var}(X_n) = \sigma^2(1 + a^2 + \dots + a^{2(n-1)}) = \sigma^2(1 - a^{2n})/(1 - a^2)$. Donc si Φ est la fdr d'une loi $\mathcal{N}(0, 1)$ on a pour tout $x \in \mathbb{R}$

$$P(X_n \leq x) = \Phi\left(\frac{x - a^n x_0}{\sigma \sqrt{(1 - a^{2n})/(1 - a^2)}}\right) \rightarrow \Phi\left(\frac{x \sqrt{1 - a^2}}{\sigma}\right)$$

qui est la fdr de la loi $\mathcal{N}(0, \sigma^2/(1 - a^2))$.

- (c) On suppose maintenant que $X_n = aX_{n-1} + b + \varepsilon_n$ pour $n \geq 1$ où $b \in \mathbb{R}$. Montrer que $(X_n)_{n \geq 0}$ converge en loi et déterminer sa loi asymptotique.

Solution : On établit par récurrence que

$$X_n = a^n x_0 + b \left(\frac{1 - a^n}{1 - a} \right) + \sum_{k=0}^{n-1} a^k \varepsilon_{n-k}$$

X_n étant une combinaison linéaire de variables aléatoires normales indépendantes, elle est normale de moyenne et variance : $\mathbb{E}(X_n) = a^n x_0 + b(1 - a^n)/(1 - a)$ et $\text{var}(X_n) = \sigma^2(1 + a^2 + \dots + a^{2(n-1)}) = \sigma^2(1 - a^{2n})/(1 - a^2)$. Donc si Φ est la fdr d'une loi $\mathcal{N}(0, 1)$ on a pour tout $x \in \mathbb{R}$

$$P(X_n \leq x) = \Phi\left(\frac{x - a^n x_0 - b(1 - a^n)/(1 - a)}{\sigma \sqrt{(1 - a^{2n})/(1 - a^2)}}\right) \rightarrow \Phi\left(\frac{(x - b/(1 - a)) \sqrt{1 - a^2}}{\sigma}\right)$$

qui est la fdr de la loi $\mathcal{N}(b/(1 - a), \sigma^2/(1 - a^2))$.

Solutions :

- $\rho_k = \rho_1 \psi_{k-1} + \varphi_2 \psi_{k-2}, k \geq 1$
- $\text{Var } X_t = \sum_i \theta_i^2$. Elle doit être minimisé sous les contraintes $\sum_i \theta_i = 1, \sum_i i \theta_i = 0$. On trouve $\theta_i = \frac{\lambda + i\mu}{2}$ et les multiplicateurs de Lagrange satisfont

$$\begin{aligned} \lambda + \frac{n\mu}{2} &= \frac{2}{n+1} \\ \lambda + \frac{(2n+1)\mu}{3} &= 0 \end{aligned}$$

3. (a) $\hat{\varphi}(Y_0, \dots, Y_n) = \frac{\sum_{t=1}^n Y_t Y_{t-1}}{\sum_{t=0}^{n-1} Y_t^2}$

(b) $\hat{\varphi} = \frac{\gamma_1}{\gamma_0}$

(c) Pour $p = 2$, on arrive au système YW :

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2$$

avec solution

$$\varphi_{2,1} = \frac{\begin{vmatrix} \rho_1 & \rho_1 \\ \rho_2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2} \quad \varphi_{2,2} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

On s'aperçoit en résolvant cet exercice que les $\varphi_i, i = 1$, sont exactement les coefficients de la régression de Y_t sur Y_{t-1}, Y_{t-2} (d'où le nom autoregression), et que pour $p = 1$, y a coïncidence entre les valeurs du coefficient de régression φ et de la corrélation ρ_1 , mais pour $p \geq 2$, la situation se complique.

(d) Pour $p = 3$, le système YW est :

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1 + \varphi_3 \rho_2$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2 + \varphi_3 \rho_1$$

$$\rho_3 = \varphi_1 \rho_2 + \varphi_2 \rho_1 + \varphi_3$$

Les deux premières équations

$$\begin{pmatrix} (1 - \varphi_2) & -\varphi_3 \\ -(\varphi_1 + \varphi_3) & 1 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$$

donnent

$$\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \frac{1}{1 - \varphi_2 - \varphi_3(\varphi_1 + \varphi_3)} \begin{pmatrix} 1 & \varphi_3 \\ \varphi_1 + \varphi_3 & 1 - \varphi_2 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \\ = \left\{ \frac{\varphi(1) + \varphi(2)\varphi(3)}{1 - (\varphi(2) + \varphi(3)(\varphi(1) + \varphi(3)))}, \frac{\varphi(1)(\varphi(3) + \varphi(1)) - \varphi(2)^2 + \varphi(2)}{1 - (\varphi(2) + \varphi(3)(\varphi(1) + \varphi(3)))} \right\}$$

(e) Les coefficients d'autocorrélation partielle $\varphi_{k,k}, \forall k \geq 1$ sont

$$\varphi_{1,1} = \rho_1 = \frac{\varphi(1) + \varphi(2)\varphi(3)}{1 - (\varphi(2) + \varphi(3)(\varphi(1) + \varphi(3)))}, \\ \varphi_{2,2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\varphi(2) + \varphi(1)\varphi(3)}{1 - \varphi(3)^2}, \varphi_{3,3} = \varphi_3, \varphi_{4,4} = 0$$

Examen series temporelles 2014

1. Trouvez un filtre causal $\theta(B) = \sum_{i=0}^3 \theta_i B^i$ qui conserve les polynômes de degré ≤ 1 , qui enlève les composantes saisonnières d'ordre 2, et qui minimise la variance quand appliqué à un BB.
2. Soit $X_t = \sum_{i=0}^q \theta_i \epsilon_{t-i}$ un processus moyenne mobile, avec ϵ_t bruit blanc de loi $N(0, \sigma^2 = 1)$.
 - (a) Quelle est la variance de X_t ?
 - (b) Trouver la covariance $\gamma_1 = E[X_t X_{t-1}]$.
 - (c) Trouver les covariances $\gamma_k = E[X_t X_{t-k}]$, $k = 2, 3, \dots$
3. On considère des processus X_t, Y_t, Z_t satisfaisant respectivement :

$$(1 - B/2)X_t = \epsilon_t, \quad (1 - B)Y_t = \epsilon_t, \quad (1 - 2B)Z_t = \epsilon_t, \quad t = T, T-1, \dots, 1, \quad (11.4)$$

où ϵ_t est un bruit blanc de loi $N(0, \sigma^2 = 1)$.

- (a) Donner les expressions de X_T , de Y_T , et de Z_T en fonction de $\epsilon_t, t = 1, 2, \dots, T$, et des valeurs au temps $t = 0$ (denotées par X_0, Y_0, Z_0).
 - (b) Est-ce que l'une des équations (11.4) admet une solution stationnaire ? Le cas échéant, donner une telle solution au temps T (pour X_T, Y_T , ou Z_T), en fonction de $\epsilon_t, t = T, T-1, \dots, 1, 0, -1, -2, \dots$. Calculer les variances des expressions obtenues et vérifier qu'elles ne dépendent pas de T .
4. Soit Y_t un processus AR(2) :

$$Y_t = -\frac{1}{2}Y_{t-1} + \frac{7}{50}Y_{t-2} + \epsilon_t, \quad \epsilon_t \text{ bruit blanc } (\sigma^2 = 1)$$

- (a) Donner la fonction de transfert $\psi(B)$ exprimant Y_t comme $Y_t = \psi(B)\epsilon_t$, en forme factorisée.
- (b) Donner une représentation moyenne mobile

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

en spécifiant les coefficients ψ_i .

Est-ce que le processus Y_t est stationnaire causal ?

- (c) Si oui, calculer la corrélation ρ_1 pour ce processus stationnaire.
- (d) Donner une récurrence pour $\rho_k, k \geq 2$, et trouver ρ_2 .
- (e) Résoudre la récurrence pour ρ_k .

5. Soit $(\varepsilon_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{N}(0, 1)$. Soit

$$\varphi(B) = I - \varphi_1 B - \varphi_2 B^2 = (I - \lambda_1 B)(I - \lambda_2 B), \lambda_1 \neq \lambda_2,$$

où B est l'opérateur de rétro-décalage et I est l'opérateur identité. Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires telles que

$$\varphi(B)X_n = \varepsilon_n, \forall n \geq 1, \quad X_0 = x_0, X_{-1} = x_{-1} \quad (11.5)$$

- (a) Trouver deux solutions indépendantes de l'équation homogène $\varphi(B)h_n = 0$.
- (b) Trouver la solution X_n de (11.5), en admettant que si ψ_k satisfait la récurrence YW

$$\psi_k = \varphi_1 \psi_{k-1} + \varphi_2 \psi_{k-2}, \psi_1 = \varphi_1, \psi_0 = 1, \psi_{-1} = 0, \dots$$

alors

$$p_n = \sum_{k=0}^{n-1} \psi_k \varepsilon_{n-k}, \forall n \geq 1$$

est une solution particulière de l'équation inhomogène $\varphi(B)p_n = \varepsilon_n$, et $p_n = 0, \forall n \leq 0$.

- (c) Déterminer la loi de X_n , et sa limite quand $n \rightarrow \infty$, en supposant $\lambda_1, \lambda_2 \in]-1, 1[$.
- (d) Choisir une loi jointe pour X_{-1}, X_0 sous la quelle le processus X_n peut être stationnaire, en supposant que (X_{-1}, X_0) sont indépendants de $\varepsilon_1, \varepsilon_2, \dots$

Solutions :

1.

2.

3. (a) *Resolvant les recurrences donne ...*

(b) $X_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t-k}$, $Z_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t+k}$, $t = 0, 1, 2, \dots$ sont des processus stationnaires.

4. (a)

$$Y_t + \frac{1}{2}Y_{t-1} - \frac{7}{50}Y_{t-2} = (1 + \frac{1}{2}B - \frac{7}{50}B^2)Y_t = \epsilon_t$$

$$Y_t = \frac{1}{(1 + \frac{7}{10}B)(1 - \frac{1}{5}B)} \epsilon_t$$

(b) $\psi_k = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2}$, $\lambda_1 = \frac{1}{5}$, $\lambda_2 = \frac{-7}{10}$.

(c) $\rho_1 = \frac{\phi_1}{1 - \phi_2} = \frac{-1/2}{1 - \frac{7}{50}} = -\frac{25}{43}$.

(d) $\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} = -\frac{1}{2} \rho_{k-1} + \frac{7}{50} \rho_{k-2}$, $k \geq 1$, $\rho_k = \frac{1}{129} \left(17 \left(\frac{1}{5}\right)^k + 112 \left(\frac{-7}{10}\right)^k \right)$.

5. (a) *Resolvant la recurrence homogène donne $h_i(n) = \lambda_i^n$, $i = 1, 2$.*

(b) $X_n = K_1 \lambda_1^n + K_2 \lambda_2^n + \sum_{k=0}^{n-1} \psi_k \epsilon_{n-k}$ doit satisfaire

$$\begin{cases} X_0 = K_1 + K_2 \\ X_{-1} = K_1/\lambda_1 + K_2/\lambda_2 \end{cases} \implies K_1 = \frac{X_0 \lambda_1 - X_{-1} \lambda_1 \lambda_2}{\lambda_1 - \lambda_2}, K_2 = \frac{X_{-1} \lambda_1 \lambda_2 - X_0 \lambda_2}{\lambda_1 - \lambda_2}$$

(c) $\lambda_1, \lambda_2 \in]-1, 1[\implies \lim_{n \rightarrow \infty} X_n = p_\infty = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}$.

$$\sigma^2 = \text{Var}(X_n) = \sum_{k=0}^{\infty} \psi_k^2 = \frac{1}{1 - \varphi_1 \rho_1 - \varphi_2 \rho_2}.$$

La loi limite est $N(0, \sigma^2)$.

Examen series temporelles 2014

1. Soit $X_t = \sum_{i=0}^q \theta_i \epsilon_{t-i}$ un processus moyenne mobile, avec ϵ_t bruit blanc de loi $N(0, \sigma^2 = 1)$.

(a) Quelle est la variance de X_t ?

(b) Trouver la covariance $\gamma_1 = E[X_t X_{t-1}]$.

(c) Trouver les covariances $\gamma_k = E[X_t X_{t-k}]$, $k = 2, 3, \dots$

2. a) Trouvez un filtre causal $\theta(B) = \sum_{i=0}^q \theta_i B^i$ de degré minimal q , qui conserve les polynômes de degré ≤ 1 , et qui enlève les composantes saisonnières d'ordre 2.

b) Trouvez un filtre causal $\theta(B) = \sum_{i=0}^3 \theta_i B^i$ qui conserve les polynômes de degré ≤ 1 et qui minimise la variance quand appliqué à un BB.

3. Soit Y_t un processus AR(2) :

$$Y_t = -\frac{1}{2}Y_{t-1} + \frac{7}{50}Y_{t-2} + \epsilon_t, \delta\epsilon_t \text{ bruit blanc } (\sigma^2 = 1)$$

(a) Donner la fonction de transfert $\psi(B)$ exprimant Y_t comme $Y_t = \psi(B)\epsilon_t$, en forme factorisée.

(b) Donner une représentation moyenne mobile

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

en spécifiant les coefficients ψ_i .

Est-ce que le processus Y_t est stationnaire causal ?

(c) Si oui, calculer la corrélation ρ_1 pour ce processus stationnaire.

(d) Donner une récurrence pour ρ_k , $k \geq 2$, et trouver ρ_2 .

(e) Résoudre la récurrence pour ρ_k .

Solutions :

1.

2.

3. (a) *Resolvant les recurrences donne ...*

(b) $X_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t-k}$, $Z_t = \sum_{k=0}^{\infty} 2^{-k} \epsilon_{t+k}$, $t = 0, 1, 2, \dots$ sont des processus stationnaires.

4. (a)

$$Y_t + \frac{1}{2}Y_{t-1} - \frac{7}{50}Y_{t-2} = (1 + \frac{1}{2}B - \frac{7}{50}B^2)Y_t = \epsilon_t$$

$$Y_t = \frac{1}{(1 + \frac{7}{10}B)(1 - \frac{1}{5}B)} \epsilon_t$$

(b) $\psi_k = \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2}$, $\lambda_1 = \frac{1}{5}$, $\lambda_2 = \frac{-7}{10}$.

(c) $\rho_1 = \frac{\phi_1}{1 - \phi_2} = \frac{-1/2}{1 - \frac{7}{50}} = -\frac{25}{43}$.

(d) $\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} = -\frac{1}{2} \rho_{k-1} + \frac{7}{50} \rho_{k-2}$, $k \geq 1$, $\rho_k = \frac{1}{129} \left(17 \left(\frac{1}{5}\right)^k + 112 \left(\frac{-7}{10}\right)^k \right)$.

5. (a) *Resolvant la recurrence homogène donne $h_i(n) = \lambda_i^n$, $i = 1, 2$.*

(b) $X_n = K_1 \lambda_1^n + K_2 \lambda_2^n + \sum_{k=0}^{n-1} \psi_k \epsilon_{n-k}$ doit satisfaire

$$\begin{cases} X_0 = K_1 + K_2 \\ X_{-1} = K_1/\lambda_1 + K_2/\lambda_2 \end{cases} \implies K_1 = \frac{X_0 \lambda_1 - X_{-1} \lambda_1 \lambda_2}{\lambda_1 - \lambda_2}, K_2 = \frac{X_{-1} \lambda_1 \lambda_2 - X_0 \lambda_2}{\lambda_1 - \lambda_2}$$

(c) $\lambda_1, \lambda_2 \in]-1, 1[\implies \lim_{n \rightarrow \infty} X_n = p_{\infty} = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}$.

$$\sigma^2 = \text{Var}(X_n) = \sum_{k=0}^{\infty} \psi_k^2 = \frac{1}{1 - \varphi_1 \rho_1 - \varphi_2 \rho_2}.$$

La loi limite est $N(0, \sigma^2)$.