



### *Remerciements*

En premier lieu, je tiens à remercier mon **Dieu** qui m'a donné le courage et la volonté pour réaliser ce travail.

Je tiens à exprimer ma reconnaissance à ma directrice de mémoire *M<sup>me</sup>* **Fatiha Mokhtari**, pour la confiance qu'elle m'a accordé.

Je la remercie profondément pour sa disponibilité, ses orientations et ses remarques fructueuses. Qu'elle trouve ici l'expression de ma profonde gratitude.

Mes remerciements s'adressent également au président et aux membres du jury qui ont accepté de juger ce travail.

Je suis très reconnaissante aux responsables du laboratoire LMSSA, pour nous avoir donné la chance de travailler dans un environnement scientifique

Je tiens à remercier à titre individuel mes enseignants : *M<sup>r</sup>* **T.Guendouzi**, *M<sup>r</sup>* **F.Madani**, *M<sup>er</sup>* **M.Kadi**, *M<sup>r</sup>* **K.Djerfi**, *M<sup>r</sup>* **A.Kandouci**, *M<sup>me</sup>* **F.Benziadi**, *M<sup>elle</sup>* **L.Bousmaha**, et *M<sup>elle</sup>* **F.Benziadi**, *M<sup>elle</sup>* **S.Rahmani**, *M<sup>elle</sup>* **S.Idriss**, *M<sup>me</sup>* **N.Ait**, *M<sup>me</sup>* **R.Rouane**.

Merci à tous mes collègues étudiantes et étudiants du master ASSPA. Enfin, un grand merci à tous les personnes qui m'ont aidé lors de la réalisation de ce modeste travail.

« *Merci à tous* »

## **Résumé**

Lorsque l'on veut étudier les interdépendances entre plusieurs variables, il est évidemment nécessaire de se situer dans un cadre multivarié. On introduit dans ce travail, la notion de processus autorégressifs multivariés vectoriels, qui est un outil très important dans la modélisation statistique de données réelles. Un intérêt particulier est porté sur l'estimation et la dynamique de ces processus. On termine ce travail par des exemples d'applications sur des données macroéconomiques.

## **Title : VAR modeling**

## **Abstract**

When one wants to study the interdependence between several variables, it is obviously necessary to be in a multivariate frame. An introduced in this work, The concept of multivariate vector autoregressive processes, which a very important tool in the statistical modeling of real data. In particular interest is focused on the estimation and dynamics of these processes. We conclude this work with examples of application on macroeconomic data.

**Mots clés :** Processus stationnaire, processus autorégressifs réels, processus autorégressifs vectoriels, estimation, prévision, dynamique d'un processus.

**Key words :** stationary Process, real autoregressive process, vector autoregressive process, estimation, forecasting.

# Table des matières

<b>Introduction</b>	<b>7</b>
<b>1 Processus autoregressif réel</b>	<b>10</b>
1.1 Processus réels stationnaires du second ordre . . . . .	10
1.1.1 Définitions . . . . .	10
1.1.2 Autocovariance et autocorrélation . . . . .	12
1.1.3 Théorème De Wold . . . . .	15
1.1.4 Opérateurs avance et retard . . . . .	15
1.2 Processus autorégressif d'ordre $p$ ( $AR(p)$ ) . . . . .	17
1.2.1 Définition . . . . .	17
1.2.2 Ecriture $MA(\infty)$ . . . . .	17
1.2.3 Estimation dans le modèle autorégressif d'ordre $p$ ( $AR(p)$ )	20
1.2.4 Prédiction dans un modèle autorégressif $AR(p)$ . . . . .	22
<b>2 Processus autorégressif vectoriel</b>	<b>24</b>
2.1 Préliminaires . . . . .	24
2.1.1 Série temporelle multivariée . . . . .	24
2.1.2 Processus stationnaire vectoriel . . . . .	25
2.1.3 Innovation d'un processus vectoriel . . . . .	27
2.1.4 Convergence des moments empiriques . . . . .	30
2.2 Motivation . . . . .	31
2.3 Représentation générale . . . . .	32
2.3.1 Processus $VAR(1)$ . . . . .	35
2.3.2 Représentation $VAR(1)$ d'un $VAR(p)$ . . . . .	36
2.4 Fonction d'autocovariance d'un processus $VAR(p)$ . . . . .	40
2.5 Représentation $MA(\infty)$ . . . . .	41

---

2.6	Estimation d'un VAR . . . . .	45
2.6.1	Estimation par le maximum de vraisemblance . . . . .	46
2.7	Choix des retards . . . . .	46
2.8	Prévision dans un VAR stationnaire . . . . .	47
2.8.1	Cas d'un VAR(1) . . . . .	47
2.8.2	Cas d'un VAR(p) . . . . .	48
2.9	Causalité . . . . .	52
2.9.1	Causalité au sens de Granger . . . . .	52
2.9.2	Causalité au sens de Sims . . . . .	54
2.10	Dynamique d'un modèle VAR . . . . .	54
2.10.1	Analyse impulsionnelle . . . . .	54
2.10.2	Décomposition de la variance . . . . .	60
2.11	La cointégration et le modèle à correction d'erreur . . . . .	61
2.11.1	Le concept de cointégration . . . . .	61
<b>3</b>	<b>Simulations et Applications</b>	<b>64</b>
3.1	Simulation d'un modèle AR . . . . .	64
3.1.1	Simulation d'un processus AR(1) . . . . .	64
3.1.2	Simulation d'un processus AR(2) . . . . .	65
3.2	Un modèle AR pour lh . . . . .	68
3.2.1	Choix d'un modèle . . . . .	68
3.2.2	Prévision . . . . .	69
3.3	Simulation d'un modèle VAR . . . . .	70
3.3.1	Spécification, hypothèses et estimation . . . . .	70
3.3.2	Tests de diagnostic . . . . .	73
3.3.3	Prévisions du processus VAR . . . . .	75
3.3.4	Analyse de causalité . . . . .	77
3.3.5	Fonctions de réponse impulsionnelle (analyse du choc) . . . . .	78
3.3.6	Décomposition de la variance des erreurs de la prévision (FEVD) . . . . .	80
3.4	Applications sur des données réelles . . . . .	81
3.4.1	Méthode d'estimation : . . . . .	81
3.4.2	Définitions . . . . .	81
3.4.3	Application aux données macroéconomiques canadiennes . . . . .	82

**TABLE DES MATIÈRES** **6**

---

3.4.4	Application aux données macroéconomiques Algériennes	95
3.5	Conclusion . . . . .	108
<b>Bibliographie</b>		<b>108</b>

# Introduction

« Une des tâches principales du macroéconomiste est d'expliquer comment les agrégats macroéconomiques se comportent au fil du temps », c'est-à-dire comment ces variables sont affectées par « la politique économique et les modifications de l'environnement économique »

Comment le PIB et l'inflation sont affectés par une augmentation temporaire des taux d'intérêt ou une baisse d'impôts ? Que se passe-t-il si une banque centrale modifie de façon permanente son objectif d'inflation ou si un gouvernement modifie son objectif d'équilibre budgétaire ? Tel est le type de questions auxquelles les lauréats ont apporté une réponse.

Le 10 octobre 2011, l'Académie Royale des Sciences de l'Suède a attribué le prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel à Thomas Sargent et Christopher Sims « pour leurs recherches empiriques sur la cause et l'effet en macroéconomie ». Selon les termes employés par l'Académie Royale suédoise.

Les recherches de Christopher Sims ont porté plus particulièrement sur la façon dont les chocs économiques ( tels que la flambée des prix du pétrole, ou la forte baisse de la consommation des ménages ) agissent sur l'économie

Il propose alors une modélisation multivariée dont les seules restrictions sont le choix des variables sélectionnées et le nombre de retards intégrés.

La représentation VAR constitue une alternative à ce type de modèles. Cette nouvelle représentation repose toutefois sur l'hypothèse que l'évolution de l'économie peut être bien approchée par la description du comportement dynamique d'un vecteur de  $N$  variables dépendant linéairement du passé.

Une première qualité des modèles VAR est qu'ils permettent d'appré-

hender le comportement dynamique de variables linéairement dépendantes du passé. De plus en tenant compte des relations entre plusieurs variables, les modèles VAR permettent ainsi d'expliquer, et non plus uniquement de décrire l'évolution d'une série.

La seconde qualité des VAR est qu'ils évitent d'avoir à décider quelles sont les variables exogènes et endogènes du modèle. Ils comportent seulement des variables endogènes. Le troisième avantage est que les VAR sont des modèles empiriques. Ainsi, seules les données décident des relations possibles entre les variables. L'intervention du chercheur pour poser des restrictions ou des hypothèses est quasi absente (Sims, 1980). Les seules et uniques restrictions dans un VAR sont le choix de variables à inclure dans le modèle et le nombre de retards.

Dans le premier chapitre de ce mémoire, nous présentons un rappel sur les processus du second ordre, ensuite nous nous intéressons aux processus autorégressifs réel d'ordre  $p$ , noté  $AR(p)$ , nous présentons quelques solutions des problèmes liés à l'estimation et la prévisions de ces processus.

Le deuxième chapitre, est consacré aux processus autorégressifs vectoriels (VAR) dont les utilisations sont fréquentes pour modéliser des séries temporelles multivariées. Après avoir donné la représentation générale, on s'intéresse aux problèmes d'estimation et de prévision liés a ce genre de processus. Ce chapitre ce termine par donner trois concepts importants dans l'analyse VAR : la causalité, la fonction d'implusion, et la décomposition de la variance des erreurs de la prévision (FEVD).

Dans le troisième chapitre nous présentons des simulations. Pour cela, nous avons utilisé le langage R version 3.4.4. et plus particulièrement les packages vars et dse. Ces packages comprennent des fonction qui servent à simuler un modèle VAR, estimer ses paramètres, prédire et encore plus effectuer des tests de diagnostique de causalité et une analyse implusionnelle .

Ces fonctions sont expliquées en employant un ensemble de données simulées et réelles telle que les données Canadiennes et les données Algeriennes



du (PIB, ICP, taux de chaumage, taux d'intérêt réel,... ).

# Chapitre 1

## Processus autoregressif réel

### 1.1 Processus réels stationnaires du second ordre

#### 1.1.1 Définitions

Dans toute la suite on considérera  $(X_t)_{t \in \mathbb{Z}}$  et on supposera  $X_t \in \mathcal{L}^2(\Omega, \mathbf{A}, \mathbb{P})$ ,  $\forall t \in \mathbb{Z}$ .  $\mathcal{L}^2(\Omega, \mathbf{A}, \mathbb{P})$  est l'espace des variables aléatoire de carré intégrable.

**Définition 1.1.1. (*Processus stochastiques*)** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. Un processus stochastique  $X$  est une famille de variables aléatoires réelles  $(X_t)_{t \in \Theta}$ , où  $\Theta \subset \mathbb{R}$  est appelé l'espace des temps. Si  $\Theta \subset \mathbb{Z}$ , le processus est dit à temps discret. Si  $\Theta$  est un intervalle de  $\mathbb{R}$ , le processus est dit à temps continu.

**Définition 1.1.2. (*Stationnarité stricte ou forte*)**  $(X_t)$  est un **processus stationnaire au sens strict** si :  $\forall n \in \mathbb{N}$ ,  $\forall (t_1, \dots, t_n)$ ,  $\forall h \in \mathbb{Z}$ , la loi de  $(X_{t_1}, \dots, X_{t_n})$  est identique à la loi de  $(X_{t_1+h}, \dots, X_{t_n+h})$ .

**Définition 1.1.3. (*Stationnarité faible*)**  $(X_t)_{t \in \mathbb{Z}}$  est un **processus stationnaire du second ordre** (ou un processus faiblement stationnaire) s'il vérifie :

(i)  $\forall t \in \mathbb{Z}$ ,  $\mathbb{E}(X_t) = m$

(ii)  $\forall t \in \mathbb{Z}$ ,  $\mathbb{V}(X_t) = \sigma^2$

(iii)  $\forall t \in \mathbb{Z}$ ,  $\forall h \in \mathbb{Z}$ ,  $\mathbb{Cov}(X_t, X_{t+h})$  (ne dépend que de  $h$ )

**Remarque 1.1.1.** (1) Dans la suite, les processus stationnaires désignent les processus de la définition (1.1.3).

(2) (iii)  $\Rightarrow$  (ii) :  $h = 0$

(3) Si un processus est stationnaire au sens strict alors il est faiblement stationnaire ;

(4) Si  $(X_t)_{t \in \mathbb{Z}}$  est un processus gaussien alors il y a équivalence entre stationnarité faible et forte.

**Exemple 1. (Processus stationnaire)**

(1) **Bruit blanc faible**,  $(\varepsilon_t)_{t \in \mathbb{Z}}$ , si et seulement si :

$$\begin{aligned}\mathbb{E}(\varepsilon_t) &= 0, \forall t \in \mathbb{Z} \\ \mathbb{V}(\varepsilon_t) &= \sigma^2, \forall t \in \mathbb{Z} \\ \text{Cov}(\varepsilon_t, \varepsilon_s) &= 0, \text{ si } t \neq s\end{aligned}$$

On notera  $\varepsilon_t \sim \mathcal{BB}(0, \sigma^2)$

(2)  $\varepsilon_t$  est un **bruit blanc fort** si et seulement si les  $\varepsilon_t$  sont i.i.d,  $\mathbb{E}(\varepsilon_t) = 0$  et  $\mathbb{V}(\varepsilon_t) = \sigma^2$ .

(3) processus **moyenne mobile** d'ordre 2 noté MA(2). Soit  $\theta \in \mathbb{R}^*$  et soit  $\varepsilon_t \sim \mathcal{BB}(0, \sigma^2)$ .

Soit  $(X_t)_{t \in \mathbb{Z}}$  défini par :  $\forall t \in \mathbb{Z}, X_t = \varepsilon_t - \theta\varepsilon_{t-1} - \theta\varepsilon_{t-2}$ .

Alors  $(X_t)_{t \in \mathbb{Z}}$  est un **processus stationnaire**, on dit que  $X_t \sim \text{MA}(2)$ .

**Exemple 2. (Processus non stationnaire)** (1) **Marche aléatoire** : Soit  $\varepsilon_t \sim \mathcal{BB}(0, \sigma^2)$ ,  $(X_t)_{t \in \mathbb{Z}}$  est une marche aléatoire sans dérive si et seulement si :

(i)  $X_t = X_{t-1} + \varepsilon_t$

(ii)  $\text{Cov}(\varepsilon_t, X_{t-k}) = 0, \forall k > 0$ .

Même si on a la propriété  $\mathbb{E}(X_t) = \mathbb{E}(X_{t-1}) \Rightarrow \mathbb{E}(X_t) = m, \forall t \in \mathbb{Z}$ ,  $X_t$  n'est pas stationnaire :

$$\left. \begin{aligned} X_t &= X_{t-1} + \varepsilon_t \\ X_{t-1} &= X_{t-2} + \varepsilon_{t-1} \\ &\vdots \\ X_1 &= X_0 + \varepsilon_1 \end{aligned} \right\} \Rightarrow X_t = X_0 + \sum_{k=1}^t \varepsilon_k$$

D'où

$$\mathbb{V}(X_t) = \mathbb{V}(X_0) + 2 \sum_{k=1}^t \text{Cov}(\varepsilon_k, X_0) + \mathbb{V}\left(\sum_{k=1}^t \varepsilon_k\right) = \mathbb{V}(X_0) + t\sigma^2$$

Le processus n'est pas stationnaire en variance.

## 1.1.2 Autocovariance et autocorrélation

L'une des principales motivations de la modélisation chronologique d'un événement aléatoire est sa structure de corrélation temporelle, en d'autres termes son niveau d'autocorrélation. Il existe, entre autres, deux outils permettant d'évaluer l'autocorrélation d'une série chronologique. Ces derniers sont pour nous d'une importance capitale puisque nous verrons dans la section suivante qu'ils possèdent des propriétés remarquables dans le cadre du processus autorégressif. Restons pour le moment dans le cadre plus général du processus  $(X_t)$  stationnaire, indexé par  $t$  et défini sur  $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ .

**Définition 1.1.4. (Fonction d'auto-covariance)** La fonction d'auto-covariance d'un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  est définie par :

$$\begin{aligned} \gamma &: \mathbb{Z} \rightarrow \mathbb{R} \\ h &\rightarrow \gamma(h) = \text{Cov}(X_t, X_{t-h}) \end{aligned}$$

**Proposition 1.1.1.** (i)  $\gamma$  est une fonction paire :

$$\gamma(-h) = \gamma(h), \forall h \in \mathbb{Z}$$

(ii)  $\gamma$  est de type positif :  $\forall n \in \mathbb{N}, \forall (t_1, \dots, t_n), \forall (a_1, \dots, a_n) \in \mathbb{R}^n$

$$\sum_{1 \leq i, j \leq n} a_i a_j \gamma(t_i - t_j) > 0$$

(iii)  $\gamma(0) = \text{Var}(X_t)$

(iiii)  $|\gamma(h)| \leq \gamma(0), \forall h \in \mathbb{Z}$

**Définition 1.1.5. (Fonction d'auto-corrélation)** La fonction d'auto-corrélation d'un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  est définie par :

$$\forall h \in \mathbb{Z}, \rho(h) = \frac{\gamma(h)}{\gamma(0)} = \text{Corr}(X_t, X_{t+h})$$

**Proposition 1.1.2.** *La fonction d'autocorrélation d'un processus stationnaire vérifie*

1.  $\forall h \in \mathbb{Z}, \rho(-h) = \rho(h)$  : elle est paire ;
2.  $\rho(0) = 1$  ;
3.  $|\rho(h)| \leq 1, \forall h \in \mathbb{Z}$ .

La fonction  $\rho(h)$  est l'expression du lien linéaire entre  $X_t$  et  $X_{t-h}$ . Si  $t$  est l'instant présent et  $h > 0$ ,  $\rho(h)$  est l'expression du lien linéaire entre le présent et le passé d'ordre  $h$  (mais aussi par parité de la fonction  $\gamma$  entre le présent et le futur d'ordre  $h$ ) ; plus  $|\rho(h)|$  est proche de 1 et plus ce lien est fort.

**Définition 1.1.6. (Propriété)** Soient  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire et  $(a_j)_{j \in \mathbb{Z}}$  une suite de réels tels que  $\sum_j |a_j| < +\infty$ .

Alors  $Y_t = \sum_{j \in \mathbb{Z}} a_j X_{t-j}$  est défini (p.s.) pour tout  $t$ . On a les propriétés suivantes :

- (i)  $Y_t \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P}), \forall t \in \mathbb{Z}$
- (ii)  $(Y_t)_t$  est un processus stationnaire tel que

$$\mathbb{E}Y_t = m_Y = \left( \sum_{j \in \mathbb{Z}} a_j \right) m_X$$

$$\gamma_Y(h) = \sum_{j,k} a_j a_k \gamma_X(h+k-j) = \sum_{j,k} a_j a_k \gamma_X(h+j-k), \forall h \in \mathbb{Z}$$

On dit que  $(Y_t)_{t \in \mathbb{Z}}$  est la transformée de  $(X_t)_{t \in \mathbb{Z}}$  par la moyenne mobile infinie associée aux  $(a_j)_{j \in \mathbb{Z}}$ .

**Définition 1.1.7. (Propriété)**

On appelle **auto-corrélation partielle d'ordre  $p$**

$$\begin{aligned} r(p) &= \text{Corr}(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1})) \\ &= \frac{\text{Corr}(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}), X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}))}{[\text{Var}(X_t - EL(X_t|X_{t-1}, \dots, X_{t-p+1}))\text{Var}(X_{t-p} - EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}))]^{1/2}} \end{aligned}$$

$EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1})$  est la projection orthogonale dans  $L^2(\Omega, \mathcal{A}, \mathbb{P})$  de  $X_t$  sur le sous espace vectoriel engendré par  $(X_{t-1}, \dots, X_{t-p+1})$

On montre que  $r(p) = a_p$  coefficient de  $X_{t-p}$  dans  $EL(X_t|X_{t-1}, \dots, X_{t-p})$ .

$$EL(X_{t-p}|X_{t-1}, \dots, X_{t-p+1}) = \sum_{j=1}^p a_j X_{t-j}$$

**Définition 1.1.8. Auto-corrélogramme partiel**

L'auto-corrélogramme partiel de  $(X_t)_{t \in \mathbb{Z}}$  est le graphe de :

$$\begin{cases} \mathbb{N} & \rightarrow ]-1; 1[ \\ p & \mapsto r(p) \end{cases}$$

**Définition 1.1.9. Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire**

(i) La régression linéaire théorique de  $X_t$  sur  $X_{t-1}, \dots, X_{t-p}, \dots$  est la projection orthogonale dans  $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$  de  $X_t$  sur  $H = \text{Vect}(X_{t-1}, \dots, X_{t-p}, \dots)$ .

(ii) La régression affine théorique de  $X_t$  sur  $1, X_{t-1}, \dots, X_{t-p}, \dots$  est la projection orthogonale dans  $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$  de  $X_t$  sur  $H^* = \text{Vect}(1, X_{t-1}, \dots, X_{t-p}, \dots)$ .

On note aussi  $\overline{\mathcal{L}}(\underline{X_{t-1}})$  l'espace  $\overline{\mathcal{L}}(1, X_{t-1}, \dots, X_{t-k}, \dots)$  et

$$EL(X_t|\underline{X_{t-1}}) = EL(X_t|X_{t-1}, \dots, X_{t-k}, \dots)$$

la régression linéaire (ou affine) sur  $\overline{\mathcal{L}}(X_{t-1})$ .

**Proposition 1.1.3.**

$$EL(X_t|\underline{X_{t-1}}) = \lim_{n \rightarrow +\infty} EL(X_t|X_{t-1}, \dots, X_{t-n})$$

au sens de  $\mathcal{L}^2$

**Théorème 1.1.1. (Admis)** Soient  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire et  $X_t^* = EL(X_t|\underline{X_{t-1}})$  la régression affine de  $X_t$  sur  $\overline{\mathcal{L}}(1, X_{t-1}, \dots, X_{t-k}, \dots)$  et  $\varepsilon_t = X_t - X_t^*$ , alors

- $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc
- $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0, \forall k > 0$

**Preuve :** Voir [5]

**Proposition 1.1.4. Processus des innovations**

Avec les notations du théorème ci-dessus :

- (i)  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est le processus des innovations de  $(X_t)_{t \in \mathbb{Z}}$
- (ii)  $\varepsilon_t$  est l'innovation de  $X_t$
- (iii)  $X_t^*$  est la prévision optimale de  $X_t$  à la date  $t - 1$

**Remarque 1.1.2.**  $\varepsilon_t = X_t - X_t^* = X_t - EL(X_t | X_{t-1})$  donc  $\varepsilon_t \perp 1$  et  $\varepsilon_t \perp X_{t-k}, \forall k > 0$ , ce qui peut aussi s'interpréter comme :

$$\mathbb{E}(\varepsilon_t) = 0$$

$$\forall k > 0, \mathbb{E}(\varepsilon_t X_{t-k}) = \text{Cov}(\varepsilon_t, X_{t-k}) = 0$$

**1.1.3 Théorème De Wold**

**Théorème 1.1.2. (De Wold)** Soient  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire et  $(\varepsilon_t)_{t \in \mathbb{Z}}$  le processus des innovations correspondant.

Alors

$$\exists (a_k)_{k \in \mathbb{Z}} / \sum_{k=0}^{+\infty} |a_k| < +\infty$$

et

$$X_t = m + \sum_{k=0}^{+\infty} a_k \varepsilon_{t-k}$$

**1.1.4 Opérateurs avance et retard**

**Définition 1.1.10.** On appelle opérateur retard l'opérateur  $L$  qui à tout processus  $(X_t)_{t \in \mathbb{Z}}$  associe le processus  $(Y_t)_{t \in \mathbb{Z}}$  défini par

$$\forall t \in \mathbb{Z}, Y_t = LX_t = X_{t-1}$$

**Remarque 1.1.3.** L'opérateur  $L$  est linéaire et inversible. Son inverse  $L^{-1} = F$  est défini par

$$\forall t \in \mathbb{Z}, FX_t = X_{t+1}$$

L'opérateur  $F$  est appelé **opérateur avance**

- Proposition 1.1.5.** (i)  $L^k = \underbrace{L \circ \dots \circ L}_{k \text{ fois}}$  vérifie  $L^k X_t = X_{t-k}$   
(ii)  $F^k = \underbrace{F \circ \dots \circ F}_{k \text{ fois}}$  vérifie  $F^k X_t = X_{t+k}$   
(iii)  $L^0 = Id$  est noté  $L^0 = I$  ; ( $L^0 X_t = X_t$ ).

**Proposition 1.1.6.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus du second ordre.

1. Si  $\mathbb{E}(X_t) = at + b$ , alors  $\mathbb{E}((I - L)X_t) = a$
2. Si  $\mathbb{E}(X_t) = \sum_{i=0}^k a_i t^i$ , alors  $\mathbb{E}((I - L)^k X_t) = c$ , où  $c$  est une cste indépendante de  $t$ .

**Définition 1.1.11.** Soit  $P$  un polynôme  $P(z) = \sum_{k=0}^p a_k z^k$ ,  $a_k \in \mathbb{R}$ , on lui associe le **polynôme retard**  $P(L)$  défini comme suit :

$$P(L) = \sum_{k=0}^p a_k L^k$$

Et

$$P(L)X_t = \left( \sum_{k=0}^p a_k L^k \right) X_t = \sum_{k=0}^p a_k X_{t-k}$$

De façon similaire on obtient le **polynôme avance**  $P(F)$  :

$$P(F) = \sum_{k=0}^p a_k F^k$$

Et

$$P(F)X_t = \left( \sum_{k=0}^p a_k F^k \right) X_t = \sum_{k=0}^p a_k X_{t+k}$$

## Inversion d'un polynôme en L

Soit  $\phi$  un polynômes de degré  $p$  à coefficients réels :

$$\phi(z) = 1 + \varphi_1 z + \dots + \varphi_p z^p$$

$$\phi(L) = 1 + \varphi_1 L + \dots + \varphi_p L^p$$



**Proposition 1.1.7.** Soit  $\phi(z) = 1 + \varphi_1 z + \dots + \varphi_p z^p$

(i)  $\phi$  est inversible si et seulement si ses racines  $z_j, j = 1, \dots, p$  sont de module distinct de 1.

(ii) Si  $|z_j| > 1, \forall j \in [1, p]$ , alors  $\phi(L)$  est inversible et

$$\phi(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k \text{ où } a_0 = 1, a_k \in \mathbb{R} \text{ et } \sum_{k=0}^{+\infty} |a_k| < +\infty$$

## 1.2 Processus autorégressif d'ordre p (AR(p))

### 1.2.1 Définition

On appelle processus autorégressif d'ordre  $p$ , noté  $AR(p)$ , un processus stationnaire  $(X_t)$  vérifiant une relation du type

$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t + \mu$  pour tout  $t \in \mathbb{Z}$ , avec  $\varphi_p \neq 0$  et  $\varepsilon_t \sim \mathcal{BB}(0, \sigma^2)$ .

On note  $\phi(L)X_t = \mu + \varepsilon_t$  où  $\phi(L) = 1 - (\varphi_1 L + \dots + \varphi_p L^p)$

**Exemple 3.**  $X_t \sim AR(2)$  i.e.  $(1 - \rho L + \rho^2 L^2)X_t = \mu + \varepsilon_t$  où  $\varepsilon_t \sim \mathcal{BB}(0, \sigma^2)$ .

**Proposition 1.2.1.** Si  $X_t \sim AR(p)$  tel que  $\phi(L)X_t = \mu + \varepsilon_t$ , alors

$$\mathbb{E}(X_t) = \frac{\mu}{\phi(1)} = \frac{\mu}{1 - (\varphi_1 + \dots + \varphi_p)}$$

**Remarque 1.2.1.** On peut se ramener ensuite à  $\mu = 0$  par centrage car  $\phi(L)(X_t - m) = \varepsilon_t$  où  $m = \mu/\phi(1)$ . On considère donc le cas où  $\phi(L)X_t = \varepsilon_t$  (et  $\mathbb{E}(X_t) = 0$ ).

### 1.2.2 Ecriture MA( $\infty$ )

**Ecriture MA( $\infty$ ) quand les racines de  $\phi$  sont de module strictement supérieur à 1**

On suppose que  $\phi(L)X_t = \mu + \varepsilon_t$  où  $\phi(L) = 1 - (\varphi_1 L + \dots + \varphi_p L^p)$  et aussi que  $|z| \leq 1 \implies \phi(z) \neq 0$ .

On suppose que  $\phi(z) = \prod_{i=1}^p (1 - \lambda_i z)$  où  $|\lambda_i| = \frac{1}{|z_i|} < 1$

Alors  $\phi(L)$  est inversible et  $\phi(L)^{-1} = \sum_{k=0}^{\infty} a_k L^k = A(L)$  tel que

$$\sum_{k \geq 0} |a_k| < \infty \text{ et } a_0 = 1.$$

On en déduit

$$\begin{aligned} X_t &= A(L)\mu + A(L)\varepsilon_t \\ &= A(1)\mu + \left(\sum_{k=0}^{\infty} a_k L^k\right)\varepsilon_t \\ &= m + \sum_{k=0}^{\infty} a_k \varepsilon_{t-k} \end{aligned}$$

$$\text{car } \phi(1)^{-1}\mu = m.$$

### Auto-covariance, auto-corrélations et équations de Yule-Walker

- L'auto-covariance :

$\gamma(h) = \text{Cov}(X_t, X_{t-h}) = \mathbb{E}(X_t X_{t-h})$  pour  $h \geq 0$  (car  $m = 0$ ), et

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t$$

Donc

$$\begin{aligned} X_t^2 &= \varphi_1 X_t X_{t-1} + \dots + \varphi_p X_t X_{t-p} + X_t \varepsilon_t \\ \gamma(0) &= \varphi_1 \gamma(1) + \dots + \varphi_p \gamma(p) + \mathbb{E}(X_t \varepsilon_t) \end{aligned}$$

Or

$$\begin{aligned} \mathbb{E}(X_t \varepsilon_t) &= \mathbb{E}[(\varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p})\varepsilon_t] + \mathbb{E}(\varepsilon_t^2) \\ &= \sigma^2 \end{aligned}$$

$\mathbb{E}[(\varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p})\varepsilon_t] = 0$ , car  $\varepsilon_t \perp \varepsilon_{t-i}$ ,  $\forall i \geq 1$

$$\gamma(0) = \varphi_1 \gamma(1) + \dots + \varphi_p \gamma(p) + \sigma_\varepsilon^2$$

Si  $h > 0$ , on procède de la même façon :

$$\begin{aligned}
X_t X_{t-h} &= \varphi_1 X_{t-1} X_{t-h} + \cdots + \varphi_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h} \\
\gamma(h) &= \varphi_1 \gamma(h-1) + \cdots + \varphi_p \gamma(h-p) + \underbrace{\mathbb{E}(\varepsilon_t X_{t-h})}_{=0}
\end{aligned}$$

- Les auto-corrélations :

A partir de la relation de récurrence de  $\gamma(h)$  on déduit celle sur  $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$ .

$$\rho(h) = \varphi_1 \rho(h-1) + \cdots + \varphi_p \rho(h-p), \forall h \geq 0$$

- Ces dernières équations sont appelées équations de Yule-Walker.

Pour  $h > 0$ , les  $\gamma(h)$  et les  $\rho(h)$  vérifient une relation de récurrence d'ordre  $p$  et

$$\begin{aligned}
1 &= \varphi_1 \rho(1) + \cdots + \varphi_p \rho(p) + \frac{\sigma_\varepsilon^2}{\gamma(0)} \\
\Rightarrow \gamma(0) &= \sigma_\varepsilon^2 \frac{1}{1 - (\varphi_1 \rho(1) + \cdots + \varphi_p \rho(p))}
\end{aligned}$$

Les équations de Yule-Walker pour  $h = 1, \dots, p$  peuvent s'écrire :

$$\begin{pmatrix} 1 & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \rho(p-2) \\ \vdots & & \ddots & \vdots \\ \rho(p-1) & \cdots & & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}$$

Les solutions de l'équation de récurrence sont complètement déterminées par la donnée de conditions initiales  $\rho(1), \dots, \rho(p)$  : elles permettent d'obtenir  $\varphi_1, \dots, \varphi_p$ . En particulier elles donneront une estimation préliminaire de  $\hat{\varphi}_1, \dots, \hat{\varphi}_p$  en fonction de  $\hat{\rho}_T(1), \dots, \hat{\rho}_T(p)$ .

$$\begin{aligned}
\rho(1) &= \varphi_1 + \varphi_2 \rho(1) + \cdots + \varphi_p \rho(p-1) \\
\rho(p) &= \varphi_1 \rho(p-1) + \cdots + \varphi_{p-1} \rho(1) + \varphi_p \\
\varphi_1 &= (1 - \varphi_2) \rho(1) - \cdots - \varphi_p \rho(p-1) \\
&\vdots \\
\varphi_p &= \rho(p) - \varphi_1 \rho(p-1) - \cdots + \varphi_{p-1} \rho(1)
\end{aligned}$$

### 1.2.3 Estimation dans le modèle autorégressif d'ordre $p$ ( $AR(p)$ ) 20

On peut donc aussi obtenir  $\rho(1), \dots, \rho(p)$  en fonction de  $\varphi_1, \dots, \varphi_p$ .

**Proposition 1.2.2.** *Si  $X_t \rightsquigarrow AR(p)$  alors les  $|\rho(h)|$  et les  $\gamma(h)$  décroissent vers 0 exponentiellement avec  $h$ .*

**Proposition 1.2.3.** *Si  $(X_t)_{t \in \mathbb{Z}} \rightsquigarrow AR(p)$  et si  $\phi(L)X_t = \mu + \varepsilon_t$  est sa représentation canonique, alors :*

$$r(h) = \begin{cases} 0 & \text{si } h > p \\ \neq 0 & \text{sinon.} \end{cases}$$

### 1.2.3 Estimation dans le modèle autorégressif d'ordre $p$ ( $AR(p)$ )

Nous allons maintenant nous intéresser à l'estimation des paramètres d'un autorégressif d'ordre  $p$ .

#### Estimation pour les modèles $AR(p)$ par la méthode des moindres carrés

Un modèle  $AR(p)$  s'écrit

$$\begin{aligned} X_t &= \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \\ &= Z_t' \beta + \varepsilon_t \end{aligned}$$

où  $Z_t' = (1, X_{t-1}, X_{t-2}, \dots, X_{t-p})$ ,  $\beta' = (\mu, \phi_1, \phi_2, \dots, \phi_p)$ , et  $(\varepsilon_t)$  est un bruit blanc.

L'estimation des paramètres du modèle

$X_t = Z_t' \beta + \varepsilon_t$  par la méthode des moindres carrés est donnée par

$$\hat{\beta} = (Z_t' Z_t')^{-1} Z_t' X_t \text{ et } \hat{\sigma}^2 = \frac{1}{T-p} \sum_{t=p+1}^T (X_t - Z_t' \hat{\beta})^2.$$

Toutefois, les résultats usuels d'économétries ne sont pas vérifiés ici, en particulier  $\mathbb{E}(\hat{\beta}) \neq \beta$ . Il est toutefois possible de montrer le résultat suivant,

**Proposition 1.2.4.** *Si les racines du polynôme caractéristique (racines de  $\phi(z) = 0$ ) sont à l'extérieur du disque unité alors,*

$$\hat{\beta} \xrightarrow{\mathbb{P}} \beta \text{ et } \hat{\sigma}^2 \xrightarrow{\mathbb{P}} \sigma^2$$

### 1.2.3 Estimation dans le modèle autorégressif d'ordre $p$ ( $AR(p)$ ) 21

et de plus

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{\mathcal{L}} N(0, \sigma^2 V), \text{ où } V = p \lim_{T \rightarrow \infty} \frac{1}{T} Z_t Z_t'$$

#### Estimation par maximum de vraisemblance

Pour déterminer la vraisemblance, il est nécessaire de supposer connue la loi des erreurs : nous supposons les erreurs normalement distribuées. Les erreurs étant normalement distribuées et indépendantes (le processus  $(\varepsilon_t)$  est, par hypothèse un bruit blanc), le vecteur  $(\varepsilon_1, \dots, \varepsilon_T)$  est un vecteur gaussien. Les composantes du vecteur  $(X_1, \dots, X_T)$  étant obtenues par combinaisons linéaires des composantes du vecteur,  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)$ ,  $X = (X_1, \dots, X_T)$  sera un vecteur gaussien, alors  $\varepsilon \sim N(0, \sigma^2 I)$ ,  $X \sim N(0, \sigma^2 \Omega)$  :

$$\mathcal{L}(X = (X_1, \dots, X_T)', \phi, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \frac{1}{[\det \Omega]^{1/2}} \exp\left(-\frac{1}{2\sigma^2} X' \Omega^{-1} X\right),$$

La log-vraisemblance est alors donnée par :

$$\ln \mathcal{L}(X = (X_1, \dots, X_T)', \phi, \sigma^2) = -\frac{T}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln |\det \Omega| - \frac{1}{2\sigma^2} X' \Omega^{-1} X$$

Le problème est que cette log-vraisemblance est difficile à calculer et donc à maximiser à cause de  $(\det \Omega)$  et de  $\Omega^{-1}$  (matrice  $T \times T$ ). De plus, il faut se donner des valeurs préliminaires pour les paramètres, on utilise alors l'algorithme de maximisation itératifs.

**Exemple 4.** Dans le cas d'un modèle  $AR(1)$ , de la forme  $X_t = c + \phi X_{t-1} + \varepsilon_t$  où  $\varepsilon_t$  est i.i.d. et distribué suivant une loi  $\mathcal{N}(0, \sigma^2)$ , avec  $|\phi| < 1$ , alors

$$X_t | X_{t-1} \sim \mathcal{N}(c + \phi X_{t-1}, \sigma^2).$$

Aussi, la loi conditionnelle de  $X_t$  est donnée par

$$f(x_t | x_{t-1}, (c, \phi, \sigma^2)) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} (x_t - c - \phi x_{t-1})^2\right],$$

c'est à dire que

$$X_t \sim \mathcal{N}(\mathbb{E}(X_t), V(X_t)) \quad \text{soit} \quad X_t \sim \mathcal{N}\left(\frac{c}{1-\phi}, \frac{\sigma^2}{1-\phi^2}\right).$$

En posant  $\beta = (c, \sigma^2)$ , la vraisemblance conditionnelle du modèle est alors donnée par

$$\mathcal{L}(\beta|X_1, \dots, X_T) = \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2\sigma^2} (X_t - c - \phi X_{t-1})^2 \right],$$

$$\log \mathcal{L}(\beta|X_1, \dots, X_T) = -\frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln \sigma^2 + \frac{1}{2\sigma^2} \sum_{t=2}^T (X_t - c - \phi X_{t-1})^2.$$

La vraisemblance marginale s'écrivant

$$\mathcal{L}(\beta, X_1) = \sqrt{\frac{1-\phi^2}{2\pi\sigma^2}} \exp \left[ -\frac{(1-\phi^2)}{2\sigma^2} \left( X_1 - \frac{c}{1-\phi} \right)^2 \right],$$

on en déduit la forme de la log-vraisemblance (exacte, et non plus conditionnelle),

$$\log \mathcal{L}(\beta, X_1, \dots, X_T) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \left( \frac{\sigma^2}{1-\phi^2} \right) - \frac{(1-\phi^2)}{2\sigma^2} \left( X_1 - \frac{c}{1-\phi} \right)^2 - \frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^T (X_t - c - \phi X_{t-1})^2.$$

On peut noter que la maximisation de la vraisemblance exacte est un problème d'optimisation non-linéaire.

## 1.2.4 Prévision dans un modèle autorégressif $AR(p)$

Le modèle s'écrit ,

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad \text{où} \quad \phi(L)X_t = \varepsilon_t$$

La prévision optimale à la date  $T+1$ , faite à la date  $T$  est  ${}_T X_{T+1}^* = EL(X_{T+1}|X_T, X_{T-1}, \dots)$ . Aussi,

$${}_T X_{T+1}^* = \phi_1 X_T + \dots + \phi_p X_{T-p}$$

car  $(\varepsilon_t)$  est l'innovation. De façon analogue,  $X_{T+h} = \phi_1 X_{T+h-1} + \dots + \phi_p X_{T+h-p} + \varepsilon_{T+h}$ , et donc  ${}_T X_{T+h}^* = EL(X_{T+h}|X_T, X_{T-1}, \dots)$  est donné, de façon récursive

$$\text{par } {}_T X_{T+h}^* = \begin{cases} \phi_{1.T} X_{T+h-1}^* + \dots + \phi_{h-1.T} X_{T+1}^* + \phi_h X_T + \dots + \phi_p X_{T+h-p} & \text{pour } h \leq p \\ \phi_{1.T} X_{T+h-1}^* + \dots + \phi_{p.T} X_{T+h-p}^* & \text{pour } h > p \end{cases}$$

**Exemple 5.** On considère un AR tel que  $\phi(L) = (1 - \varphi L)^2$  où  $|\varphi| < 1$ . On a donc le processus :

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

On veut prévoir  ${}_tX_{t+h}^*$ .

On commence par centrer :  $\phi(L)Y_t = \varepsilon_t$

$$\phi(L) \underbrace{({}_tX_{t+h}^* - m)}_{{}_tY_{t+h}^*} = 0$$

${}_tY_{t+h}^*$  s'écrit :  ${}_tY_{t+h}^* = \varphi^h(ah + b)$ .

Si  $h = 0$ ,  ${}_tY_t^* = Y_t = b$

Si  $h = 1$ ,  ${}_tY_{t+1}^* = 2\varphi Y_t - \varphi^2 Y_{t-1} = \varphi(a + b)$

Par identification  $a = Y_t - \varphi Y_{t-1}$  et  $b = Y_t$ .

On en déduit

$${}_tY_{t+h}^* = \varphi^h((Y_t - \varphi Y_{t-1})h + Y_t)$$

**Exemple 6.** Dans le cas d'un processus  $AR(1)$  tel que  $X_t = \mu + \phi X_{t-1} + \varepsilon_t$  alors

(i)  ${}_T X_{T+1}^* = \mu + \phi X_T$ ,

(ii)  ${}_T X_{T+2}^* = \mu + \phi {}_T X_{T+1}^* = \mu + \phi[\mu + \phi X_T] = \mu[1 + \phi] + \phi^2 X_T$ ,

(iii)  ${}_T X_{T+3}^* = \mu + \phi {}_T X_{T+2}^* = \mu + \phi[\mu + \phi[\mu + \phi X_T]] = \mu[1 + \phi + \phi^2] + \phi^3 X_T$ ,

et récursivement, on peut obtenir  ${}_T X_{T+h}^*$  de la forme

$${}_T X_{T+h}^* = \mu + \phi {}_T X_{T+h-1}^* = \mu [1 + \phi + \phi^2 + \dots + \phi^{h-1}] + \phi^h X_T$$

# Chapitre 2

## Processus autorégressif vectoriel

### 2.1 Préliminaires

#### 2.1.1 Série temporelle multivariée

La compréhension de certains phénomènes aléatoires nécessite une modélisation simultanée de plusieurs séries. Les séries chronologiques multivariées se trouvent ainsi très utiles dans de nombreuses applications. C'est d'ailleurs de ces séries dont il sera question tout au long de ce travail. Introduisons les définitions suivantes.

**Définition 2.1.1.** Soit  $(\Omega, \mathcal{F}, P)$  un espace probabilisé, sur lequel nous considérons une suite de variables aléatoires réelles  $\{X_t, t \in \mathbb{Z}\}$ . Pour tout  $t \in \mathbb{Z}$ , et tout  $\omega \in \Omega$ ,  $X_t(\omega)$  est une réalisation de  $X_t$ .

**Définition 2.1.2.** Une série chronologique multivariée est définie comme une suite d'observations d'un processus stochastique multivarié  $\{X_t, t \in \mathbb{Z}\}$ . Pour  $t$  fixé,  $X_t = (X_{1t}, \dots, X_{nt})'$  représente un vecteur de dimension  $n$ .

**Définition 2.1.3.** (*Processus vectoriel du second ordre*)  $(X_t)$  est un processus du second ordre si et seulement si

$$\Leftrightarrow \forall t, X_t \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, \mathcal{A}, \mathbb{P})$$

$$\Leftrightarrow \forall i, \forall t, x_{it} \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$$



$$\Leftrightarrow X_t' X_t = \| X_t \|_2^2 = \sum_{i=1}^n X_{it}^2 \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$$

$$\text{On a } \mathbb{E}X_t = \begin{pmatrix} \mathbb{E}X_{1t} \\ \vdots \\ \mathbb{E}X_{nt} \end{pmatrix} = \mu_t \text{ et } \mathbb{V}X_t = \mathbb{E}(X_t - \mu)(X_t - \mu)' = (\text{Cov}(X_{it}, X_{jt}))_{1 \leq i, j \leq n}.$$

## 2.1.2 Processus stationnaire vectoriel

L'un des concepts jouant un rôle important dans l'étude des séries chronologiques est la notion de stationnarité. Ce concept permet des simplifications importantes dans l'analyse de données dépendantes et facilite définitivement l'inférence statistique. On distingue deux types de stationnarité, la stationnarité stricte ou forte et la stationnarité au sens large ou faiblement.

**Définition 2.1.4. (*Stationnarité stricte ou forte*)** Un processus  $\{X_t, t \in \mathbb{Z}\}$  de dimension  $n$  est stationnaire au sens strict ou strictement stationnaire dans le cas vectoriel si pour tout  $k$  entier naturel et  $t_1, \dots, t_k$  et  $h$  entiers relatifs, la distribution de  $X_{t_1}, \dots, X_{t_k}$  est la même que celle de  $X_{t_1+h}, \dots, X_{t_k+h}$ .

C'est une notion assez difficile à vérifier en pratique ; la notion suivante est beaucoup plus facile à vérifier empiriquement.

**Définition 2.1.5. (*stationnaire au sens large ou faiblement stationnaire*)** Soit  $(X_t)$  est un processus du second ordre.

On dit que  $(X_t)$  est faiblement stationnaire si et seulement si :

- (i)  $\mathbb{E}X_t = \mu$
- (ii)  $\mathbb{V}X_t = \Sigma_X = \Gamma_X(0)$
- (iii)  $\mathbb{E}(X_t - \mu)(X_{t+h} - \mu)' = \Gamma_X(h) = (\gamma_{ij})_{i,j \in \{1, \dots, n\}}$  ( $\Gamma$  est appelée la fonction d'autocovariance de  $X$ )

**Remarque 2.1.1.** (1) (iii)  $\Rightarrow$  (ii)

(2) (iii)  $\Rightarrow \text{Cov}(X_{it}, X_{j,t-h}) = \gamma_{ij}(h), \forall h \in \mathbb{Z}$

(3)  $(X_t)$  stationnaire  $\Rightarrow (x_{it})$  stationnaire  $\forall i$ , mais la réciproque est fausse en général car les conditions :

$$\mathbb{E}X_t = \mu \Rightarrow \mathbb{E}X_{it} = \mu_i, \forall i$$

$$\text{Cov}(X_{it}, X_{i,t-h}) = \gamma_i(h)$$

ne suffisent pas pour avoir la stationnarité de  $(X_t)$ .

**Exemple 7.** (1)  $\varepsilon_t \rightsquigarrow \mathbf{BB}(0, \Sigma_\varepsilon)$  où  $\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{nt} \end{pmatrix}$

C'est un processus stationnaire car par définition on a :

(i)  $\mathbb{E}\varepsilon_t = 0$

(ii)  $\mathbb{V}\varepsilon_t = \Sigma_\varepsilon$

(iii)  $\mathbb{E}\varepsilon_t \varepsilon_\tau' = 0$  si  $t \neq \tau$

En général, il n'est pas nécessaire que les vecteurs constituant le processus bruit blanc soient indépendants mais seulement non-corrélés. Lorsque les vecteurs aléatoires  $\varepsilon_t$  sont non-corrélés, nous parlons alors de bruit blanc faible. Si les éléments de la séquence sont aussi indépendants alors on dit que le bruit blanc est fort.

(2)  $X_t \rightsquigarrow \text{VMA}(1)$

Où  $\varepsilon_t \rightsquigarrow \mathbf{BB}(0, \Sigma_\varepsilon)$ ,  $A \in \mathcal{M}_n(\mathbb{R})$  fixée. Alors  $X_t = \varepsilon_t - A\varepsilon_{t-1}$  est stationnaire car :

(i)  $\mathbb{E}X_t = 0$

(ii)

$$\begin{aligned} \mathbb{V}X_t &= \mathbb{E}(\varepsilon_t - A\varepsilon_{t-1})(\varepsilon_t - A\varepsilon_{t-1})' \\ &= \mathbb{E}[\varepsilon_t \varepsilon_t' - A\varepsilon_{t-1} \varepsilon_t' - \varepsilon_t \varepsilon_{t-1}' A' + A\varepsilon_{t-1} \varepsilon_{t-1}' A'] \\ &= \Sigma_\varepsilon + A\Sigma_\varepsilon A' \end{aligned}$$

(iii)  $\mathbb{E}X_t X_{t-1}' = \mathbb{E}(\varepsilon_t - A\varepsilon_{t-1})(\varepsilon_{t-1} - A\varepsilon_{t-2})'$

Pour  $h > 1$ ,  $\mathbb{E}X_t X_{t-h}' = 0$

$$\begin{aligned} \mathbb{E}X_t X_{t+1}' &= -\Sigma_\varepsilon A' \\ \mathbb{E}X_t X_{t+h}' &= 0, \text{ si } h > 1 \end{aligned}$$

**Définition 2.1.6.** Dans le cas où le processus  $\{X_t\}$  est stationnaire avec une matrice de covariance  $\Gamma(\cdot)$ , la série univariée  $\{X_{it}\}$  est stationnaire pour tout  $i$  avec fonction d'autocovariance  $\gamma_{ii}(\cdot)$ . La fonction  $\gamma_{ij}(\cdot)$ ,  $i \neq j$ , est appelée la fonction de covariance croisée des deux séries  $\{X_{it}\}$  et  $\{X_{jt}\}$ .

**Propriété 1.** Voici quelques propriétés de la matrice de covariance  $\Gamma(\cdot)$ .

(i)  $\gamma_{ij}(h) = \gamma_{ji}(-h)$ .

(ii)  $\Gamma(h) = \Gamma'(-h)$ .

(iii)  $|\gamma_{ij}(h)| \leq [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}, i, j = 1, \dots, n;$

(iv)  $\Gamma(\cdot)$  est définie semi-positive, c'est-à-dire  $\sum_{k=1}^m \sum_{l=1}^m \alpha'_k \Gamma(k-l) \alpha_l \geq 0$ , pour tout  $m \in \mathbb{N}$  et pour tout vecteur réel,  $\alpha_1, \dots, \alpha_m \in \mathbb{R}^d$ .

**Définition 2.1.7.** La matrice de corrélation  $R(\cdot)$  du processus  $\{X_t\}$  est définie par

$$R(h) = D^{-1/2} \Gamma(h) D^{-1/2} = [\rho_{ij}(h)]_{i,j=1}^n = [\gamma_{ij}(h) / [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}]_{i,j=1}^n$$

où  $D = \text{diag}[\gamma_{11}(0), \dots, \gamma_{nn}(0)]$  est une matrice diagonale dont le  $i^{\text{me}}$  élément de la diagonale est la variance de la  $i^{\text{me}}$  composante du processus  $\{X_t\}$  et  $\rho_{ii}(\cdot)$  est la fonction d'autocorrélation de la  $i^{\text{me}}$  composante du processus  $\{X_t\}$ .

**Remarque 2.1.2.** La matrice de corrélation  $R(\cdot)$  possède les mêmes propriétés que la matrice de covariance  $\Gamma(\cdot)$ , à l'exception de la propriété additionnelle suivante

$$\rho_{ii}(0) = 1, \forall i = 1, \dots, n.$$

**Proposition 2.1.1.** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire.

Soit  $(A_j)_{j \in \mathbb{Z}}$  une suite de matrices telle que  $\sum_{j \in \mathbb{Z}} \|A_j\| < +\infty$ . Alors

(i)  $Y_t = \sum_j A_j X_{t-j} \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, \mathcal{A}, \mathbb{P})$

(ii)  $Y_t$  est stationnaire avec :

$$- \mathbb{E}Y_t = \mu_Y = \left( \sum_j A_j \right) \mu_X$$

$$- \Gamma_Y(h) = \sum_{j,k} A_j \Gamma_X(h+k-j) A'_k$$

### 2.1.3 Innovation d'un processus vectoriel

$$\text{Soit } X_t = \begin{pmatrix} X_{1t} \\ \vdots \\ X_{nt} \end{pmatrix}$$

**Définition 2.1.8. (Prévision)**  $X_t^* = EL(X_t | \underline{X_{t-1}}) = \begin{pmatrix} EL(X_{1t} | \underline{X_{t-1}}) \\ \vdots \\ EL(X_{nt} | \underline{X_{t-1}}) \end{pmatrix}$

où  $\forall j \in \mathbb{Z}$

$$\begin{aligned} EL(X_{jt} | \underline{X_{t-1}}) &= EL(X_{jt} | \underline{X_{1t-1}}, \dots, \underline{X_{nt-1}}) \\ &= EL(X_{jt} | \underline{X_{1s}}, \dots, \underline{X_{ns}}, s \leq t-1) \\ &= EL(X_{jt} | \underline{X_{is}}, i = 1, \dots, n, s \leq t-1) \end{aligned}$$

**Proposition 2.1.2.** Soit  $X_t^*$  la *prévision linéaire optimale* de  $X_t$  à la date  $t-1$ .  $\hat{X}_t$  est une autre *prévision linéaire* de  $X_t$  à la date  $t-1$ , alors

$$\mathbb{V}(X_t - X_t^*) \leq \mathbb{V}(X_t - \hat{X}_t)$$

au sens des matrices symétriques positives

On a égalité si et seulement si  $\hat{X}_t = X_t^*$ .

**Remarque 2.1.3.** (1)  $\hat{X}_t$  est une *prévision linéaire* de  $X_t$  en  $t-1$ , i.e.  $\hat{X}_t$  est une fonction linéaire de  $X_{js}$ ,  $s \leq t-1$ ,  $j = 1, \dots, n$ .

$$\begin{aligned} \mathbb{V}(X_t - X_t^*) &\leq \mathbb{V}(X_t - \hat{X}_t) \\ \Rightarrow \forall u \in \mathbb{R}^n - \{0\}, u' \mathbb{V}(X_t - X_t^*) u &< u' \mathbb{V}(X_t - \hat{X}_t) u \end{aligned}$$

En particulier pour  $u = (0, \dots, 0, \underbrace{1}_{i^{me} position}, 0, \dots, 0)$ , on a :

$$\mathbb{V}(X_{it} - X_{it}^*) \leq \mathbb{V}(X_{it} - \hat{X}_{it})$$

$X_{it}^*$  est la *prévision optimale* de  $X_{it}$  comme fonction linéaire des  $X_{is}$ ,  $s \leq t-1$ .

(2)  $\forall u \in \mathbb{R}^n - \{0\}$ ,  $\mathbb{V}u'(X_t - X_t^*) < \mathbb{V}u'(X_t - \hat{X}_t)$   
 $\Rightarrow u' X_t^*$  est une *prévision* de  $u' X_t$  plus précise que  $u' \hat{X}_t$ .

(3) Si on note

$$\begin{aligned} \bar{\mathcal{L}}(\underline{X_{t-1}}) &= \bar{\mathcal{L}}(\underline{X_{1t-1}}, \dots, \underline{X_{nt-1}}) \\ &= \bar{\mathcal{L}}(\underline{X_{1s}}, \dots, \underline{X_{ns}}, s \leq t-1) \end{aligned}$$

On a  $\bar{\mathcal{L}}(X_{it-1}) \subsetneq \bar{\mathcal{L}}(\underline{X_{t-1}})$

et si  $\tilde{X}_{it} = EL(X_{it} | \underline{X_{t-1}})$ ,  $\mathbb{V}(X_{it} - \tilde{X}_{it}) > \mathbb{V}(X_{it} - X_{it}^*)$  alors

C'est normal car on a plus d'informations dans  $\bar{\mathcal{L}}(\underline{X_{t-1}})$  que dans  $\bar{\mathcal{L}}(X_{it-1})$ .

**Proposition 2.1.3. (Processus des innovations)** Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stationnaire et  $X_t^* = EL(X_t | \underline{X}_{t-1})$ . L'innovation de  $X_t$  est

$$\varepsilon_t = X_t - X_t^*$$

On a plus précisément

$$\begin{aligned} \varepsilon_t &= \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{nt} \end{pmatrix} \\ &= \begin{pmatrix} X_{1t} - EL(X_{1t} | \underline{X}_{t-1}) \\ \vdots \\ X_{nt} - EL(X_{nt} | \underline{X}_{t-1}) \end{pmatrix} \\ &= \begin{pmatrix} X_{1t} - EL(X_{1t} | \{X_{1s}, \dots, X_{ns}, s \leq t-1\}) \\ \vdots \\ X_{nt} - EL(X_{nt} | \{X_{1s}, \dots, X_{ns}, s \leq t-1\}) \end{pmatrix} \end{aligned} \quad (2.1)$$

On montre que  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc.

**Remarque 2.1.4.** (1) ce qui a été présenté est valable dans le cas où  $\mathbb{E}X_t = 0$ . Quand  $\mathbb{E}X_t \neq 0$ , on définit

$$\bar{\mathcal{L}}(\underline{X}_{t-1}) = \bar{\mathcal{L}}(1, \underline{X}_{1t-1}, \dots, \underline{X}_{nt-1})$$

Puis

$$X_t^* = \begin{pmatrix} EL(X_{1t} | 1, \underline{X}_{1t-1}, \dots, \underline{X}_{nt-1}) \\ \vdots \\ EL(X_{nt} | 1, \underline{X}_{1t-1}, \dots, \underline{X}_{nt-1}) \end{pmatrix}$$

$$(2) \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{nt} \end{pmatrix} \text{ où } \varepsilon_{it} \rightsquigarrow \mathbf{BB}, \forall i \in \{1, \dots, n\}$$

Mais  $\varepsilon_{it}$  n'est pas l'innovation du processus univarié  $(X_{it})$ . En effet

$$X_{it} - EL(X_{it} | 1, \underline{X}_{1t-1}, \dots, \underline{X}_{nt-1}) \neq X_{it} - EL(X_{it} | \underline{X}_{1t-1})$$

**Théorème 2.1. (Théorème de Wold)** Soit  $(X_t)$  un processus vectoriel stationnaire et  $(\varepsilon_t)$  le processus des innovations.

On a alors  $\exists (A_k)_{k \in \mathbb{Z}}$  tel que

$$X_t = \mu_t + \sum_{\mathbb{Z}} A_k \varepsilon_{t-k}$$

avec  $A_0 = I$  et  $\sum_{\mathbb{Z}} \|A_k\| < +\infty$

### 2.1.4 Convergence des moments empiriques

Soit  $X_t = \begin{pmatrix} X_{1t} \\ \vdots \\ X_{nt} \end{pmatrix}$  et  $\mathbb{E}X_t = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$

On définit les estimateurs :

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T X_t = \bar{X}_T = \begin{pmatrix} \bar{X}_{1t} \\ \vdots \\ \bar{X}_{nt} \end{pmatrix}$$

$$\hat{\Gamma}(0) = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X}_T)(X_t - \bar{X}_T)'$$

$$\hat{\Gamma}(h) = \frac{1}{T-h} \sum_{t=1}^T (X_t - \bar{X}_T)(X_{t-h} - \bar{X}_T)'$$

**Proposition 2.1.4.** (i) **Convergence** :  $\hat{\mu}$  et  $\hat{\Gamma}(h)$  sont des estimateurs convergents respectivement de  $\mu$  et  $\Gamma(h)$ .

(ii) **Normalité asymptotique** :

$$\sqrt{T}(\hat{\mu} - \mu) \xrightarrow{loi} \mathcal{N}(0, \sum_{h \in \mathbb{Z}} \Gamma_X(h)).$$

## 2.2 Motivation

### Motivation

On considère deux processus stationnaires  $\{x_{1t}, t \in \mathbb{Z}\}$  et  $\{x_{2t}, t \in \mathbb{Z}\}$  définies par les relations suivantes :

$$x_{1t} = \mu_1 + \sum_{i=1}^p b_{1i}x_{1t-i} + \sum_{i=1}^p c_{1i}x_{2t-i} + \varepsilon_{1t} \quad (2.2)$$

$$x_{2t} = \mu_2 + \sum_{i=1}^p b_{2i}x_{1t-i} + \sum_{i=1}^p c_{2i}x_{2t-i} + \varepsilon_{2t} \quad (2.3)$$

Où les innovations  $\{\varepsilon_{1t}, t \in \mathbb{Z}\}$  et  $\{\varepsilon_{2t}, t \in \mathbb{Z}\}$  sont des bruits de variance respective  $\sigma_1^2$  et  $\sigma_2^2$ , et non corrélés :  $E(\varepsilon_{1t} \varepsilon_{2t-j}) = 0, \forall j \in \mathbb{Z}$  : On constate immédiatement que le processus vectoriel  $X_t = (x_{1t}, x_{2t})'$  peut s'écrire sous la forme d'un processus AR(p) : En effet :

$$A_0 = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad A_i = \begin{pmatrix} b_{1i} & c_{1i} \\ b_{2i} & c_{2i} \end{pmatrix}, \quad \forall i \in [1, p]$$

On définit un processus vectoriel  $\varepsilon_t$  i.i.d. tel que :

$$\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \quad E(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

Alors, on montre immédiatement que :

$$X_t = A_0 + \sum_{i=1}^p A_i X_{t-i} + \varepsilon_t \quad (2.4)$$

On qualifie cette représentation de processus VAR (Vectorial Autoregressive) d'ordre  $p$ , noté VAR(p), ce système initial donnée par les équations (2.2) et (2.3), ou par la définition matricielle (2.4). On constate que dans cette représentation le niveau de  $x_{2t}$  a un effet immédiat sur  $x_{1t}$  et vice versa.

C'est pourquoi on travaille généralement à partir de la forme réduite du modèle VAR.

## 2.3 Représentation générale

**Définition 2.3.1.** *Un processus stationnaire  $\{X_t\}$  autorégressif de dimension  $n$  et d'ordre  $p$ , noté  $\text{VAR}(p)$ , est défini par :*

$$X_t = \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \Rightarrow \phi(L)X_t = \mu + \varepsilon_t, \forall t \in \mathbb{Z},$$

où  $\phi(L) = I_n - \sum_{i=1}^p \phi_i L^i$  est une matrice de dimension  $n \times n$ , qui est l'opérateur autorégressif multivarié tel que  $\phi_p \neq 0$ ,  $(\phi_j)_{1 \leq j \leq p} \in (M_n(\mathbb{R}))^p$ ,  $\mu$  est un vecteur constant de dimension  $n \times 1$  et  $\varepsilon_t \rightsquigarrow \mathbf{BB}(0, \Sigma_\varepsilon)$ ,  $\Sigma_\varepsilon \in M_n(\mathbb{R})$ .

Le processus vectoriel des innovations  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est *i.i.d.*  $(0, \Sigma_\varepsilon)$ , et satisfait par conséquent les propriétés suivantes :

$$E(\varepsilon_t) = 0$$

$$E(\varepsilon_t \varepsilon_{t-j}') = \begin{cases} \Sigma_\varepsilon & j = 0; \\ 0 & j \neq 0. \end{cases}$$

De la même façon que pour un  $\text{AR}(1)$ , on peut donc exprimer le polynôme matriciel  $\phi(L)$ , de dimension  $(n, n)$ , de la façon suivante :

$$\phi(L) = \sum_{i=0}^p \phi_i L^i = I_n - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p, \phi_0 = 1$$

où  $I_n$  désigne la matrice identité  $(n, n)$ . On pose les définitions suivantes :

$$X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \\ \dots \\ X_{nt} \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \dots \\ \varepsilon_{nt} \end{pmatrix}, \quad \phi_i = \begin{pmatrix} \phi_{11}^i & \phi_{12}^i & \dots & \phi_{1n}^i \\ \phi_{21}^i & \phi_{22}^i & \dots & \dots \\ \dots & \dots & \phi_{j,k}^i & \dots \\ \phi_{n1}^i & \phi_{n2}^i & \dots & \phi_{nn}^i \end{pmatrix}, \quad \forall i \in [1, p]$$

On retrouve alors la forme réduite évoquée dans la motivation puisque, les processus  $\{X_{it}, t \in \mathbb{Z}\}$  sont respectivement définis en fonctions de leur passé et du passé des processus  $\{X_{jt}, t \in \mathbb{Z}\}$  pour  $j \neq i$ . Par exemple, pour  $X_{1t}$  on obtient,  $\forall t \in \mathbb{Z}$  :

$$\begin{aligned} X_{1t} &= \mu_j + \phi_{11}^1 X_{1t-1} + \phi_{12}^1 X_{2t-1} + \dots + \phi_{1n}^1 X_{nt-1} \\ &+ \phi_{11}^2 X_{1t-2} + \phi_{12}^2 X_{2t-2} + \dots + \phi_{1n}^2 X_{nt-2} \\ &+ \dots \\ &+ \phi_{11}^p X_{1t-p} + \phi_{12}^p X_{2t-p} + \dots + \phi_{1n}^p X_{nt-p} + \varepsilon_{1t} \end{aligned}$$



De façon plus générale, pour  $X_{jt}$ ,  $\forall j \in [1, n]$ , on a :

$$\begin{aligned} X_{jt} &= \mu_1 + \phi_{j1}^1 X_{1t-1} + \phi_{j2}^1 X_{2t-1} + \dots + \phi_{jj}^1 X_{jt-1} + \dots + \phi_{jn}^1 X_{nt-1} \\ &\quad + \phi_{j1}^2 X_{1t-2} + \phi_{j2}^2 X_{2t-2} + \dots + \phi_{jj}^2 X_{jt-2} + \dots + \phi_{jn}^2 X_{nt-2} \\ &\quad + \dots \\ &\quad + \phi_{j1}^p X_{1t-p} + \phi_{j2}^p X_{2t-p} + \dots + \phi_{jj}^p X_{jt-p} + \dots + \phi_{jn}^p X_{nt-p} + \varepsilon_{jt} \end{aligned}$$

**Proposition 2.3.1.** *Le processus VAR(p) représenté par l'équation  $\phi(L)X_t = \mu + \varepsilon_t$  est dit stable si :*

$$\det\{\phi(z)\} \neq 0, \text{ pour tout } z \in \mathbb{C} \text{ tel que } |z| \leq 1 :$$

Dans ce cas, le processus VAR(p) admet la représentation linéaire suivante :

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \psi(L)\varepsilon_t,$$

où les matrices  $\psi_i$  sont déterminées par :

$$\psi(L) = \sum_{i=0}^{\infty} \psi_i L^i = \phi^{-1}(L),$$

et les coefficients  $\{\psi_i\}$  sont absolument sommables, c'est-à-dire que :

$$\sum_{i=0}^{\infty} \|\psi_i\| < \infty,$$

avec  $\|\cdot\|$  correspondant à la norme euclidienne d'une matrice.

**Démonstration 1.**  $\phi(L) = I_n - \phi_1 L - \dots - \phi_p L^p = (\varphi_{ij}(L))_{1 \leq i, j \leq n}$  avec  $\varphi_{ij}(L) = \delta_{ij} - \phi_{ij}^1 L - \dots - \phi_{ij}^p L^p$

On note  $\phi(z) = I_n - \phi_1 z - \dots - \phi_p z^p = (\varphi_{ij}(z))_{1 \leq i, j \leq n}, \forall z \in \mathbb{C}$  avec  $\varphi_{ij}(z) = \delta_{ij} - \phi_{ij}^1 z - \dots - \phi_{ij}^p z^p$ .

On note  $\tilde{\phi}(L)$  la comatrice de  $\phi(L)$ .

On sait que

$$\begin{aligned} \phi(z)\tilde{\phi}(z) &= \tilde{\phi}(z)\phi(z) = \det \phi(z) I_n \\ \phi(L)\tilde{\phi}(L) &= \tilde{\phi}(L)\phi(L) = \det \phi(L) I_n \end{aligned}$$

On aura donc  $\phi(L)$  inversible  $\Rightarrow \det \phi(L)$  inversible.

On sait que si les racines de  $\det \phi(z)$  sont de module strictement plus grand que 1,  $\det \phi(L)$  a pour inverse

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

tel que  $a_0 = 1$  et  $\sum_{k \geq 0} |a_k| < +\infty$

On a donc

$$\begin{aligned} \phi(L)\tilde{\phi}(L) &= \det \phi(L)I_n \\ \Rightarrow \phi(L)\underbrace{\tilde{\phi}(L)(\det \phi(L))^{-1}}_{\phi(L)^{-1}} &= I_n \end{aligned}$$

avec  $\tilde{\phi}(L) = (\tilde{\varphi}_{ij}(L))_{1 \leq i, j \leq n}$  et  $d^\circ \tilde{\varphi}_{ij} \leq p^{n-1}$

$$\phi(L)^{-1} = \left( \sum_0^{\infty} a_k L^k \right) \tilde{\phi}(L) = \sum_0^{\infty} A_k L^k$$

avec  $A_0 = a_0 \tilde{\phi}(1)I_n$  et  $\sum_k \|A_k\| < +\infty$  car  $\sum_k |a_k| < +\infty$

**Exemple 8.** On considère un processus bi-varié  $\{X_t, t \in \mathbb{Z}\}$  admettant une représentation VAR(1) telle que :

$$\begin{aligned} x_{1t} &= 3 + 0.2x_{1t-1} + 0.7x_{2t-1} + \varepsilon_{1t} \\ x_{2t} &= 1 + 0.3x_{1t-1} + 0.4x_{2t-1} + \varepsilon_{2t} \end{aligned}$$

On pose  $X_t = (x_{1t}, x_{2t})'$  et  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ . Cette représentation peut s'exprimer sous la forme réduite suivante :

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

avec  $\mu = (3, 1)'$  et :

$$\phi(L) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix} L = \begin{pmatrix} 1 - 0.2L & -0.7L \\ -0.3L & 1 - 0.4L \end{pmatrix}$$

Donc la condition de stationnarité du processus  $\{X_t, t \in \mathbb{Z}\}$  se ramène à déterminer les racines du polynôme :

$$\det[\phi(L)] = (1 - 0.2L)(1 - 0.4L) - 0.21L^2$$

Les deux racines  $\lambda_1 = 1.30$  et  $\lambda_2 = -5.91$  sont supérieures à 1 en module, le processus VAR est stationnaire et stable. Donc les processus univariés  $\{x_{1t}, t \in \mathbb{Z}\}$  et  $\{x_{2t}, t \in \mathbb{Z}\}$  sont stationnaires.

On peut facilement démontrer que cette condition de stationnarité peut être exprimée en fonction des valeurs propres de la matrice  $\phi(L)$

### 2.3.1 Processus VAR(1)

Examinons le cas où le processus  $\{X_t\}$  est un VAR(1) admettant une écriture de la forme :

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

En procédant à des remplacements successifs, il est possible d'obtenir une expression pour  $\{X_t\}$  en fonction des erreurs  $\{\varepsilon_t\}$ . Ainsi, pour  $t = 1$ , on a l'expression suivante :

$$X_1 = \mu + \phi_1 X_0 + \varepsilon_1.$$

Un premier remplacement donne :

$$\begin{aligned} X_2 &= \mu + \phi_1 X_1 + \varepsilon_2, \\ &= \mu + \phi_1(\mu + \phi_1 X_0 + \varepsilon_1) + \varepsilon_2, \\ &= (I_n + \phi_1)\mu + \phi_1 \varepsilon_1 + \varepsilon_2 + \phi_1^2 X_0 \end{aligned}$$

En continuant les récursions,  $\{X_t\}$  s'écrit sous cette nouvelle forme :

$$X_t = (I_n + \phi_1 + \dots + \phi_1^{t-1})\mu + \phi_1^t X_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}.$$

À la limite, il est raisonnable de penser que le processus  $\{X_t\}$  écrit de la forme précédente admet l'écriture suivante, sous certaines conditions :

$$X_t = m + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i},$$

où  $m = (I_n - \phi_1)^{-1}\mu$  représente l'espérance du processus  $\{X_t\}$ . En fait, il peut être montré que  $\sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$  existe en moyenne quadratique. Cette représentation comme somme infinie d'un bruit blanc est la limite de l'équation

$$X_t = (I_n + \phi_1 + \dots + \phi_1^{t-1})\mu + \phi_1^t X_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}.$$

La condition de stabilité dans la proposition (2.3.1) est valide pour les processus VAR(p). Dans le cas  $p = 1$ , il est particulièrement simple de donner la condition de stabilité en fonction des valeurs propres de  $\phi_1$ . Ceci fait l'objet de la proposition suivante.

**Proposition 2.3.2.** *Un processus VAR(1) est stable si les valeurs propres de  $\phi_1$  sont toutes inférieures à un en module.*

**Démonstration 2.** *Afin de montrer la proposition précédente, notons que selon la proposition (2.3.1), on doit avoir*

$$\det(I_n - \phi_1 z) \neq 0, \forall |z| \leq 1.$$

*Et on doit distinguer deux cas, à savoir  $\det(\phi_1) \neq 0$  et  $\det(\phi_1) = 0$ . Si  $\det(\phi_1) \neq 0$ , alors  $\lambda \neq 0$  n'est pas solution de l'équation  $\det(\lambda I_n - \phi_1) = 0$ . De plus, les valeurs propres de  $\phi_1$  satisfont l'équation suivante :*

$$\det(\lambda I_n - \phi_1) = \lambda^n \det(I_n - \lambda^{-1} \phi_1) = 0.$$

*De cette relation nous concluons que les solutions satisfont  $|\lambda^{-1}| > 1$  ou encore que les valeurs propres doivent être strictement inférieures à un. Si  $\det(\phi_1) = 0$ , alors  $\lambda = 0$  est une solution de l'équation  $\det(\lambda I_n - \phi_1) = 0$  et cette valeur propre satisfait la condition. Pour les autres valeurs non nulles, la même argumentation que précédemment est valide, d'où le résultat.*

## 2.3.2 Représentation VAR(1) d'un VAR(p)

Un résultat qui s'avérera important pour la suite concerne l'écriture d'un processus VAR(p) sous la forme d'un certain processus VAR(1) de grande dimension.

Ceci permet d'appliquer des résultats obtenus dans le cas particulier d'un VAR(1) au processus plus général VAR(p). Ainsi, le processus VAR(p) peut se représenter comme suit :

$$X_t^* = \mu^* + \phi^* X_{t-1}^* + \varepsilon_t^*,$$

où

$$X_t^* = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix}, \varepsilon_t^* = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\mu^* = \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \phi^* = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{pmatrix}$$

Ainsi,  $X_t^*$ ,  $\varepsilon_t^*$  et  $\mu^*$  sont de dimension  $(np) \times 1$ , et  $\phi^*$  est une matrice de dimension  $(np) \times (np)$ . L'équation précédemment citée est la représentation VAR(1) d'un VAR(p). Ainsi, d'un point de vue théorique, la condition de stabilité d'un VAR(p) peut être reformulée en fonction des valeurs propres de la matrice  $\phi^*$ .

**Exemple 9.** On considère un processus bi-varié  $\{X_t, t \in \mathbb{Z}\}$  admettant une représentation VAR(2) telle que :

$$\begin{aligned} x_{1t} &= 3 + 0.2x_{1t-1} + 0.7x_{2t-1} - 0.4x_{1t-2} - 0.6x_{2t-2} + \varepsilon_{1t} \\ x_{2t} &= 1 + 0.3x_{1t-1} + 0.4x_{2t-1} - 0.1x_{1t-2} - 0.8x_{2t-2} + \varepsilon_{2t} \end{aligned}$$

On pose  $X_t = (x_{1t}, x_{2t})'$  et  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ . Cette représentation peut s'exprimer sous la forme réduite suivante :

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

avec  $\mu = (3, 1)'$  et :

$$\begin{aligned}\phi(L) &= \phi_0 + \phi_1 L + \phi_2 L^2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix} L + \begin{pmatrix} 0.4 & 0.6 \\ 0.1 & 0.8 \end{pmatrix} L^2 \\ &= \begin{pmatrix} 1 - 0.2L + 0.4L^2 & -0.7L + 0.6L^2 \\ -0.3L + 0.1L^2 & 1 - 0.4L + 0.8L^2 \end{pmatrix}\end{aligned}$$

Déterminons  $E(X)$  :

$$E(X_t) = \phi(1)^{-1}\mu = \begin{pmatrix} 1.2 & -0.1 \\ -0.2 & 1.4 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.59 \\ 1.8 \end{pmatrix} = m$$

On pose  $\forall t \in \mathbb{Z}$

$$X_{(4,1)}^* = \begin{pmatrix} X_t - \mu \\ X_{t-1} - \mu \end{pmatrix} = \begin{pmatrix} x_{1t} - 2.59 \\ x_{2t} - 1.8 \\ x_{1t-1} - 2.59 \\ x_{2t-1} - 1.8 \end{pmatrix}, \quad \varepsilon_{(4,1)}^* = \begin{pmatrix} \varepsilon_t \\ 0_{(2,1)} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \\ 0 \end{pmatrix}$$

et

$$\phi_{(4,4)}^* = \begin{pmatrix} \phi_1 & \phi_2 \\ I_2 & 0_{(2,2)} \end{pmatrix} = \begin{pmatrix} -0.2 & -0.7 & 0.4 & 0.6 \\ -0.3 & -0.4 & 0.1 & 0.8 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Alors on montre que la représentation VAR(1) du processus  $X_t^*$  est identique à la représentation VAR(p) du processus  $X_t$  puisque :

$$\begin{aligned}X_t^* &= \phi^* X_{t-1}^* + \varepsilon_t^* \\ \Leftrightarrow \begin{pmatrix} x_{1t} - 2.59 \\ x_{2t} - 1.8 \\ x_{1t-1} - 2.59 \\ x_{2t-1} - 1.8 \end{pmatrix} &= \begin{pmatrix} -0.2 & -0.7 & 0.4 & 0.6 \\ -0.3 & -0.4 & 0.1 & 0.8 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{1t-1} - 2.59 \\ x_{2t-1} - 1.8 \\ x_{1t-2} - 2.59 \\ x_{2t-2} - 1.8 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \begin{cases} x_{1t} &= 3 + 0.2x_{1t-1} + 0.7x_{2t-1} - 0.4x_{1t-2} - 0.6x_{2t-2} + \varepsilon_{1t} \\ x_{2t} &= 1 + 0.3x_{1t-1} + 0.4x_{2t-1} - 0.1x_{1t-2} - 0.8x_{2t-2} + \varepsilon_{2t} \end{cases}\end{aligned}$$

**Proposition 2.3.3.** *Le processus VAR(p) représenté par l'équation  $\phi(L)X_t = \mu + \varepsilon_t$ , est stable si toutes les valeurs propres de  $\phi^*$  sont inférieures à un en module ou de façon équivalente si :*

$$\det(I_{np} - \phi^* z) \neq 0; \text{ pour } |z| \leq 1.$$

et on a dans ce cas

$$(i) X_t = m + \sum_0^{\infty} A_k \varepsilon_{t-k} \text{ avec } A_0 = I \text{ et } \sum_{k \geq 0} \|A_k\| < +\infty$$

(ii)  $(\varepsilon_t)$  est le processus des innovations de  $(X_t)$

**Démonstration 3.** (i) Soit le modèle  $\phi(L)X_t = \mu + \varepsilon_t$  avec  $\phi(L)$  inversible. L'inverse de  $\phi(L)$  est donné par

$$\phi(L)^{-1} = (\det \phi(L))^{-1} \tilde{\phi}(L) = \sum_{k=0}^{\infty} A_k L^k$$

où  $A_0 = I_n$  et  $\sum_{k \geq 0} \|A_k\| < +\infty$

On a

$$\begin{aligned} X_t &= \phi(L)^{-1}(\mu + \varepsilon_t) \\ &= \phi(L)^{-1}\mu + \phi(L)^{-1}\varepsilon_t \\ &= m + \sum_0^{\infty} A_k \varepsilon_{t-k} \end{aligned}$$

(ii)

$$\begin{aligned} \phi(L)X_t &= \mu + \varepsilon_t \\ \Leftrightarrow X_t &= \mu + \sum_1^p \phi_k X_{t-k} + \varepsilon_t \\ \Leftrightarrow X_{it} &= \mu_i + \sum_{k=1}^p \left( \sum_{j=1}^n \varphi_{ij}^k X_{jt-k} \right) + \varepsilon_{it} \end{aligned}$$

D'où  $\varepsilon_{it} \in \mathcal{L}(X_{it}, 1, x_{1t-1}, \dots, x_{1t-p}, \dots, x_{nt-p})$

Puis

$$\overline{\mathcal{L}}(1, \underline{\varepsilon}_t) = \overline{\mathcal{L}}(1, \{\varepsilon_{is}, i = 1, \dots, n, s \leq t\}) \subset \overline{\mathcal{L}}(\underline{X}_t)$$

$$\begin{aligned} \forall i, X_{it} &= \mu_i + \sum_{k=0}^{\infty} \left( \sum_{j=1}^n a_{ij}^k \varepsilon_{jt-k} \right) \in \overline{\mathcal{L}}(1, \{\varepsilon_{is}, i = 1..n, s \leq t\}) = \overline{\mathcal{L}}(1, \underline{\varepsilon}_t) \\ &\Rightarrow \overline{\mathcal{L}}(1, \underline{X}_t) \subset \overline{\mathcal{L}}(1, \underline{\varepsilon}_t) \\ &\Rightarrow \overline{\mathcal{L}}(1, \underline{X}_t) = \overline{\mathcal{L}}(1, \underline{\varepsilon}_t) \end{aligned}$$

Montrons que  $\varepsilon_t = X_t - X_t^*$

$$\begin{aligned} X_t^* &= EL(X_t | 1, \underline{X}_{t-1}) \\ &= EL \left( \mu + \sum_1^p \phi_k X_{t-k} + \varepsilon_t | 1, \underline{X}_{t-1} \right) \\ &= \left( \begin{array}{c} EL \left( \mu_1 + \sum_{k=1}^p \left( \sum_{j=1}^n \varphi_{1j}^k X_{jt-k} \right) \varepsilon_{1t} | 1, \underline{X}_{t-1} \right) \\ \vdots \\ EL \left( \mu_n + \sum_{k=1}^p \left( \sum_{j=1}^n \varphi_{nj}^k X_{jt-k} \right) \varepsilon_{nt} | 1, \underline{X}_{t-1} \right) \end{array} \right) \\ &= \underbrace{\mu + \sum_1^p \phi_k X_{t-k}}_{X_t - \varepsilon_t} + \underbrace{EL(\varepsilon_t | 1, \underline{X}_{t-1})}_{\substack{EL(\varepsilon_t | 1, \varepsilon_{t-1}) \\ 0 \text{ car } \varepsilon_t \sim \text{BB}}} \end{aligned}$$

CQFD

## 2.4 Fonction d'autocovariance d'un processus VAR(p)

La fonction d'autocovariance d'un processus VAR(p) peut s'écrire de la façon suivante en considérant l'équation  $X_t^* = \mu^* + \phi^* X_{t-1}^* + \varepsilon_t^*$  :

$$\begin{aligned} \Gamma(h) &= E\{(X_t - \mu)(X_{t-h} - \mu)'\}, \\ &= E\left\{ \left( \sum_{i=0}^{h-1} \phi_i^* \varepsilon_{t-i} + \sum_{i=0}^{\infty} \phi_{h+i}^* \varepsilon_{t-h-i} \right) \left( \sum_{i=0}^{\infty} \phi_i^* \varepsilon_{t-h-i} \right)' \right\}, \\ &= \sum_{i=0}^{\infty} \phi_{h+i}^* \Sigma_{\varepsilon} \phi_i^{*'} \end{aligned}$$



La fonction d'autocovariance  $\{\Gamma(h), h \in \mathbb{Z}\}$  d'un VAR( $p$ ) peut également s'écrire sous la forme d'un système de Yule-Walker :

$$\begin{cases} \Gamma(0) = \sum_{i=1}^p \phi_i \Gamma'(i) + \Sigma_\varepsilon, & h = 0, \\ \Gamma(h) = \sum_{i=1}^p \phi_i \Gamma(h-i), & h > 0. \end{cases} \quad (*)$$

Les démonstrations des résultats précédentes s'obtiennent en prenant l'espérance du produit  $X_t - \mu$  avec  $\dot{X}_{t-h}$ , où  $\dot{X}_{t-h} = \phi_1 \dot{X}_{t-h-1} + \phi_2 \dot{X}_{t-h-2} + \dots + \phi_p \dot{X}_{t-h-p-1} + \varepsilon_{t-h}$ ,  $\dot{X}_{t-h} = X_{t-h} - \mu$ . On note pour  $h = 0$  que  $\Gamma(0) = \Gamma'(0)$ , et que  $E[(X_t - \mu)\varepsilon_t'] = \Sigma_\varepsilon$ , en utilisant  $X_t = \mu + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$ . On note que pour (\*) la stabilité du processus est invoquée, et le résultat est en fait obtenu en effectuant des manipulations analogues à ce qui est fait dans le cas univarié, en prenant cependant en considération que  $\Gamma(h) = \Gamma'(-h)$  pour  $h < 0$ .

## 2.5 Représentation MA( $\infty$ )

Considérons un processus vectoriel  $\{X_t, t \in \mathbb{Z}\}$ , de dimension  $(n, 1)$ , stationnaire.

**Remarque 2.5.1.** *Tout comme dans le cas univarié, sous la condition de stationnarité, il est possible d'appliquer le théorème de Wold et de représenter  $X_t$  sous la forme d'un processus vectoriel moyenne mobile infini VMA( $\infty$ ).*

*Nous allons à nouveau donner l'intuition du théorème de Wold appliqué aux processus vectoriels. On considère un processus stationnaire satisfaisant la représentation VAR( $p$ ) suivante,  $\forall t \in \mathbb{Z}$  :*

$$\phi(L)X_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t$$

*On sait d'après la propriété précédente que ce processus VAR( $p$ ) peut être réexprimer sous la forme d'un processus VAR(1) tel que :*

$$\phi^*(L)X_t^* = \varepsilon_t^* \iff X_t^* = \phi^* X_{t-1}^* + \varepsilon_t^*$$

où les processus  $X_t^*$  et  $\varepsilon_t^*$ , ainsi que la matrice  $\phi^*$  ont été définis précédemment. Dès lors, en itérant vers le passé, on montre que le processus  $X_t^*$  peut s'écrire sous la forme d'une moyenne mobile d'ordre fini à  $t$  donné.

$$X_t^* = \varepsilon_t^* + \phi^* \varepsilon_{t-1}^* + \phi^{*2} \varepsilon_{t-2}^* + \dots + \phi^{*k} X_{t-k}^*$$

Si l'on suppose que les valeurs propres de la matrice  $\phi^*$  sont strictement inférieures à 1 en module (condition de stabilité), alors  $\lim_{k \rightarrow \infty} \phi^{*k} = 0$  : Dès lors, le processus  $X_t^*$  peut s'écrire sous la forme :

$$X_t^* = \lim_{k \rightarrow \infty} \sum_{i=0}^k \phi^{*i} \varepsilon_{t-i}^* = \sum_{i=0}^{\infty} \phi^{*i} \varepsilon_{t-i}^* = \psi^*(L) \varepsilon_t^*$$

En reprenant la définition du processus  $X_t^*$ , on peut alors déterminer la décomposition de Wold associée au processus VAR( $p$ ).

**Proposition 2.5.1.** *Tout processus  $\{X_t, t \in \mathbb{Z}\}$ , de dimension  $(n, 1)$ , stationnaire, satisfaisant une représentation VAR( $p$ ), telle que,  $\forall t \in \mathbb{Z}$  :*

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

admet une représentation moyenne mobile convergente définie par :

$$X_t = m + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = m + \psi(L) \varepsilon_t$$

avec  $m = E(X_t) = \phi(1)^{-1} \mu$ , où  $\varepsilon_t$  est un bruit blanc vectoriel et où la séquence des matrices de dimension  $(n, n)$ ,  $\{\psi_i\}_{i=0}^{\infty}$  satisfait  $\psi_0 = I_n$  et est absolument sommable au sens où les éléments  $\psi_{j,k}^i$  de la matrice  $\psi_i$  satisfont la condition :

$$\sum_{s=0}^{\infty} |(\psi_{jk}^i)^s| < \infty, \forall i \geq 1, \forall (j, k) \in [1, n]^2$$

La condition de sommabilité, qui garantit la convergence des moments d'ordre deux du processus  $X_t$ , doit se comprendre dans ce contexte comme une condition de sommabilité sur une séquence de matrices  $\{\psi_i\}_{i=0}^{\infty}$ . Cette condition se ramène à démontrer la sommabilité de tous les éléments  $\psi_{jk}^i$  de

ces matrices, au sens traditionnel de la sommabilité d'une séquence de scalaires .

Il est possible de déterminer de façon générale la forme des matrices de l'opérateur polynômial associée à la représentation VMA( $\infty$ ).

**Proposition 2.5.2.** *Le polynôme matriciel  $\psi(L) = \sum_{i=0}^{\infty} \psi_i L^i$  associé à la représentation VMA( $\infty$ ) d'un processus VAR( $p$ ) stationnaire,  $\{X_t, t \in \mathbb{Z}\}$ ,  $\forall t \in \mathbb{Z}$  :*

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

satisfait la relation de récurrence suivante :

$$\psi_0 = I_n$$

$$\psi_s = \phi_1 \psi_{s-1} + \phi_2 \psi_{s-2} + \dots + \phi_p \psi_{s-p}, \quad \forall s \geq 1$$

avec  $\psi_s = 0, \forall s < 0$ .

**Démonstration 4.** *Une façon simple d'obtenir la représentation VMA( $\infty$ ) d'un processus VAR( $p$ ) consiste à identifier les polynômes matriciels  $\phi(L)^{-1}$  et  $\psi(L)$  : En effet, dans le cas d'un processus centré ( $\mu = 0$ ), on a :*

$$X_t = \phi(L)^{-1} \varepsilon_t = \psi(L) \varepsilon_t \iff \phi(L) \psi(L) = I_n$$

Cette égalité peut se réécrire sous la forme :

$$\lim_{k \rightarrow \infty} [(I_n - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(I_n - \psi_1 L - \psi_2 L^2 - \dots - \psi_p L^p - \dots - \psi_k L^k)] = I_n$$

Dès lors, on montre par identification des termes de même ordre que les matrices de l'opérateur polynômial associée à la représentation VMA( $\infty$ ) satisfont une équation de récurrence qui correspond à celle de la proposition (2.3.5).

**Exemple 10.** *On considère un processus bi-varié stationnaire  $\{Y_t, t \in \mathbb{Z}\}$  admettant une représentation VAR(1) telle que :*

$$\phi(L)Y_t = c + \varepsilon_t$$

avec  $\varepsilon_t$  i.i.d.  $(0, \Omega)$ ,  $c = (3, 1)'$  et :

$$\phi(L) = \phi_0 + \phi_1 L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix} L = \begin{pmatrix} 1 - 0.2L & -0.7L \\ -0.3L & 1 - 0.4L \end{pmatrix}$$

Par application du théorème de Wold, on sait que ce processus peut être représenté sous forme VMA( $\infty$ ) telle que :

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \mu + \psi(L) \varepsilon_t$$

Immédiatement, on montre que

$$\mu = E(Y_t) = \begin{pmatrix} 0.8 & -0.7 \\ -0.3 & 0.6 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 9.25 \\ 6.29 \end{pmatrix}$$

Par définition, on a  $\phi(L)\psi(L) = I_2$ , ce qui peut se réécrire sous la forme :

$$\lim_{k \rightarrow \infty} [(I_2 - \phi_1 L)(\psi_0 - \psi_1 L - \psi_2 L^2 - \dots - \psi_p L^p - \dots - \psi_k L^k)] = I_2$$

Par identification des membres de même terme, on montre que :

$$\psi_0 = I_2$$

$$\psi_1 = \phi_1 = \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}$$

$$\psi_2 = \phi_1 \psi_1 = \phi_1^2 = \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}^2$$

et de façon générale, on a :

$$\psi_n = \phi_1 \psi_{n-1} = \phi_1^n = \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}^n, \forall n \geq 1$$

On retrouve ainsi la formule générale que l'on avait établi par itération vers le passé dans le cas d'un VAR(1) :

$$Y_t = \mu + \psi(L) \varepsilon_t = \mu + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$$

On montre ainsi que :

$$Y_t = \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} = \begin{pmatrix} 9.25 \\ 6.29 \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}^i \begin{pmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \end{pmatrix}$$

**Remarque 2.5.2.** (1) d'après le théorème de Wold on a

$$X_t = m + \sum_0^{\infty} A_k \varepsilon_{t-k} = m + \phi(L)^{-1} \varepsilon_t \text{ avec } m = \mathbb{E}X_t = \phi(L)^{-1} \mu.$$

(2) Soit le modèle  $\phi(L)X_t = \mu + \varepsilon_t$ .

- a) Si  $\det \phi(z)$  a ses racines de module strictement plus grand que 1, alors  $X_t = m + \phi(L)^{-1} \varepsilon_t$ . On en déduit que  $(X_t)_{t \in \mathbb{Z}}$  est stable (représentation VMA( $\infty$ )).
- b) Si  $\det \phi(1) = 0$ , alors  $(X_t)$  ne peut pas être stationnaire car  $\det \phi(Z) = (1 - Z)^d \psi(Z)$  tel que les racines de  $\psi$  sont de module strictement plus grand que 1.

$$(1 - L)^d X_t = \psi(L)^{-1} \tilde{\phi}(L)(\mu + \varepsilon_t)$$

- c) Si  $\det \phi(Z)$  a toutes ses racines de module strictement supérieur à 1, alors
- il existe une solution  $(X_t)$  stationnaire,
  - il existe aussi des solutions non stationnaires en espérance :
- $Z_t = X_t + Y_t$  où  $Y_t$  est déterministe tel que  $\phi(L)Y_t$ .

(3) En statistique en économique, on applique un test dit test de Dickey-Fuller (ADF) pour tester la stationnarité des séries chronologiques il teste existence d'une racine unitaire

## 2.6 Estimation d'un VAR

L'estimation d'un processus VAR(p) peut se faire par :

- la méthode des moindres carrés ordinaires (MCO), pour chaque équation ;
- la méthode du maximum de vraisemblance.

La méthode du maximum de vraisemblance est la principale méthode utilisée (Hamilton, 1994). On présent brièvement cette méthode.

## 2.6.1 Estimation par le maximum de vraisemblance

Considérons un processus VAR(p) :

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

où  $\varepsilon_t$  est un bruit blanc de matrice de variance-covariance  $\Sigma_\varepsilon$ .

La vraisemblance conditionnellement à toutes les valeurs passées du processus s'écrit :

$$L(X_1, \dots, X_T) = \prod_{t=1}^T L(X_t | \underline{X}_{t-1})$$

où  $\underline{X}_{t-1}$  désigne la mémoire du processus  $X_t$  jusqu'à la date  $t - 1$  compris.

La vraisemblance du processus VAR(p) s'écrit alors :

$$L(X_1, \dots, X_T) = \prod_{t=1}^T \frac{1}{(\sqrt{2\pi})^N \sqrt{\det \Sigma_\varepsilon}}$$

$$\times \exp \left[ -\frac{1}{2} \sum_{t=1}^T (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}) \Sigma_\varepsilon^{-1} (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}) \right]$$

La log-vraisemblance du processus VAR(p), quant à elle, s'écrit :

$$\log L(X_1, \dots, X_T) = -\frac{TN}{2} \log 2\pi - \frac{T}{2} \log \Sigma_\varepsilon - \frac{1}{2} \sum_{t=1}^T \varepsilon_t' \Sigma_\varepsilon^{-1} \varepsilon_t$$

La maximisation de cette log-vraisemblance permet alors d'obtenir les estimations des paramètres  $\phi_1, \dots, \phi_p$  et de la matrice de variance-covariance  $\Sigma_\varepsilon$ .

## 2.7 Choix des retards

Pour déterminer le nombre de retards d'un modèle à retards échelonnés, nous avons présenté les critères d'Akaike et de Schwarz. Dans le cas de la représentation VAR, ces critères peuvent être utilisés pour déterminer l'ordre  $p$  du modèle. La procédure de sélection de l'ordre de la représentation consiste à estimer tous les modèles VAR pour un ordre allant de 0 à  $h$  ( $h$  étant le retard maximum admissible par la théorie économique ou par les données

disponibles). Les fonctions  $AIC(p)$  et  $SC(p)$  sont calculées de la manière suivante :

$$AIC(p) = \ln \left[ \det \left| \sum_{\varepsilon} \right| \right] + \frac{2n^2p}{k}$$

$$SC(p) = \ln \left[ \det \left| \sum_{\varepsilon} \right| \right] + \frac{n^2p \ln(k)}{k}$$

avec :  $n$  = nombre de variables du système ;  $k$  = nombre d'observations ;  
 $p$  = nombre de retards ;  $\Sigma_{\varepsilon}$  = matrice des variances covariances des résidus du modèle.

Le retard  $p$  qui minimise les critères  $AIC$  ou  $SC$  est retenu.

## 2.8 Prédiction dans un VAR stationnaire

### 2.8.1 Cas d'un VAR(1)

Considérons le cas d'un modèle VAR(1) centré tel que  $\forall t \in \mathbb{Z}$  :

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t$$

Supposons que l'on dispose de  $T$  réalisations  $(X_1, X_2, \dots, X_T)$ , d'un estimateur convergent  $\hat{\phi}_1^*$  de  $\phi_1$  et d'un estimateur convergent  $\hat{\phi}_0^*$  de  $\phi_0$ . La formule qui permet d'obtenir une prédiction de la réalisation à la date  $T+1$  du processus  $X_t$  est donc naturellement donnée par :

$$\hat{X}_{T+1} = E(X_{t+1} | X_T, X_{T-1}, \dots, X_1) = \hat{\phi}_0 + \hat{\phi}_1 X_T$$

A l'horizon  $T+2$ , on a :

$$\hat{X}_{T+2} = E(X_{t+2} | X_T, X_{T-1}, \dots, X_1) = \hat{\phi}_0 + \hat{\phi}_1 \hat{X}_{T+1} = (I + \hat{\phi}_1) \hat{\phi}_0 + \hat{\phi}_1^2 X_T$$

**Proposition 2.8.1.** *De la même façon à un horizon  $h$ , la prédiction d'un VAR(1) est donnée par :*

$$\hat{X}_{T+h} = E(X_{T+h} | X_T, X_{T-1}, \dots, X_1) = (I + \hat{\phi}_1 + \hat{\phi}_1^2 + \dots + \hat{\phi}_1^{h-1}) \hat{\phi}_0 + \hat{\phi}_1^h X_T$$

Dès lors, l'erreur de prévision s'écrit sous la forme :

$$\begin{aligned} X_{T+h} - \hat{X}_{T+h} &= X_{T+h} - E(X_{T+h} | X_T, X_{T-1}, \dots, X_1) \\ &= X_{T+h} - E(X_{T+h} | \varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1) \\ &= \sum_{i=0}^{h-1} \psi_1^i \varepsilon_{T+h-i} \end{aligned}$$

Par définition des bruits blancs, cette erreur de prévision a une espérance nulle. La matrice de variance covariance de l'erreur de prévision est donc :

$$\begin{aligned} &E \left[ (X_{T+h} - \hat{X}_{T+h})(X_{T+h} - \hat{X}_{T+h})' | X_T, X_{T-1}, \dots, X_1 \right] \\ &= E \left[ (\varepsilon_{T+h} + \psi_1 \varepsilon_{T+h-1} + \dots + \psi_1^{h-1} \varepsilon_{T+1})(\varepsilon_{T+h} + \psi_1 \varepsilon_{T+h-1} + \dots + \psi_1^{h-1} \varepsilon_{T+1})' \right] \\ &= \sum_{i=0}^{h-1} \psi_1^i (\psi_1^i)' \end{aligned}$$

**Proposition 2.8.2.** Pour un processus VAR(1), la matrice de variance covariance de l'erreur de prévision à un horizon  $h$  est déterminée par la relation :

$$E \left[ (X_{T+h} - \hat{X}_{T+h})' (X_{T+h} - \hat{X}_{T+h}) | X_T, X_{T-1}, \dots, X_1 \right] = \sum_{i=0}^{h-1} \psi_1^i (\psi_1^i)'$$

Les variances des erreurs de prévisions pour les processus univariés  $(X_{1,T}, X_{2,T}, \dots, X_{n,T})$  sont déterminés par la diagonale principale de cette matrice.

## 2.8.2 Cas d'un VAR(p)

A la date  $T$ , on souhaite effectuer une prévision de  $X_{T+h}$  et déterminer une région de confiance de la prévision de  $X_{T+h}$ .

On considère le modèle

$$\phi(L)X_T = \mu + \varepsilon_t$$

avec  $\det \phi(Z)$  ayant toutes ses racines de module strictement supérieur à 1.

On s'intéresse à la solution stationnaire  $(X_t)$ .

$$\begin{aligned} \hat{X}_{T+1} &= EL(X_{T+1} | 1, \underline{X}_T) \\ &\vdots \\ \hat{X}_{T+h} &= EL(X_{T+h} | 1, \underline{X}_T) \end{aligned}$$



Plus précisément on a

$$\begin{aligned}
\hat{X}_{T+1} &= \mu + \phi_1 X_T + \cdots + \phi_p X_{T+1+p} + \underbrace{EL(\varepsilon_{T+1}|1, X_T)}_0 \\
X_{T+2} &= \mu + \phi_1 X_{T+1} + \cdots + \phi_p X_{T+2+p} + \varepsilon_{T+2} \\
\hat{X}_{T+2} &= \mu + \phi_1 \hat{X}_{T+1} + \cdots + \phi_p X_{T+2+p} + \underbrace{EL(\varepsilon_{T+2}|1, X_T)}_0 \\
&\vdots \\
X_{T+h} &= \mu + \sum_1^p \phi_k X_{T+h-k} + \varepsilon_{T+h} \\
\hat{X}_{T+h} &= \mu + \sum_1^p \phi_k \hat{X}_{T+h-k} + 0
\end{aligned}$$

avec  $\hat{X}_{T+h-k} = X_{T+h-k}$  si  $k \geq h$ .

C'est la méthode itérative de calcul de la prévision de  $X_{t+h}$

### Région de confiance sous hypothèse de normalité

- Si  $h = 1$ ,  $X_{T+1} - \hat{X}_{T+1} = \varepsilon_{T+1} \rightsquigarrow \mathcal{N}(0, \Sigma_\varepsilon)$  car les  $\varepsilon_T$  sont iid et  $\rightsquigarrow \mathcal{N}(0, \Sigma_\varepsilon)$ . La région de confiance de  $\mathbb{R}^n$  dans laquelle  $X_{T+1}$  a une probabilité  $1 - \alpha$  de se trouver est

$$\left\{ (X_{T+1} - \hat{X}_{T+1})' \Sigma_\varepsilon^{-1} (X_{T+1} - \hat{X}_{T+1}) \leq q_{1-\alpha}(\mathcal{X}^2(n)) \right\}$$

- Si  $h = 2$ ,  $X_{T+2} - \hat{X}_{T+2} = \phi_1 (X_{T+1} - \hat{X}_{T+1}) \varepsilon_{T+2}$

$$\mathbb{V}(X_{T+2} - \hat{X}_{T+2}) = \phi_1 \Sigma_\varepsilon \phi_1' + \Sigma_\varepsilon$$

D'où

$$(X_{T+2} - \hat{X}_{T+2})' (\phi_1 \Sigma_\varepsilon \phi_1' + \Sigma_\varepsilon)^{-1} (X_{T+2} - \hat{X}_{T+2}) \rightsquigarrow \mathcal{X}^2(n)$$

On peut en déduire comme précédemment une région de confiance au niveau  $1 - \alpha$ .

### Exemple 11. Spécification, estimation et prévision d'un modèle VAR

Nous cherchons à modéliser, sous la forme VAR, la demande ( $x_{1t}$ ) et les prix ( $x_{2t}$ ) d'une matière première. Nous disposons des données trimestrielles CVS et en différences premières sur 18 ans (tableau 2.1).

Tableau 2.1 - Demande et prix CVS et en différences premières (extrait des données)

Date	$x_{1t}$	$x_{2t}$
0001Q1	-146,4	-75,6
0001Q2	19,2	1,8
...	...	...
0018Q3	19,2	-16,2
0018Q4	50,4	3,6

On demande :

1. de rechercher l'ordre du modèle VAR,
2. d'estimer les paramètres du modèle,
3. de calculer une prévision pour l'année 19

1) Rechercher l'ordre du modèle VAR

Nous allons utiliser les critères de Akaike et de Schwarz pour des décalages  $h$  allant de 0 à 4. Nous devons donc estimer quatre modèles différents et retenir celui dont les critères AIC et SC sont les plus faibles.

Estimation du modèle pour  $h = 1$

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} a_1^0 \\ a_2^0 \end{bmatrix} + \begin{bmatrix} a_{11}^1 & a_{11}^2 \\ a_{21}^1 & a_{21}^2 \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Ou encore :

$$\begin{aligned} x_{1t} &= a_1^0 + a_{11}^1 x_{1t-1} + a_{11}^2 x_{2t-1} + \varepsilon_{1t} \\ x_{2t} &= a_2^0 + a_{21}^1 x_{1t-1} + a_{21}^2 x_{2t-1} + \varepsilon_{2t} \end{aligned}$$

Pour estimer les paramètres de ce modèle, nous pouvons utiliser la méthode des moindres carrés ordinaires (MCO) pour chaque équation.

$$\begin{aligned} x_{1t} &= 0,00676 \times x_{1t-1} - 0,6125 \times x_{2t-1} + 17,129 \\ x_{2t} &= -0,1752 \times x_{1t-1} + 0,2992 \times x_{2t-1} - 12,862 \end{aligned}$$

On calcule

$$\det \sum_{\varepsilon} = \begin{bmatrix} \text{var}\varepsilon_1^2 & \text{var}\varepsilon_1\varepsilon_2 \\ \text{var}\varepsilon_1\varepsilon_2 & \text{var}\varepsilon_2^2 \end{bmatrix} = \begin{bmatrix} 1500,002 & -674,339 \\ -674,339 & 1200,074 \end{bmatrix} = 1345379.$$

Dans toute la suite des calculs nous retenons la méthode 1, soit :

$$\text{Var}(\varepsilon_{1t}) = 1500,002; \text{Var}(\varepsilon_{2t}) = 1200,074; \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = -674,339$$

$$\text{Soit } \det(\sum_{\varepsilon}) = \begin{bmatrix} 1500,002 & -674,339 \\ -674,339 & 1200,074 \end{bmatrix} = 1345379.$$

Pour  $p = 1$  :

$$AIC(p) = \ln \left[ \det \sum_{\varepsilon} \right] + \frac{2k^2p}{n} = \ln(1345379) + (2 \times 2^2 \times 1)/71 = 14,28$$

et

$$\begin{aligned} SC(p) &= \ln \left[ \det \left( \sum_{\varepsilon} \right) \right] + \frac{k^2p \ln(n)}{n} = \ln(1345379) + (2^2 \times 1 \times \ln(71))/71 \\ &= 14,47 \end{aligned}$$

Nous pouvons calculer de la même manière :

$$AIC(2) = 14,31; AIC(3) = 14,39; AIC(4) = 14,45$$

$$SC(2) = 14,56; SC(3) = 14,78; SC(4) = 14,97$$

Le nombre de décalages retenu correspond à la valeur la plus faible des critères soit :  $p = 1$ .

2) Estimation des paramètres

Le modèle VAR(1) estimé s'écrit :

$$\begin{aligned} x_{1t} &= \underset{(0,06)}{0,00676} \times x_{1t-1} - \underset{(4,47)}{0,6125} \times x_{2t-1} + \underset{(2,87)}{17,1292} + \varepsilon_{1t} \\ x_{2t} &= \underset{(1,84)}{-0,1752} \times x_{1t-1} + \underset{(2,44)}{0,2992} \times x_{2t-1} - \underset{(2,41)}{12,862} + \varepsilon_{2t} \end{aligned}$$

3) Prévision

La prévision est calculée en faisant « tourner » le modèle :

$$\begin{aligned} \hat{x}_{1,73} &= 0,0067 \times x_{1,72} - 0,6125 \times x_{2,72} + 17,129 \\ \hat{x}_{2,73} &= -0,1752 \times x_{1,72} + 0,2992 \times x_{2,72} - 12,862 \end{aligned}$$

$$\begin{aligned}\hat{x}_{1,73} &= 0,0067 \times 50,4 - 0,6125 \times 3,6 + 17,129 = 15,3 \\ \hat{x}_{2,73} &= -0,1752 \times 50,4 + 0,2992 \times 3,6 - 12,862 = -20,6 \\ \\ \hat{x}_{1,74} &= 0,0067 \times 15,3 - 0,6125 \times -20,6 + 17,129 = 29,9 \\ \hat{x}_{2,74} &= -0,1752 \times 15,3 + 0,2992 \times -20,6 - 12,862 = -21,7 \\ \\ \hat{x}_{1,75} &= 0,0067 \times 29,9 - 0,6125 \times -21,7 + 17,129 = 30,6 \\ \hat{x}_{2,75} &= -0,1752 \times 29,9 + 0,2992 \times -21,7 - 12,862 = -24,6 \\ \\ \hat{x}_{1,76} &= 0,0067 \times 30,6 - 0,6125 \times -24,6 + 17,129 = 32,4 \\ \hat{x}_{1,76} &= -0,1752 \times 30,6 + 0,2992 \times -24,6 - 12,862 = -25,6\end{aligned}$$

## 2.9 Causalité

Une des questions que l'on peut se poser à partir d'un VAR est de savoir s'il existe une relation de causalité entre les différentes variables du système. Il existe ainsi plusieurs définitions de la causalité :

- causalité au sens de Granger (1969) ;
- causalité au sens de Sims (1980).

### 2.9.1 Causalité au sens de Granger

Granger (1969) a proposé les concepts de causalité et d'exogénéité : la variable  $X_{2t}$  est la cause de  $X_{1t}$ , si la prédictibilité de  $X_{1t}$  est améliorée lorsque l'information relative à  $X_{2t}$  est incorporée dans l'analyse. Soit le modèle VAR(p) pour lequel les variables  $X_{1t}$  et  $X_{2t}$  sont stationnaires :

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} a_1^1 & b_1^1 \\ a_1^2 & b_1^2 \end{bmatrix} \begin{bmatrix} X_{1t-1} \\ X_{2t-1} \end{bmatrix} + \begin{bmatrix} a_2^1 & b_2^1 \\ a_2^2 & b_2^2 \end{bmatrix} + \dots + \begin{bmatrix} X_{1t-2} \\ X_{2t-2} \end{bmatrix} \begin{bmatrix} a_p^1 & b_p^1 \\ a_p^2 & b_p^2 \end{bmatrix} \begin{bmatrix} X_{1t-p} \\ X_{2t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Le bloc de variables  $(x_{2t-1}, x_{2t-2}, \dots, x_{2t-p})$  est considéré comme exogène par rapport au bloc de variables  $(x_{1t-1}, x_{1t-2}, \dots, x_{1t-p})$  si le fait de rajouter le bloc  $X_{2t}$  n'améliore pas significativement la détermination des variables  $X_{1t}$ . Ceci consiste à effectuer un test de restrictions sur les coefficients des variables  $X_{2t}$  de la représentation VAR (noté RVAR = Restricted VAR). Soit :

- $X_{2t}$  ne cause pas  $X_{1t}$  si l'hypothèse suivante est acceptée

$$H_0 : b_1^1 = b_2^1 = \dots = b_p^1 = 0.$$

- $X_{1t}$  ne cause pas  $X_{2t}$  si l'hypothèse suivante est acceptée

$$H_0 : a_1^2 = a_2^2 = \dots = a_p^2 = 0.$$

Si nous sommes amenés à accepter les deux hypothèses que  $X_{1t}$  cause  $X_{2t}$  et que  $X_{2t}$  cause  $X_{1t}$ , on parle de boucle rétroactive « feedback effect ».

### Application au cas d'un VAR(p) avec $n = 2$

Pour un VAR(p) avec  $n = 2$ , la condition de la causalité de Granger est immédiate à obtenir.

**Résultat,** Dans le système bi-varié suivant VAR(2) :  $\forall t \in \mathbb{Z}$

$$Z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^2 & \phi_{12}^2 \\ \phi_{21}^2 & \phi_{22}^2 \end{pmatrix} \begin{pmatrix} y_{t-2} \\ x_{t-2} \end{pmatrix} + \dots + \begin{pmatrix} \phi_{11}^p & \phi_{12}^p \\ \phi_{21}^p & \phi_{22}^p \end{pmatrix} \begin{pmatrix} y_{t-p} \\ x_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix}$$

la variable  $x_t$  ne cause pas la variable  $y_t$  si et seulement si :

$$\phi_{12}^1 = \phi_{12}^2 = \phi_{12}^3 = \dots = \phi_{12}^p = 0$$

Autrement dit, la variable  $x_t$  ne cause pas la variable  $y_t$  si et seulement si les matrices  $\phi_i$  sont toutes triangulaires inférieures,  $\forall i \in [1, p]$ .

En effet, réécrivons le processus sous cette condition, on a :

$$Z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11}^1 & 0 \\ \phi_{21}^1 & \phi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^2 & 0 \\ \phi_{21}^2 & \phi_{22}^2 \end{pmatrix} \begin{pmatrix} y_{t-2} \\ x_{t-2} \end{pmatrix} + \dots + \begin{pmatrix} \phi_{11}^p & 0 \\ \phi_{21}^p & \phi_{22}^p \end{pmatrix} \begin{pmatrix} y_{t-p} \\ x_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix}$$

Dès lors,

$$E(y_{t+h}|y_t, y_{t-1}, \dots, y_1) = c_1 + \phi_{11}^1 y_t + \phi_{11}^2 y_{t-1} + \dots + \phi_{11}^p y_{t-p+1}$$

$$E(y_{t+h}|y_t, y_{t-1}, \dots, y_1, x_t, x_{t-1}, \dots, x_1) = c_1 + \phi_{11}^1 y_t + \phi_{11}^2 y_{t-1} + \dots + \phi_{11}^p y_{t-p+1}$$

On a bien alors :

$$E(y_{t+h}|y_t, y_{t-1}, \dots, y_1) = E(y_{t+h}|y_t, y_{t-1}, \dots, y_1, x_t, x_{t-1}, \dots, x_1)$$

## 2.9.2 Causalité au sens de Sims

Sims (1980) présente une spécification de test légèrement différente, en considérant que si les valeurs futures de  $y_{1t}$  permettent d'expliquer les valeurs présentes de  $y_{2t}$ , alors  $y_{2t}$  est la cause de  $y_{1t}$ .

Ceci se traduit par la représentation suivante :

$$\begin{aligned}
 y_{1t} &= a_1^0 + \sum_{i=1}^p a_{1i}^1 y_{1t-i} + \sum_{i=1}^p a_{1i}^2 y_{2t-i} + \sum_{i=1}^p b_i^2 y_{2t+i} + \varepsilon_{1t} \\
 y_{2t} &= a_2^0 + \sum_{i=1}^p a_{2i}^1 y_{1t-i} + \sum_{i=1}^p a_{2i}^2 y_{2t-i} + \sum_{i=1}^p b_i^1 y_{1t+i} + \varepsilon_{2t}
 \end{aligned}$$

- $y_{1t}$  ne cause pas  $y_{2t}$  si l'hypothèse suivante est acceptée  
 $H_0 : b_1^2 = b_2^2 = \dots = b_p^2 = 0.$
- $y_{2t}$  ne cause pas  $y_{1t}$  si l'hypothèse suivante est acceptée  
 $H_0 : b_1^1 = b_2^1 = \dots = b_p^1 = 0.$

## 2.10 Dynamique d'un modèle VAR

Les modèles VAR sont souvent analysés au travers leur dynamique et ce via la simulation de chocs aléatoires et l'analyse de la décomposition de leurs variance.

### 2.10.1 Analyse impulsionnelle

Nous arrivons maintenant au cœur de l'analyse des modèles VAR. Un modèle VAR modélise par essence les relations dynamiques entre un groupe de variables choisies pour caractériser un phénomène économique particulier. L'analyse impulsionnelle va permettre de déterminer l'influence d'un choc relié à l'évolution d'une des variables sur les autres variables du système. La non-causalité de Granger va jouer un rôle important si elle est introduite dans le système car elle spécifie justement que les chocs associés à une variable n'ont possiblement pas d'influence sur d'autres variables.

Quelle est l'idée générale de l'analyse impulsionnelle (analyse des chocs) ?

**Idée Générale** Une fonction de réponse aux innovations résume l'information concernant l'évolution d'une composante  $x_{it}$  qui intervient suite à une impulsion sur  $x_{jt}$  à la date  $T$ , en supposant que toutes les autres variables sont constantes pour  $t \leq T$ .

**Exemple 12.** On considère un processus bi-varié  $\{Y_t, t \in \mathbb{Z}\}$  admettant une représentation VAR(1) telle que :

$$\begin{aligned} y_{1t} &= 3 + 0.2y_{1t-1} + 0.7y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= 1 + 0.3y_{1t-1} + 0.4y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

On pose  $Y_t = (y_{1t} \ y_{2t})'$  et  $\varepsilon_t = (\varepsilon_{1t} \ \varepsilon_{2t})'$ . On suppose que les chocs  $\varepsilon_{1t}$  et  $\varepsilon_{2t}$  sont corrélés. Cette hypothèse est particulièrement importante pour la suite de cet exemple.

$$E(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

On cherche à identifier l'impact d'un choc unitaire sur  $y_{2T}$  à la date  $T$  sur la dynamique de la variable  $y_{1t}$  aux périodes postérieures à  $T$ , en supposant les évolutions de ces deux variables pour  $t \leq T$  connues et données. Cela revient à supposer :

$$\varepsilon_{2T} = 1, \quad \varepsilon_{2t} = 0, \quad \forall t > T$$

On cherche donc à déterminer la variation de  $y_{1t}$  engendrée par ce choc. Pour cela considérons la décomposition de Wold du processus  $Y_t$  déterminée par :

$$y_{1t} = \mu_1 + \sum_{i=0}^{\infty} \psi_{1i} \varepsilon_t$$

où  $\psi_{1i}$  désigne la première ligne de la matrice  $\psi_i$  issue de la représentation VMA :

$$Y_t = \psi(L)\varepsilon_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

$$Y_t = \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} = \begin{pmatrix} 9.25 \\ 6.29 \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}^i \begin{pmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{pmatrix}$$

On pourrait penser que suite au choc  $\varepsilon_{2T} = 1$ , dans ce cas, la suite des réalisations  $y_{1T+h}$  soit donnée directement par les coefficients correspondants du vecteur  $\psi_{1i}$ . On obtiendrait ainsi une fonction de réponse de la variable  $y_1$  à une impulsion de la variable  $y_2$ . C'est une première possibilité de fonctions de réponse.

Le problème ici c'est qu'en raison de la corrélation des deux chocs, l'impulsion initiale sur  $\varepsilon_{2T}$  n'est pas sans influence sur l'innovation  $\varepsilon_{1T}$  qui entre elle aussi dans la représentation moyenne mobile infinie de  $y_{1t}$ . Conditionnellement à la réalisation de  $\varepsilon_{2T}$ , du fait de la corrélation des deux chocs, la probabilité d'engendrer une innovations  $\varepsilon_{1T}$  non nulle est-elle même non nulle.

Or généralement, ce qui est intéressant sur le plan économique c'est d'envisager une impulsion sur la composante orthogonale de  $\varepsilon_{2t}$  à  $\varepsilon_{1t}$ . C'est à dire, il convient d'isoler l'innovation "propre" au processus  $y_{2t}$  non "polluée" par la réaction de l'innovation  $y_{1t}$ . C'est pourquoi, il convient dans le cas général où  $E(\varepsilon_t \varepsilon_t') \neq I_n$ , d'orthogonaliser les innovations. On considère donc la décomposition suivante de la matrice de covariance des innovations :

$$ADA'$$

où  $A$  est une matrice (2,2) triangulaire inférieure et où  $D$  est une matrice diagonale. Dans notre exemple, on a :

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\sigma_{12}}{\sigma_1^2} & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2} \end{pmatrix} \begin{pmatrix} 1 & \frac{\sigma_{12}}{\sigma_1^2} \\ 0 & 1 \end{pmatrix}$$

D'où dans notre exemple :

$$A = \begin{pmatrix} 1 & 0 \\ 0.5 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix}$$

On pose :

$$v_t = A^{-1}\varepsilon_t$$



On remarque alors que les innovations  $v_t$  sont des combinaisons linéaires des innovations du modèle initial  $\varepsilon_t$  qui possèdent une propriété d'indépendance puisque :

$$\begin{aligned} E(v_t v_t') &= A^{-1} E(\varepsilon_t \varepsilon_t') (A^{-1})' \\ &= A^{-1} (A^{-1})' \\ &= A^{-1} A D A' (A')^{-1} \\ &= D \end{aligned}$$

Donc la matrice de variance covariance des innovations  $v_t$  est diagonale, ce qui prouve que ces innovations ne sont pas corrélées. Dans notre exemple, il est très simple de constater que cette orthogonalisation correspond à la projection linéaire des  $\varepsilon_{2t}$  sur les  $\varepsilon_{1t}$ . Le résidu  $v_{2t}$  correspond à la composante orthogonale des  $\varepsilon_{2t}$ .

$$v_{2t} = \varepsilon_{2t} - \frac{\sigma_{12}}{\sigma_1^2} \varepsilon_{1t} = \varepsilon_{2t} - \frac{\varepsilon_{1t}}{2}$$

$$E(v_{2t} \varepsilon_{1t}') = E(v_{2t} v_{1t}') = 0$$

Reprenons la décomposition de Wold associée à  $Y_t$  il vient :

$$Y_t = m + \psi(L)\varepsilon_t = m + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$$

Or on pose  $\varepsilon_t = Av_t$ , dès lors cette représentation VMA( $\infty$ ) peut se réécrire en fonction d'innovations  $v_t$  non corrélées.

$$Y_t = \tilde{\psi}(L)\varepsilon_t = m + \sum_{i=0}^{\infty} \tilde{\psi}_i v_{t-i} = m + \sum_{i=0}^{\infty} (\phi_1^i A) v_{t-i}$$

On obtient donc dans notre exemple :

$$Y_t = \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} = \begin{pmatrix} 9.25 \\ 6.29 \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} -0.2 & -0.7 \\ -0.3 & -0.4 \end{pmatrix}^i \begin{pmatrix} 1 & 0 \\ 0.5 & 1 \end{pmatrix} \begin{pmatrix} v_{1t-1} \\ v_{2t-1} \end{pmatrix}$$

De façon équivalente on peut réécrire le VAR en fonction des seules innovations orthogonales :

$$\varepsilon_t = Av_t \iff \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0.5 & 1 \end{pmatrix} = \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}$$

$$\begin{aligned}y_{1t} &= 3 + 0.2y_{1t-1} + 0.7y_{2t-1} + v_{1t} \\y_{2t} &= 1 + 0.3y_{1t-1} + 0.4y_{2t-1} + 0.5v_{1t} + v_{2t}\end{aligned}$$

Dès lors, on construit de la même façon la séquence des  $y_{1,T+h}$  obtenus conditionnellement à un choc unitaire sur la composante orthogonale  $v_{2T}$ . Cela revient à supposer :

$$v_{2T} = 1, \quad v_{2t} = 0, \quad \forall t > T$$

### Cas général

On cherche ainsi de façon générale à se ramener à une représentation où les innovations sont orthogonales.

**Proposition 2.10.1.** *Dans le cas général où  $E(\varepsilon_t \varepsilon_t') \neq I_n$ , on orthogonalise les innovations de la façon suivante. On pose :*

$$v_t = A^{-1} \varepsilon_t$$

où la matrice  $A$  est issue de l'orthogonalisation de :

$$ADA'$$

où  $A$  est une matrice  $(2,2)$  triangulaire inférieure et où  $D$  est une matrice diagonale. On cherche donc à récrire le système VAR non plus en fonction des innovations corrélés  $\varepsilon_t$ , mais en fonction des innovations orthogonales  $v_t$  qui satisfont :

$$E(v_t v_t') = D$$

**Proposition 2.10.2.** *Cette phase d'orthogonalisation implique toutefois que l'ordre dans lequel sont disposées les variables du VAR affecte l'analyse dynamique et en particulier l'analyse des fonctions de réponse.*

**Exemple 13.** *En effet, reprenons l'exemple précédent. On considère un processus bi-varié  $\{Y_t, t \in \mathbb{Z}\}$  admettant une représentation VAR(1) telle que :*

$$\begin{aligned}y_{1t} &= 3 + 0.2y_{1t-1} + 0.7y_{2t-1} + \varepsilon_{1t} \\y_{2t} &= 1 + 0.3y_{1t-1} + 0.4y_{2t-1} + \varepsilon_{2t}\end{aligned}$$

On suppose que les deux chocs  $\varepsilon_{1t}$  et  $\varepsilon_{2t}$  sont corrélés et ont des variances différentes.

$$\text{cov}(\varepsilon_{1t}\varepsilon_{2t}) = \frac{1}{2}\sigma_1^2 = 1\sigma_2^2 = 2$$

Il existe deux façons d'écrire le VAR, soit on pose  $Y_t = (y_{1t} \ y_{2t})'$ , soit l'on écrit  $Y_t = (y_{2t} \ , y_{1t})$ . Le choix de l'ordre des variables va dès lors conditionner notre schéma d'orthogonalisation :

1. **Cas 1** : on écrit  $Y_t = (y_{1t} \ , y_{2t})'$  et  $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ . Dès lors, on pose :

$$E(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 2 \end{pmatrix}$$

Dans ce cas, on a :

$$ADA' = \begin{pmatrix} 1 & 0 \\ 0.5 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1.75 \end{pmatrix} \begin{pmatrix} 1 & 0.5 \\ 0 & 1 \end{pmatrix}$$

Les innovations orthogonales sont donc définies par :

$$\begin{aligned} v_t = A^{-1}\varepsilon_t &\iff \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.5 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -0.5 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \\ &\iff \begin{cases} v_{1t} = \varepsilon_{1t} \\ v_{2t} = -\frac{1}{2}\varepsilon_{1t} + \varepsilon_{2t} \end{cases} \end{aligned}$$

Dès lors  $v_{2t}$  mesure la composante de  $\varepsilon_{2t}$  orthogonale à  $\varepsilon_{1t}$ .

2. **Cas 2** : on écrit  $Y_t = (y_{2t} \ y_{1t})'$  et  $\varepsilon_t = (\varepsilon_{2t} \ \varepsilon_{1t})'$ . Dès lors, on pose :

$$E(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_2^2 & \sigma_{12} \\ \sigma_{12} & \sigma_1^2 \end{pmatrix} = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

Dans ce cas, on a :

$$ADA' = \begin{pmatrix} 1 & 0 \\ 0.25 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \frac{7}{8} \end{pmatrix} \begin{pmatrix} 1 & 0.25 \\ 0 & 1 \end{pmatrix}$$

Les innovations orthogonales sont donc définies par :

$$v_t = A^{-1}\varepsilon_t \iff \begin{pmatrix} v_{2t} \\ v_{1t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.25 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{1t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -0.25 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{1t} \end{pmatrix}$$

$$\iff \begin{cases} v_{1t} = -\frac{1}{4}\varepsilon_{2t} + \varepsilon_{1t} \\ v_{2t} = \varepsilon_{2t}. \end{cases}$$

Dans ce cas,  $v_{2t}$  n'est rien d'autre que  $\varepsilon_{2t}$ , qui par nature est corrélé avec  $\varepsilon_{1t}$ .

3. En conséquence, dans cet exemple on montre que (i) si l'on désire étudier l'impact d'une innovation "pure" du processus  $y_{2t}$  sur le niveau de  $y_{1t}$ , et que (ii) on retient cette méthode d'orthogonalisation, il convient d'écrire le système VAR sous la forme  $Y_t = (y_{1t}, y_{2t})'$  : Bien entendu, les fonctions de réponse obtenues dans les deux cas ne sont pas identiques.

**Résultat** De façon générale dans l'écriture d'un VAR, la variable, que l'on suppose économiquement être la variable explicative, doit figurer après la variable dont on veut expliquer les évolutions.

## 2.10.2 Décomposition de la variance

La décomposition de la variance de l'erreur de prévision a pour objectif de calculer pour chacune des innovations sa contribution à la variance de l'erreur. Par une technique mathématique, on peut écrire la variance de l'erreur de prévision à un horizon  $h$  en fonction de la variance de l'erreur attribuée à chacune des variables ; il suffit ensuite de rapporter chacune de ces variances à la variance totale pour obtenir son poids relatif en pourcentage.

Reprenons notre modèle VAR(1) à deux variables  $x_{1t}$  et  $x_{2t}$ , la variance de l'erreur de prévision pour  $x_{1t+h}$  peut s'écrire :

$$\sigma_{x_1}^2(h) = \sigma_{\varepsilon_1}^2 [\psi_{11}^2(0) + \psi_{11}^2(1) + \dots + \psi_{11}^2(h-1)]$$

$$+ \sigma_{\varepsilon_2}^2 [\psi_{22}^2(0) + \psi_{22}^2(1) + \dots + \psi_{22}^2(h-1)]$$

où les  $\psi_{ii}$  sont les termes de la matrice  $\Psi$  de la Représentation VMA d'un processus VAR. À l'horizon  $h$ , la décomposition de la variance, en pourcentage, des propres innovations de  $x_{1t}$  sur  $x_{1t}$ , est donnée par :

$$\frac{\sigma_{\varepsilon_1}^2 [\psi_{11}^2(0) + \psi_{11}^2(1) + \dots + \psi_{11}^2(h-1)]}{\sigma_{x_1}^2(h)}.$$

Et la décomposition de la variance, en pourcentage, des innovations de  $x_{1t}$  sur  $x_{2t}$ , est donnée par :

$$\frac{\sigma_{\varepsilon_2}^2 [\psi_{22}^2(0) + \psi_{22}^2(1) + \dots + \psi_{22}^2(h-1)]}{\sigma_{x_1}^2(h)}$$

L'interprétation des résultats est importante :

- si un choc sur  $\varepsilon_{1t}$  n'affecte pas la variance de l'erreur de  $x_{2t}$  quel que soit l'horizon de prévision, alors  $x_{2t}$  peut être considéré comme exogène car  $x_{2t}$  évolue indépendamment de  $\varepsilon_{1t}$  ;
- si un choc sur  $\varepsilon_{1t}$  affecte fortement - voir totalement - la variance de l'erreur de  $x_{2t}$ , alors  $x_{2t}$  est considéré comme endogène.

Dans la pratique, les résultats ne sont pas aussi marqués mais indiquent la contribution de chacune des variables à la variance de l'erreur.

## 2.11 La cointégration et le modèle à correction d'erreur

### 2.11.1 Le concept de cointégration

L'analyse de la cointégration permet d'identifier clairement la relation véritable entre deux variables en recherchant l'existence d'un vecteur de cointégration et en éliminant son effet, le cas échéant.

#### Propriétés de l'ordre d'intégration d'une série

Une série est intégrée d'ordre  $d$  (notée  $x_t \rightarrow I(d)$ ), s'il convient de la différencier  $d$  fois afin de la stationnariser.

Soit une série  $x_{1t}$  stationnaire et une série  $x_{2t}$  intégrée d'ordre 1 :

$$\begin{array}{l} x_{1t} \rightarrow I(0) \\ x_{2t} \rightarrow I(1) \end{array} \Rightarrow x_{1t} + x_{2t} \rightarrow I(0)$$

La série  $y_t = x_{1t} + x_{2t}$  est non stationnaire puisque l'on somme une série affectée d'une tendance et une série stationnaire.

Soit deux séries  $x_{1t}$  et  $x_{2t}$  intégrées d'ordre  $d$  :

$$\begin{aligned} x_{1t} &\rightarrow I(d) \\ x_{2t} &\rightarrow I(d) \end{aligned} \Rightarrow x_{1t} + x_{2t} \rightarrow I(?)$$

La combinaison linéaire  $\alpha x_{1t} + \beta x_{2t} \rightarrow I(?)$ . En effet, le résultat dépend des signes des coefficients  $\alpha, \beta$  et de l'existence d'une dynamique non stationnaire commune. Examinons un autre cas de figure :

$$\begin{aligned} x_{1t} &\rightarrow I(d) \\ x_{2t} &\rightarrow I(d') \quad d' \neq d \end{aligned} \Rightarrow x_{1t} + x_{2t} \rightarrow I(?)$$

Il s'avère impossible de conclure car on somme deux séries d'ordre d'intégration différent.

Soit deux séries affectées d'une tendance :

- dans un premier cas, les deux séries ont une tendance d'évolution constante sur une première période et puis une tendance d'évolution divergente en une deuxième période, les séries ne sont pas cointégrées (cf. figure 2.1).
- dans un deuxième cas, les deux séries ont une évolution constante sur toute la période, les séries sont cointégrées (cf. figure 2.2), il existe une évolution de long terme constante entre les chroniques.

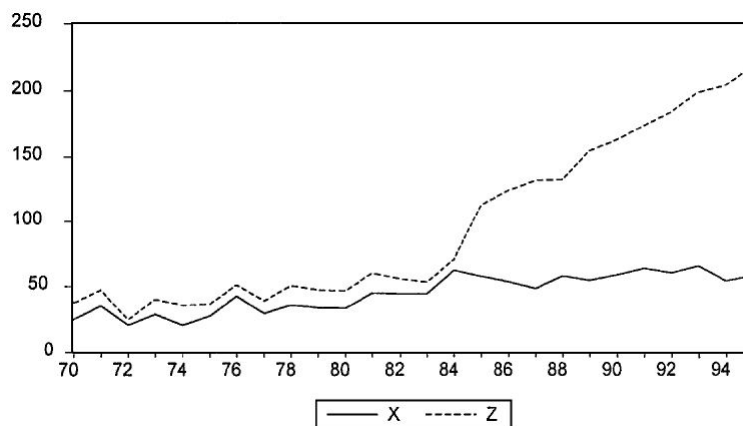


Figure 2.1 : Les variables  $x_t$  et  $z_t$  ne sont pas cointégrées.

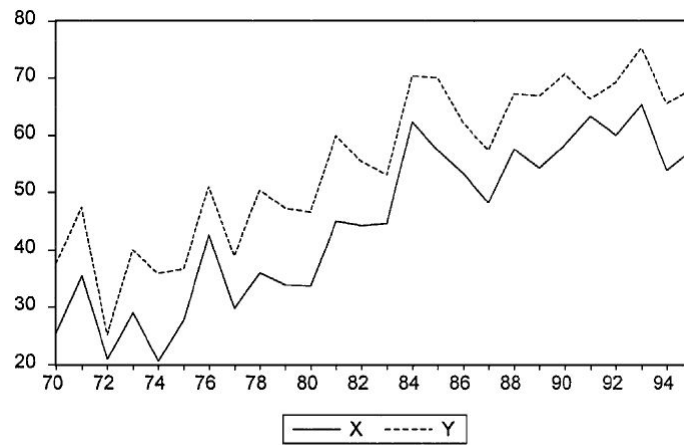


Figure 2.2 : Les variables  $x_t$  et  $y_t$  sont cointégrées.

### Conditions de cointégration

Deux séries  $x_t$  et  $y_t$  sont dites cointégrées si les deux conditions sont vérifiées :

- elles sont affectées d'une tendance stochastique de même ordre d'intégration  $d$ ,
- une combinaison linéaire de ces séries permet de se ramener à une série d'ordre d'intégration inférieur.

Soit :

$$x_t \rightarrow I(d)$$

$$y_t \rightarrow I(d)$$

tel que  $\alpha_1 x_t + \alpha_2 y_t \rightarrow I(d - b)$ , avec  $d \geq b > 0$ .

On note :  $x_t, y_t \rightarrow CI(d, b)$  où  $[\alpha_1 \ \alpha_2]$  est le vecteur de cointégration.

Dans le cas général à  $k$  variables, on a :

$$x_{1t} \rightarrow I(d)$$

$$x_{2t} \rightarrow I(d)$$

...

$$x_{kt} \rightarrow I(d)$$

on note  $X_t = [x_{1t} \ x_{2t} \ \dots \ x_{kt}]$

s'il existe un vecteur de cointégration  $a = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_k]$  de dimension  $(k, 1)$  tel que  $\alpha X_t \rightarrow I(d - b)$ , alors les  $k$  variables sont cointégrées et le vecteur de cointégration est  $a$ . On note que  $X_t \rightarrow CI(d, b)$  avec  $b > 0$ .

# Chapitre 3

## Simulations et Applications

La simulation, en général demande beaucoup d'attention, car les erreurs peuvent causer des problèmes du modèle, soit dans la réalisation du processus, soit dans l'estimation des paramètres du modèle. Ces erreurs sont liées d'une part avec le fait qu'on a réalisé en réalité des pseudo aléatoires-aléatoires et d'autre part avec les erreurs de calculs. Les simulations suivantes sont réalisées à l'aide de logiciel R 3.4.4 (logiciel libre disponible sur le site : <https://cran.r-project.org>) et en utilisant les packages `dse`, `vars` et `tseries`.

### 3.1 Simulation d'un modèle AR

#### 3.1.1 Simulation d'un processus AR(1)

Pour simuler un AR(1) ( $\phi = 0.7$ ), tracer son graphe, ses autocorrélations et ses autocorrélations partielles on utilise le code suivant :

• **Code R 1.1. Simulation du processus AR(1) avec  $\phi = 0.7$  :**

```
y=arima.sim(n=100,list(ar = 0.7 ),innov=rnorm (100) )
y
op=par(no.readonly=TRUE)
layout(matrix(c(1,1,2,3),2,2,byrow=TRUE))
plot.ts(y,ylab='')
acf(y,main='Autocorrelations',ylab='',ylim=c(-1,1)
    ,ci.col='black')
```



```
pacf(y,main='PartialAutocorrelations',ylab='',ylim=c(-1,1)
      ,ci.col='black')
```

Dans le code R 1.1, les fonctions utilisées sont contenues dans le package `stats`. Cependant, Il est à noter que les mêmes fonctions sont fournies dans le package `fArma` de Weurtz [2007a]. Ces fonctions incluent la simulation (`armaSim()`), l'estimation (`armaFit()`) et la prévision (`predict()`) ainsi que l'évaluation de la stabilité (`armaRoots()`) et le calcul de l'autocorrélation théorique et les fonctions d'autocorrélation partielle (`armaTrueacf()`).

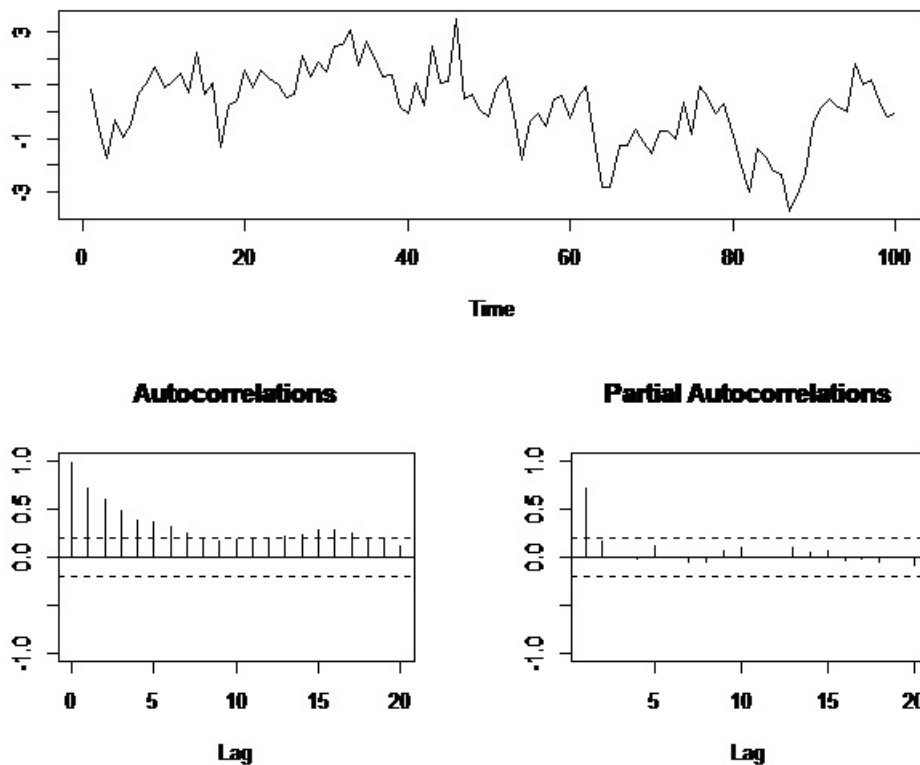


Figure 3.1 : Graphe des séries temporelles du processus AR(1),  $\phi = 0.7$ .

### 3.1.2 Simulation d'un processus AR(2)

La stabilité d'un tel processus peut facilement être vérifiée avec la fonction `polyroot()`. Dans le code R 1.2, ce processus AR(2) ( $y_t = 0.7y_{t-1} - 0.64y_{t-2} + \varepsilon_t$ ), est généré en utilisant la fonction `filter()` à la place de

`arima.sim()` comme dans le code

R 1.1. L'avantage d'utiliser `filter()` est que des processus AR(p) instables peuvent également être générés. Ensuite, le processus AR(2) généré est estimé avec la fonction `arima()`. Les estimations sont proches de leurs valeurs théoriques, comme on pouvait s'y attendre avec un échantillon de 1000. Les modules du polynôme caractéristique sont récupérés avec `Mod()` et les parties réelles et complexes avec les fonctions `Re()` et `Im()`, respectivement. Veuillez noter que les signes des coefficients estimés doivent être inversés pour le calcul des racines. Les racines peuvent être représentées dans un système de coordonnées cartésiennes avec un cercle d'unité, comme le montre la figure (3.3)

```

• Code R 1.2. Simulation d'un processus AR(2) avec  $\phi_1 = 0,7$  et  $\phi_2 = -0,64$ 
series= rnorm(1000)
y.st=filter(series,filter=c(0.7 , -0.64) ,method= 'recursive')
y.st
ar2.st=arima(y.st,c(2,0,0),include.mean=FALSE,transform.pars=FALSE
             ,method='ML')
ar2.st$coef
op=par(no.readonly=TRUE)
layout(matrix(c(1,1,2,3),2,2,byrow=TRUE))
plot(y.st)
acf(y.st,main='Autocorrelations',ylab='',ylim=c(-1,1)
    ,ci.col='black')
pacf(y.st,main='Partial Autocorrelations',ylab='',ylim=c(-1,1)
    ,ci.col='black')
polyroot(c(1,-ar2.st$coef))
Mod(polyroot(c(1,-ar2.st$coef)))
root.comp= Im(polyroot(c(1,-ar2.st$coef)))
root.real=Re(polyroot(c(1,-ar2.st$coef)))
‡ Tracer les racines dans un cercle d'unités
x=seq(-1,1,length = 1000)
y1= sqrt(1- x^2)
y2=-sqrt(1- x^2)

```

```

plot(c(x,x),c(y1,y2),xlab='Real part',ylab='Complex part',type='l',
      ,main='Unit Circle',ylim=c(-2,2),xlim=c(-2,2))
abline(h=0)
abline(v=0)
points(Re(polyroot(c(1,-ar2.st$coef))),
      Im(polyroot(c(1,-ar2.st$coef))),pch=19)
legend(-1.5,-1.5,legend='Roots of AR(2)', pch=19)

```

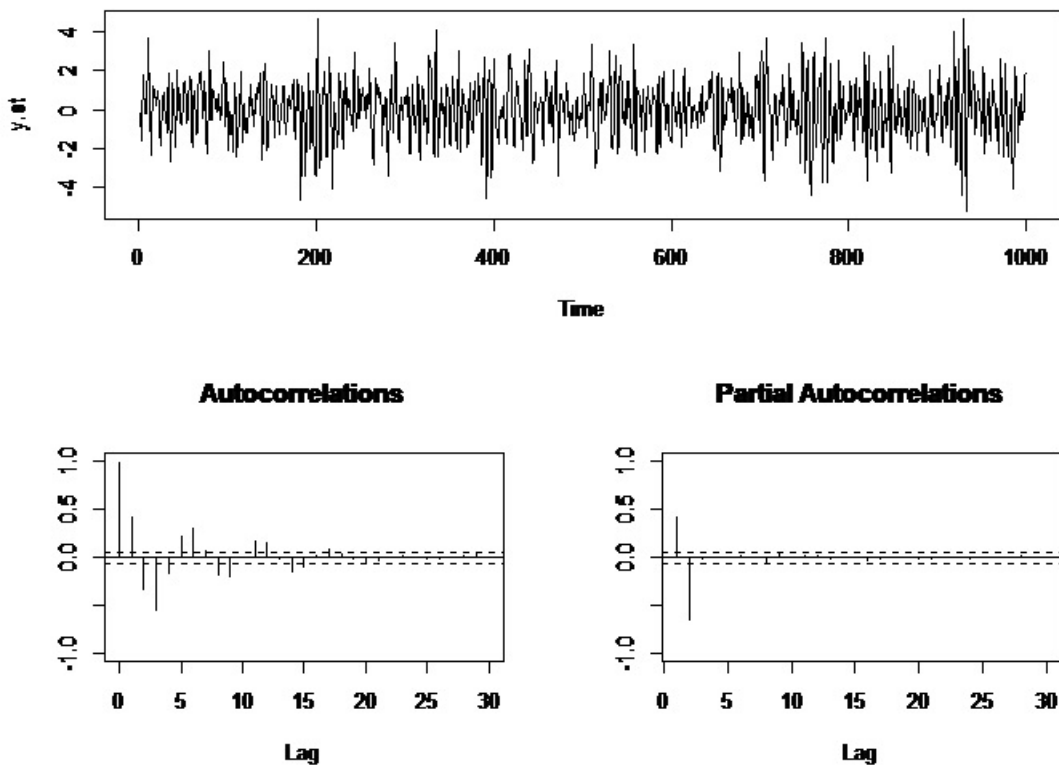


Figure 3.2 : Graphe des séries temporelles du processus AR(2),  $\phi_1 = 0.7$  et  $\phi_2 = -0.64$ .

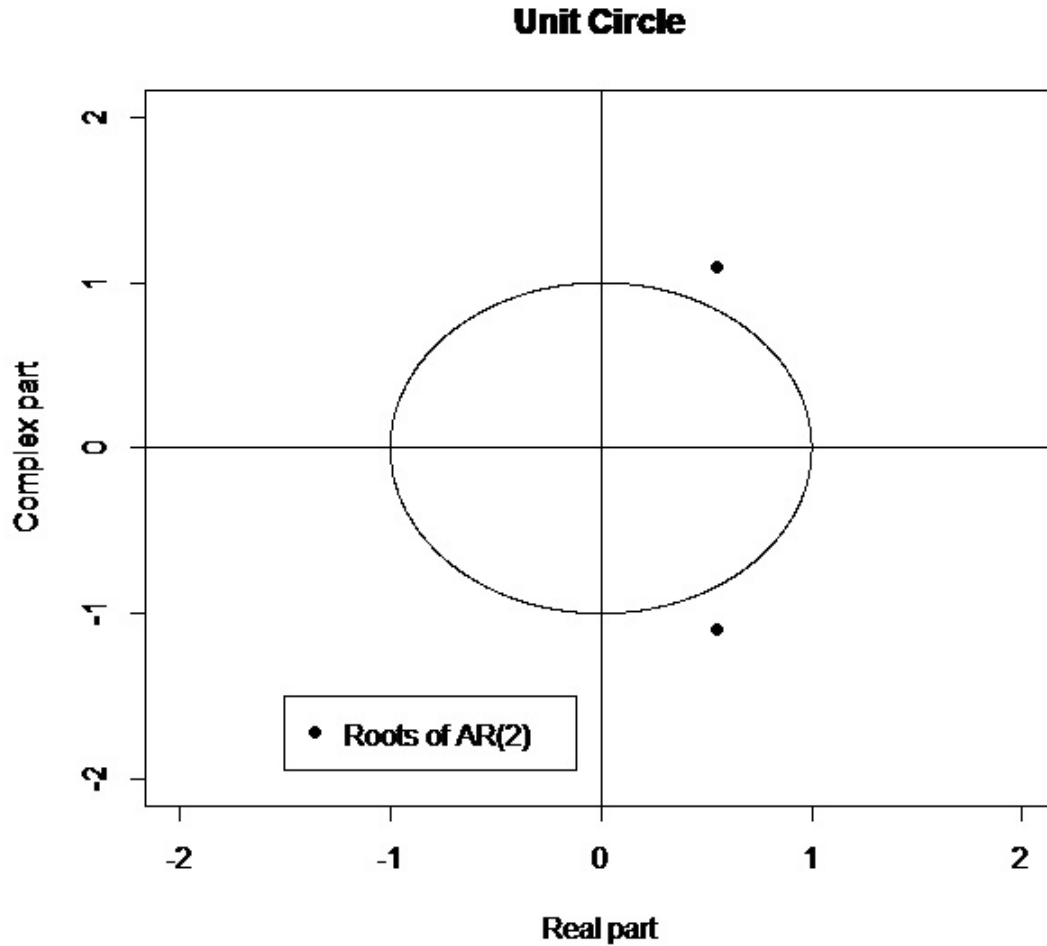


Figure 3.3 : Cercle unitaire et racines du processus AR(2) stable,  $\phi_1 = 0.7$  et  $\phi_2 = -0.64$ .

## 3.2 Un modèle AR pour lh

### 3.2.1 Choix d'un modèle

On s'intéresse maintenant au jeu de données `lh`. On trace sa trajectoire, son ACF et son PACF avec `myplot`. Peut-on utiliser un AR(1) ou un AR(3) pour fitter `lh` de façon simple ?

On peut utiliser la fonction `arima` pour fitter. La fonction `arima` estime

les paramètres du modèle par maximum de vraisemblance et/ou par moindres carrés (par défaut). Elle calcule également la valeur des critères AIC, AICc et BIC.

La fonction `auto.arima` du package `forecast` effectue la sélection du modèle selon les critères AIC, AICc et BIC.

```
X =lh
X
lh.ar1= arima(X, order = c(1, 0, 0))
lh.ar3= arima(X, order = c(3, 0, 0))
lh.auto= auto.arima(X)
```

Le critère AIC préfère le modèle AR(3) (de peu), par contre les critères AICc et BIC préfèrent le modèle AR(1). Pour comparer les deux fits, on peut utiliser la fonction `tsdiag` sur un fit obtenu avec la fonction `arima`.

```
tsdiag(lh.ar1)
tsdiag(lh.ar3)
```

Cela affiche les résidus du fit, l'ACF des résidus, et les p-values du test de Ljung-Box. On constate que le fit utilisant un modèle AR(3) semble meilleur qu'avec l'AR(1).

### 3.2.2 Prévision

En utilisant le fit `lh.ar3`, on peut prédire les prochaines valeurs de la série `lh`. On utilise pour cela la fonction `predict`.

```
lh.fore= predict(lh.ar3, 12)
lh.fore
ts.plot(lh, lh.fore$pred, lh.fore$pred + 1.96 * lh.fore$se,
+ lh.fore$pred -1.96 * lh.fore$se, gpars = list(lty =
+ c(1, 1, 2, 2), lwd = c(1,2, 1, 1)),main = 'lh prediction',
+ xlab = 'Time', ylab = 'lh')
```

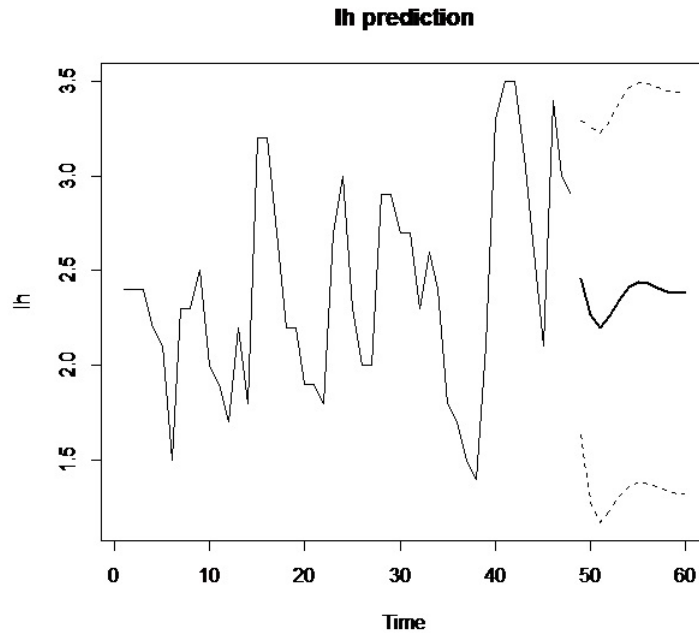


Figure 3.4 : Graphe de prévision du processus lh.

### 3.3 Simulation d'un modèle VAR

#### 3.3.1 Spécification, hypothèses et estimation

Pour les processus VAR on utilise les fonctions `VAR()` et `VARselect()` contenues dans le package `vars`. Dans la première fonction, un modèle VAR(p) approprié sera estimé en fournissant le nombre de retard maximal, `lag.max`, et le critère désiré. Le résultat de la fonction `VARselect()` est une liste d'objets le premier élément dit `selection` et le deuxième `criteria`.

Nous allons maintenant générer un processus VAR(2) artificiel bidimensionnel qui obéit à la forme suivante :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \begin{bmatrix} 5.0 \\ 10.0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1} + \begin{bmatrix} -0.3 & -0.7 \\ -0.1 & 0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}_t.$$

Le processus ci-dessus est simulé dans le code R 2.1. Ceci est réalisé en employant la fonction `ARMA()` et `simulate()`, contenue dans le package `dse1`.

Dans la première étape, le polynôme de retard  $\phi(L)$  décrit dans l'équation  $(\phi(L)X_t = \varepsilon_t)$  est créé comme un tableau par `Apoly`. La forme de la matrice de variance-covariance du processus d'erreur est une matrice d'identité stockée en tant qu'objet `B`, et finalement le terme constant est assigné comme `TRD`. Un objet `ARMA` est ensuite créé et le modèle est simulé pour un échantillon de 500 observations. Les séries résultantes sont extraites de la sortie de l'élément de `output` et représentées sur la Figure (3.5). À l'étape suivante, l'ordre de décalage est déterminé empiriquement en utilisant `VARselect()`. Alternativement, le modèle `VAR(p)` aurait pu être estimé directement en définissant `lag.max = 4` et `type = 'AIC'`. Tous les critères indiquent un ordre de décalage de deux. Enfin, un `VAR(2)` avec une constante est estimé avec la fonction `VAR()`, et on vérifie la stabilité de ses racines en appliquant la fonction `roots()` à l'objet `varsimest`. La fonction a un argument `'modulus'` de type logique qui retourne par défaut les modules des valeurs propres ; si non, un vecteur de nombres complexes est renvoyé.

Les résultats du `VAR(2)` pour les variables  $y_1$  et  $y_2$  sont présentés dans les tableaux (3.1) et (3.2), respectivement. Comme prévu, les coefficients estimés sont proches de leurs valeurs théoriques, et tous sont significativement différents de zéro. Enfin, les valeurs propres de la forme associée sont inférieures à un et sont fournis dans le tableau (3.3).

**Tableau 3.1.** *résultat VAR pour  $y_1$*

<i>Variable</i>	<i>estimation Std. Erreur</i>		<i>t - valeur Pr(&gt;  t )</i>	
<b>Niveaux décalés</b>				
<i>y1.l1</i>	<b>0.4998</b>	0.0354	14.1003	0e + 00
<i>y2.l1</i>	<b>0.1551</b>	0.0407	3.8085	2e - 04
<i>y1.l2</i>	<b>-0.3291</b>	0.0352	-9.3468	0e + 00
<i>y2.l2</i>	<b>-0.7550</b>	0.0454	-16.6466	0e + 00
<b>Déterministe</b>				
<i>const.</i>	<b>5.9196</b>	0.6197	9.5531	0e + 00

Tableau 3.2. résultat VAR pour  $y_2$ 

<i>Variable</i>	<i>estimation Std. Erreur</i>		<i>t – valeur Pr(&gt;  t )</i>	
<b>Niveaux décalés</b>				
<i>y1.l1</i>	-0.1499	0.0358	-4.1920	0e + 00
<i>y2.l1</i>	-0.4740	0.0411	-11.5360	0e + 00
<i>y1.l2</i>	-0.1184	0.0355	-3.3328	9e - 04
<i>y2.l2</i>	0.3006	0.0458	6.5684	0e + 00
<b>Déterministe</b>				
<i>const.</i>	9.7620	0.6253	15.6124	0e + 00

Tableau 3.3. Valeurs propres de la forme associée

	1	2	3	4
<i>Valeurs propres</i>	0.8311	0.6121	0.6121	0.6049

• **Code R 2.1. Simulation d'un modèle VAR(2)**

```
Apoly= array(c(1.0,-0.5,0.3,0,0.2,0.1,0,-0.2,0.7,1,0.5,-0.3)
            ,c(3,2,2))
```

```
Apoly
```

```
‡ Définir la covariance à l'identité -matrix
```

```
B = diag (2)
```

```
‡ Réglage du terme constant à 5 et 10
```

```
TRD = c(5,10)
```

```
TRD
```

```
‡ Générer le modèle VAR(2)
```

```
var2=ARMA(A=Apoly,B = B,TREND = TRD)
```

```
‡ Simulation de 500 observations
```

```
varsim= simulate(var2,sampleT = 500,
```

```
    noise = list(w = matrix(rnorm(1000)
```

```
,nrow = 500,ncol=2)),rng =list( seed = c(123456) ) )
```

```
‡ Obtention des séries générées
```

```
vardat=matrix(varsim$output,nrow = 500,ncol=2)
```

```
colnames(vardat)=c( 'y1 ', 'y2')
```

```
‡ Plot de serie
```



```
plot.ts( vardat,main ='', xlab ='')
# Déterminer un décalage approprié
infocrit=VARselect(vardat,lag.max = 3,type = 'const')
# Estimer le modèle
varsimest= VAR(vardat,p = 2,type = 'const',season = NULL,
  exogen = NULL)
# Alternativement, la sélection selon AIC
varsimest= VAR(vardat,type = 'const',lag .max = 3,ic = 'SC')
# Vérification des racines
roots=roots(varsimest)
```

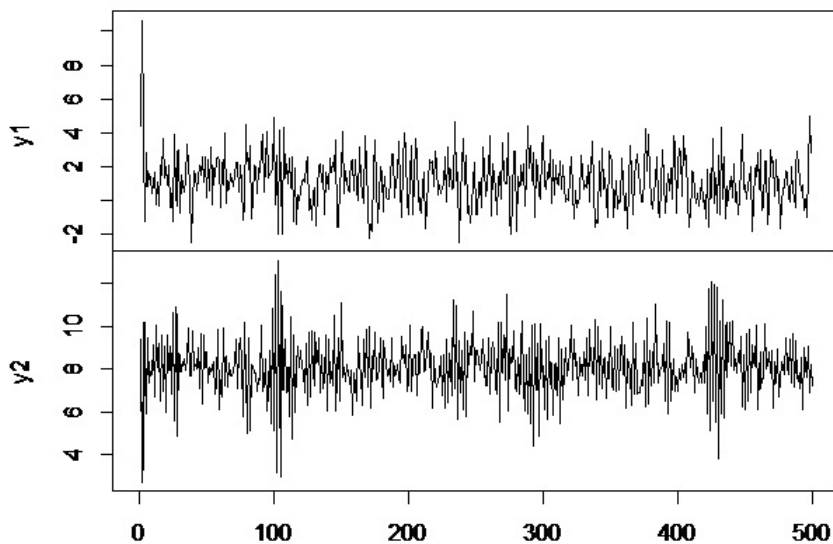


Figure 3.5 : graphes des séries chronologiques de la simulation du processus VAR(2) .

### 3.3.2 Tests de diagnostic

Pour tester la corrélation on utilise la fonction `serial.test()`. Et pour la normalité on utilise la fonction `normality.test`.

**• Code R 2.2. Tests de diagnostic du processus VAR(2)**

```
‡ tester la corrélation
args(serial.test)
‡ Portmanteau-Test
var2c.serial= serial.test(varsimest,lags.pt =16,type = 'PT. asymptotic')
var2c.serial
plot(var2c.serial,names = 'y1')
plot(var2c.serial,names = 'y2')
‡ tester la normalité
args(normality.test)
var2c.norm = normality.test(varsimest,multivariate.only = TRUE)
var2c.norm
```

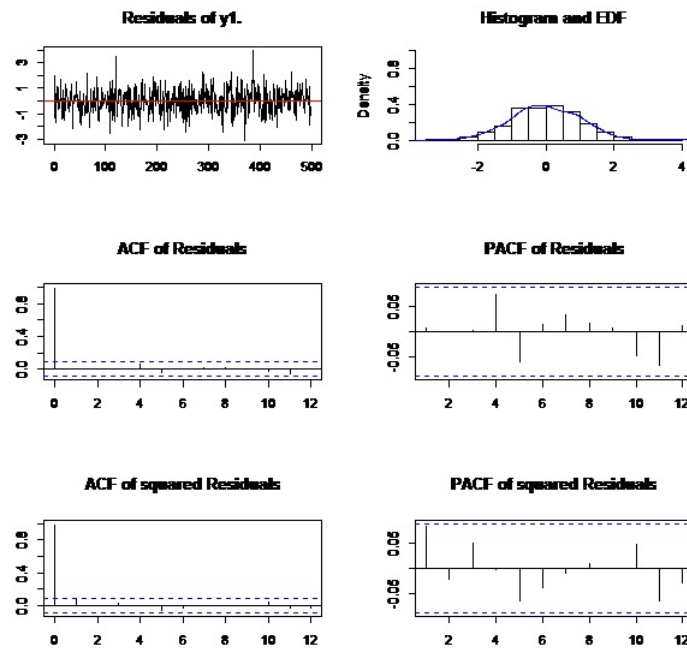


Figure 3.6 : diagnostic graphique des résidus de y1.

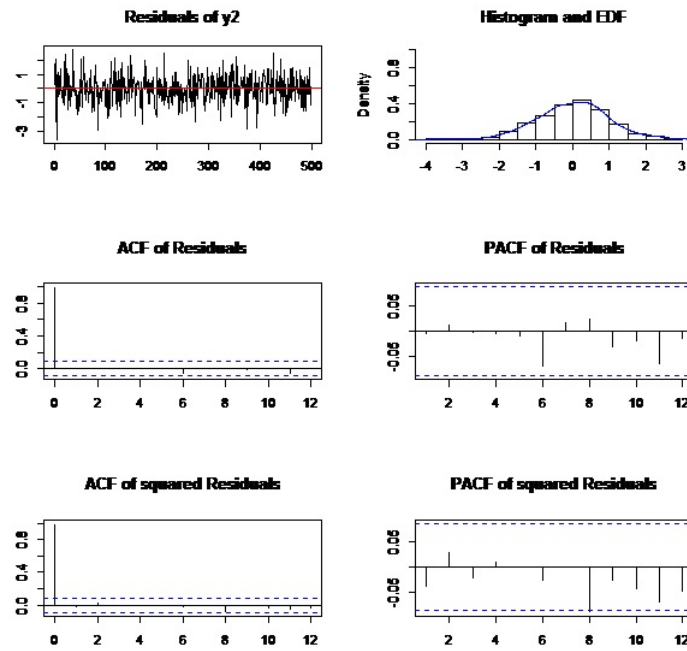
Figure 3.7 : diagnostic graphique des résidus de  $y_2$ .

Tableau 3.4. Tests de diagnostics du VAR(2)

<i>Test</i>	<i>Statistique</i>	<i>D.F.</i>	<i>p - value</i>
<i>Portmanteau</i>	52.44	56	0.61
<i>ARCHVAR</i>	32.58	45	0.92
<i>JBVAR</i>	0.54	4	0.97
<i>Kurtosis</i>	0.42	2	0.81
<i>kewness</i>	0.12	2	0.94

### 3.3.3 Prévisions du processus VAR

Dans le package `vars`, la prévision des processus VAR est effectuée par `prédicte`. Autre argument de cette fonction est `n.ahead`, Une valeur pour l'intervalle de confiance de prévision peut être également données. Sa valeur par défaut est 0,95. La méthode `predict` renvoie un objet list de la classe `varprd` avec trois éléments. Le premier élément, `fcst`, est une liste de matrices contenant les valeurs prédites, les limites inférieure et supérieure en

fonction de l'intervalle de confiance choisi, `ci`, et de sa taille. Le deuxième élément, `endog`, est un objet matriciel contenant les variables endogènes, et le troisième est l'objet `varest`. Une méthode pour tracer les objets de la classe `varprd` existe grâce à la fonction `fanchart()`

• **Code R 2.4. Prévisions du processus VAR**

```
# Prédiction d'objets de classe varest
args(vars:::predict.varest )
predictions = predict(varsimest,n.ahead =2 ,ci = 0.95)
class(predictions)
args(vars:::plot.varpr )
# Plot de prévisions pour y1
plot(predictions,names = 'y1 ')
# Fanchart pour y2 10args(fanchart)
fanchart(predictions,names = 'y2')
```

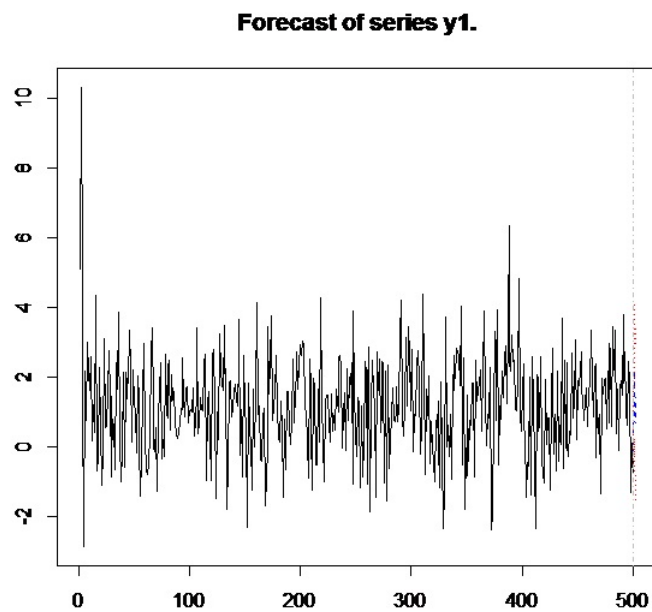


Figure 3.8 : Prédiction de `y1`.

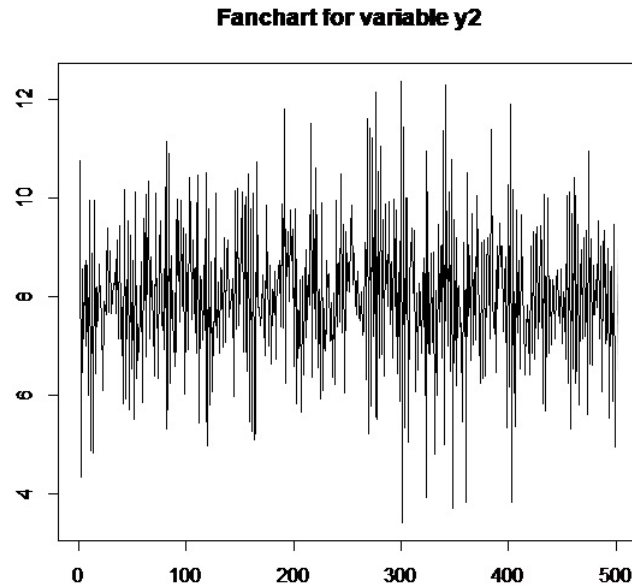


Figure 3.9 : prévision de  $y_2$  .

### 3.3.4 Analyse de causalité

Les deux tests de causalité test de Granger et test instantané sont implémentés dans la fonction `causality()` contenue dans le package `vars`. La fonction a deux arguments. Le premier argument, `x`, est un objet de classe `varest`, et le second, `cause`, est un vecteur de caractères des noms de variables, qui sont supposés être causaux aux variables restantes dans un processus VAR(p). Si cet argument n'est pas donnée, la variable de la première colonne de `x$y` est utilisée comme variable cause et un avertissement est imprimé. Dans le code R 2.3, cette fonction est appliquée au processus VAR(2) simulé. Les résultats sont fournis dans le tableau 3.5. Il est clair que l'hypothèse nulle de la non-causalité de Granger doit être écartée, alors que l'hypothèse d'absence de causalité instantanée ne peut pas être rejetée

- **Code R 2.3. Analyse de causalité du processus VAR(2)**

- ‡ Tests de causalité

- ‡ Granger et causalité instantanée

```
var.causal=causality( varsimest, cause = 'y2')
```

**Tableau 3.5.** *Tests de causalité*

<i>Test</i>	<i>Statistique</i>	<i>p – value</i>
<i>Granger</i>	250.07	0.00
<i>Instant</i>	0.00	0.99

### 3.3.5 Fonctions de réponse impulsionnelle (analyse du choc)

Deux tests de causalité ont été introduits, ce qui est très utile pour déduire si une variable aide à en prédire une autre. Cependant, cette analyse ne permet pas de quantifier l'impact de la variable d'impulsion sur la variable réponse au fil du temps. L'analyse de réponse impulsionnelle est utilisée pour étudier ces types d'interactions dynamiques entre les variables endogènes et est basé sur la représentation moyenne mobile (Wold) d'un processus VAR(p).

Dans le code R 2.5, une analyse de réponse impulsionnelle est effectuée pour un processus VAR(2). Pour plus de clarté, les réponses impulsionnelles de y1 à y2 et vice versa ont été divisées en deux lignes de commande distinctes. Les résultats sont présentés aux figures (3.10) et (3.11), respectivement.

#### • Code R 2.5. IRA du procédé VAR

```
‡ Analyse de réponse impulsionnelle
irf.y1=irf(varsimest,impulse ='y1',response='y2',n.ahead=10,
  ortho=FALSE,cumulative=FALSE,boot=FALSE,seed=12345)
args(vars:::plot.varirf)
plot(irf.y1)
irf.y2=irf(varsimest,impulse ='y2',response='y1',n.ahead=10,
  ortho=TRUE,cumulative=TRUE,boot=FALSE,seed=12345)
plot(irf.y2)
```

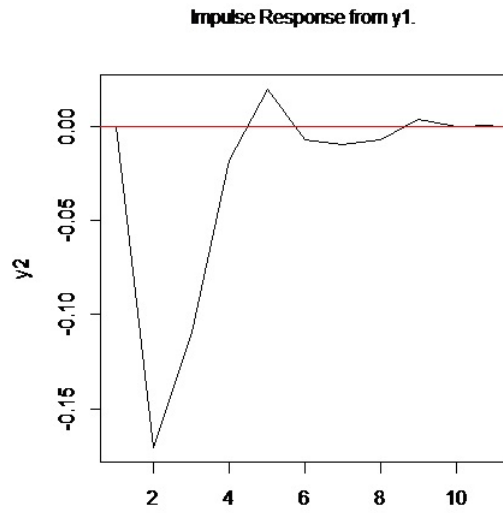


Figure 3.10 : Réponses d'impulsion de y1 à y2.

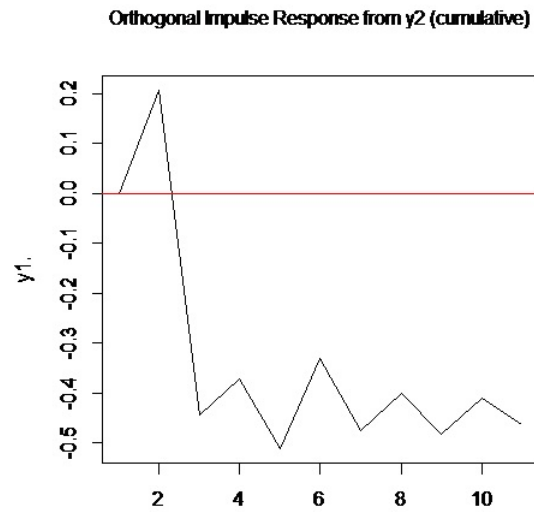


Figure 3.11 : Réponses d'impulsion de y2 à y1.

### 3.3.6 Décomposition de la variance des erreurs de la prévision (FEVD)

La méthode `fevd` dans le package `vars` est disponible pour la décomposition de la variance des erreurs de la prévision (FEVD). L'argument `n.ahead` définit le nombre d'étapes de prévision, un FEVD est appliqué au processus VAR(2) simulé, et sa sortie graphique est présentée à la figure (3.12).

• **Code R 2.6. FEVD du processus VAR**

```

‡ Décomposition de la variance des erreurs de la prévision (FEVD)
fevd.var2=fevd(varsimest,n.ahead=10)
args(vars:::plot.varfevd)
plot(fevd.var2,addbars=2)

```

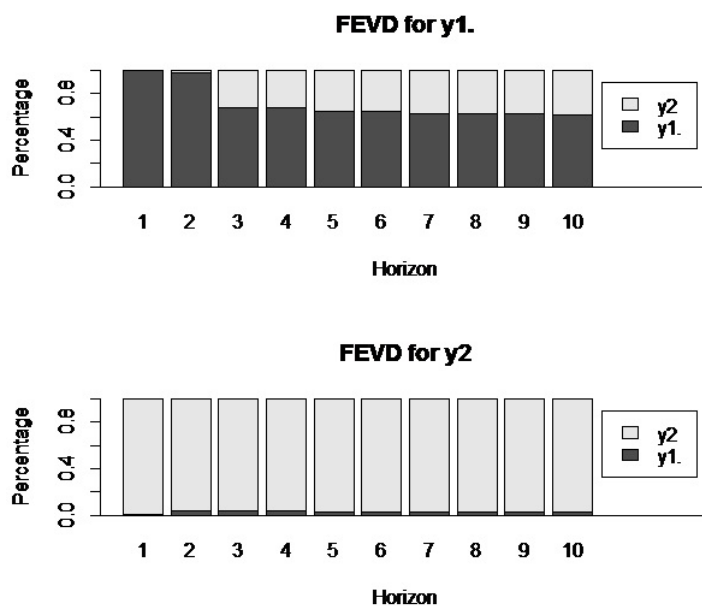


Figure 3.12 : FEVD pour le processus VAR(2)



## 3.4 Applications sur des données réelles

### 3.4.1 Méthode d'estimation :

L'estimation des paramètres du modèle de la représentation VAR a été faite à travers plusieurs tests et dans un ordre bien déterminé.

1) Test de racine unitaire : Test adf.test pour l'investigation des propriétés stochastiques des séries considérées dans le modèle en analysant leur ordre d'intégration.

2) Test de cointégration de Johansen afin de détecter les relations de cointegrations existantes entre les variables du modèle.

En cas d'absence de relation de cointégration, on passe à estimer le modèle VAR, sinon on se trouve dans l'obligation de déterminer un modèle autorégressif à correction d'erreur (VECM). (Voir [4])

3) Estimation du modèle VAR : dans notre étude, on va essayer de présenter l'estimation du modèle VAR et de trouver les paramètres.

4) Test de causalité de Granger : afin de voir l'effet de causalité des différentes variables .

5) Les fonctions de réponse impulsionnelle : pour étudier l'impact des chocs des variables. La fonction de réponse impulsion est une fonction dynamique permettant de voir la propagation du choc dans le temps, c'est-à-dire le temps nécessaire pour observer l'effet du choc dans la période contemporaine.

6) Estimation de la décomposition de la variance orthogonale : l'erreur anticipée de la décomposition de la variance renvoie à la proportion de la succession des mouvements due au propre choc ou contre choc d'une autre variable.

### 3.4.2 Définitions

#### **Indice des prix à la consommation (croissance annuelle en %)**

Ou taux d'inflation. La croissance de l'indice des prix à la consommation (IPC) mesure la variation annuelle en pourcentage de ce dernier afin de déterminer le taux d'inflation. Ce taux reflète le changement subi par les prix payés par le consommateur moyen pendant une période donnée lors de ses

### 3.4.3 Application aux données macroéconomiques canadiennes 82

---

achats de biens et services. Évidemment, le panier de biens et services sur lequel on base les calculs change avec le temps dû aux changements des habitudes de consommation. Une croissance de l'IPC, sans croissance des revenus, conduit à une diminution du pouvoir d'achat. Inversement, une diminution de l'IPC ou son maintien, alors que les revenus sont à la hausse conduit à une augmentation du pouvoir d'achat des ménages.

#### **Croissance annuelle du PIB (%)**

La croissance annuelle du produit intérieur brut (PIB) en % représente la variation relative du volume du PIB en dollars constants entre deux années. Elle reflète l'augmentation (ou la baisse dans le cas d'une croissance négative) du niveau d'activité économique dans un pays. Il s'agit d'un indicateur souvent retenu lorsque l'on veut faire des prévisions à court et à moyen terme sur la situation économique d'un pays.

#### **Taux d'intérêt réel (%)**

taux d'intérêt nominal ou taux facial est le taux qui est fixé lors de l'opération d'emprunt ou de prêt. Il est inscrit dans le contrat qui lie emprunteur et prêteur et sert à calculer les intérêts dus. Pour une obligation, c'est le taux d'intérêt défini lors de son émission. Il sert à déterminer les intérêts versés lors du détachement du coupon.

Le taux d'intérêt réel est le taux nominal après déduction du taux de l'inflation. Cette définition n'est valable que si le taux d'inflation est faible, sinon il faut utiliser l'équation :

$$1 + \text{taux d'intérêt nominal} = (1 + \text{taux d'intérêt réel}) \times (1 + \text{taux d'inflation})$$

### 3.4.3 Application aux données macroéconomiques canadiennes

Les séries chronologiques originales sont publiées par l'OCDE. La gamme de l'échantillon est du 1er trimestre 1980 au 4ème trimestre 2000. Les séries suivantes ont été utilisées dans la construction de la série prévue à Canada :

- U : Mesure du chômage canadien.

### 3.4.3 Application aux données macroéconomiques canadiennes 83

- rw : Mesure de Salaire réel dans le secteur manufacturier.
- e : Mesure de l'emploi.
- prod : Mesure de production intérieur brute.

La méthodologie utilisée :

● Code R 3.1. Analyse préliminaire des données

```
library(vars)
data(Canada)
Canada
donnees=window(Canada,start=1980,end=c(1998,4),freq=4)
summary(Canada)
plot(Canada,nc =2)
```

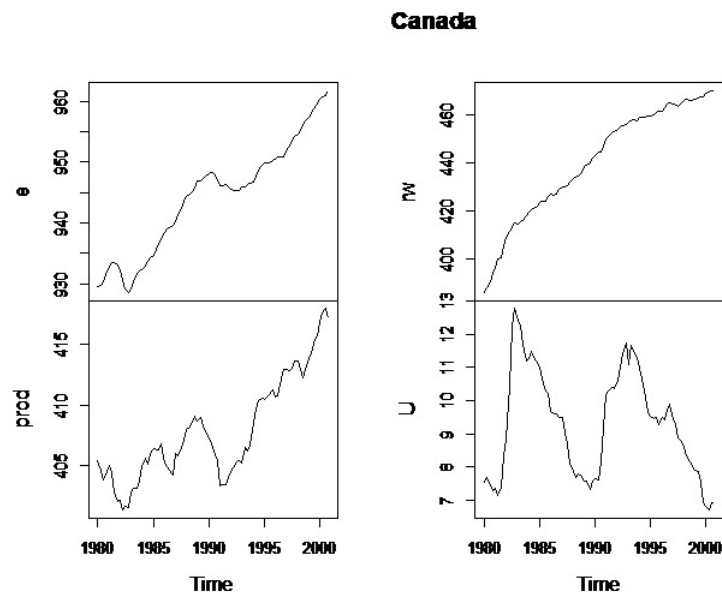


Figure 3.13 : Ensemble de données canadiennes : plot de séries temporelles .

Une analyse préliminaire des graphique (figure 3.13) des quatre séries concernées montre une non stationnarité de ces dernières. On procède alors à des tests de stationnarité

#### Test de stationnarité

il convient de s'assurer de la stationnarité des variables retenues car la stationnarité constitue une condition nécessaire pour éviter les relations factices. Les tests de racine unitaire permettent de mettre en évidence le caractère stationnaire ou non d'une chronique par la détermination d'une tendance déterministe ou stochastique.

Il existe deux types de processus non stationnaires : Les processus TS (Trend Stationary Process) qui présentent une non stationnarité de type déterministe et les processus DS (Difference Stationary Process) pour lesquels la stationnarité est de type aléatoire. Ces processus sont respectivement stationnarisés par écart à la tendance et par un filtre aux différences permet de déterminer l'ordre d'intégration de la variable. Afin de discriminer entre les deux types de processus et d'appliquer la méthode de stationnarité adéquate, nous utilisons le test ADF. Ce test permet de déterminer l'ordre de différenciation d'une série chronologique suivant son évolution au cours du temps.

#### • Code R 3.2. : Ensemble de données canadiennes : Analyse préliminaire

```
summary(ur.df(Canada[, 'prod'], type='trend', lags=2))
summary(ur.df(Canada[, 'e'], type='trend', lags=2))
summary(ur.df(Canada[, 'U'], type='trend', lags=1))
summary(ur.df(Canada[, 'rw'], type='trend', lags=4))
summary(ur.df(diff(Canada[, 'prod'], type='trend', lags=1)))
summary(ur.df(diff(Canada[, 'e'], type='trend', lags=1)))
summary(ur.df(diff(Canada[, 'U'], type='trend', lags=0)))
summary(ur.df(diff(Canada[, 'rw'], type='trend', lags=3)))
summary(ur.df(diff(diff(Canada[, 'rw'], type='trend', lags=0))))
```

**Tableau 3.6.** Tests ADF pour les données canadiennes

Variable	Termes de retard	Valeur de test	<i>Valeurs critiques</i>			
			1%	5%	10%	
prod	constant, trend	2	-1.99	-4.04	-3.45	-3.15
$\Delta$ prod	constant	1	-5.16	-3.51	-2.89	-2.58
e	constant, trend	2	-1.91	-4.04	-3.45	-3.15
$\Delta$ e	constant	1	-4.51	-3.51	-2.89	-2.58
U	constant	1	-2.22	-3.51	-2.89	-2.58
$\Delta$ U		0	-4.75	-2.6	-1.95	-1.61
rw	constant, trend	4	-2.06	-4.04	-3.45	-3.15
$\Delta$ rw	constant	3	-2.62	-3.51	-2.89	-2.58
$\Delta$ rw	constant	0	-5.6	-3.51	-2.89	-2.58

Il s'agit de tester l'hypothèse nulle de non stationnarité  $H_0$ , si cette dernière est acceptée, le processus admet une racine unitaire. Sinon, le processus est stationnaire

Il ressort de ce tableau que la valeur du test ADF est inférieure à la valeur critique à 5%. Les données canada sont non stationnaires, c'est pourquoi on a passé à la différence première.

En fin, la valeur de la statistique du test ADF est supérieure à la valeur critique à 5%.

Les différentes séries du modèle exhibent le même ordre d'intégration, il est donc licite de chercher d'éventuelles relations de cointegration. (Voir [2])

### Sélection de l'ordre de décalage

La première étape de cette analyse, consiste à déterminer le nombre de retards du modèle VAR(p). Pour ce faire, nous estimons un certain nombre de processus autorégressifs et nous retenons celui qui minimise conjointement les critères d'Akaike et Schwartz.

les auteurs ont déterminé une longueur de décalage optimale pour un VAR illimité pour une longueur de retard maximale de 8. Cela peut être accompli rapidement avec la fonction `VARselect()`, comme dans le code R 3.3. Les résultats sont rapportés dans le tableau 3.7.

• **Code R 3.3. Canada VAR : Sélection de l'ordre de décalage**

```
VARselect(Canada, lag.max =8, type='both')
```

**Tableau 3.7.** Canada VAR : sélection de l'ordre de décalage

<i>Lag Order</i>	<i>AIC(n)</i>	<i>HQ(n)</i>	<i>SC(n)</i>	<i>FPE(n)</i>
<i>p</i> = 1	-6.2726	-5.9784	-5.5366	0.0019
<i>p</i> = 2	-6.6367	-6.1464	-5.4100	0.0013
<i>p</i> = 3	-6.7712	-6.0848	-5.0538	0.0012
<i>p</i> = 4	-6.6346	-5.7522	-4.4265	0.0014
<i>p</i> = 5	-6.3981	-5.3196	-3.6994	0.0018
<i>p</i> = 6	-6.3077	-5.0331	-3.1183	0.0020
<i>p</i> = 7	-6.0707	-4.6000	-2.3906	0.0028
<i>p</i> = 8	-6.0616	-4.3947	-1.8908	0.0031

Selon l'AIC et FPE, le nombre de retard optimal est  $p = 3$ , tandis que le critère HQ indique  $p = 2$  et le critère SC indique une longueur de retard de  $p = 1$ .

**Estimation du modèle VAR**

Après avoir stationnariser les variables, ce qui nous intéresse en fait, c'est l'évaluation des mécanismes de propagation des chocs sur les différents variables de la Canada. Pour cela nous allons commencer par présenter les moyens d'analyse choisis pour évaluer les mécanismes de transmission des chocs avant de passer à l'analyse des résultats empiriques obtenus pour ce pays. Alors pour estimer ces modèles on utilise le code 3.4

• **Code R 3.4. Estimer le modèle**

```
varsimest=VAR(Canada,type='const',ic='SC')
varsimest
```

VAR Estimation Results:

=====

Estimated coefficients for equation prod:

=====

### 3.4.3 Application aux données macroéconomiques canadiennes 87

Call:

prod = prod.l1 + e.l1 + U.l1 + rw.l1 + const

prod.l1	e.l1	U.l1	rw.l1	const
1.00580112	-0.08012537	0.02249361	0.03093849	59.50401304

Estimated coefficients for equation e:

=====

Call:

e = prod.l1 + e.l1 + U.l1 + rw.l1 + const

prod.l1	e.l1	U.l1	rw.l1	const
0.1737566	1.3623097	0.8764326	-0.1454582	-356.5398351

Estimated coefficients for equation U:

=====

Call:

U = prod.l1 + e.l1 + U.l1 + rw.l1 + const

prod.l1	e.l1	U.l1	rw.l1	const
-0.1072831	-0.3686967	0.1483751	0.1371228	339.2913904

Estimated coefficients for equation rw:

=====

Call:

rw = prod.l1 + e.l1 + U.l1 + rw.l1 + const

prod.l1	e.l1	U.l1	rw.l1	const
-0.13471227	0.06531416	-0.18689438	0.96790330	10.14832207

On a par exemple :

$prod_t = 1.00580112 prod_{t-1} - 0.08012537 e_{t-1} + 0.02249361 U_{t-1} + 0.03093849 rw_{t-1} + 59.50401304$

### 3.4.3 Application aux données macroéconomiques canadiennes 88

---

#### Test de diagnostic

Pour les test de diagnostic utilise le code 3.5

- **Code R 3.5. Test de diagnostic pour les données canadiennes**

```
p1ct=VAR(Canada,p=1,type ='both')
# Serial
serial.test(p1ct,lags.pt=16,type='PT.asymptotic')
```

Portmanteau Test (asymptotic)

```
data: Residuals of VAR object p1ct
Chi-squared = 234.08, df = 240, p-value = 0.5956
```

```
serial.test(p1ct,lags.pt=16,type='PT.adjusted')
```

Portmanteau Test (adjusted)

```
data: Residuals of VAR object p1ct
Chi-squared = 261.38, df = 240, p-value = 0.1639
```

```
# JB
normality.test(p1ct)
```

```
$JB
```

JB-Test (multivariate)

```
data: Residuals of VAR object p1ct
Chi-squared = 5.6297, df = 8, p-value = 0.6886
```

```
$Skewness
```

Skewness only (multivariate)



### 3.4.3 Application aux données macroéconomiques canadiennes 89

---

```
data: Residuals of VAR object plct
Chi-squared = 3.209, df = 4, p-value = 0.5235
```

```
$Kurtosis
```

```
      Kurtosis only (multivariate)
```

```
data: Residuals of VAR object plct
Chi-squared = 2.4207, df = 4, p-value = 0.6589
```

La règle de décision pour le test de normalité si pvalue supérieur à 0.05 c'est une loi normale.

La règle de décision pour le test portemanteau si pvalue supérieur à 0.05 c'est une un bruit blanc.

#### Tests de causalité

Nous nous proposons d'illustrer la notion de causalité au sens de Granger en procédant à un test de non causalité. Les résultats obtenus, pour un nombre de retards  $p$  égal à 1, sont donnés dans le tableau 3.8

- **Code R 3.6. Tests de causalité.**

```
‡ Granger et causalité instantanée
```

```
var.causal=causality(varsimest,cause=c('rw','prod')
var.causal=causality(varsimest,cause=c('prod','e')
var.causal=causality(varsimest,cause=c('prod','U')
var.causal=causality(varsimest,cause=c('U','e')
var.causal=causality(varsimest,cause=c('rw','e')
var.causal=causality(varsimest,cause=c('U','rw')
```

### 3.4.3 Application aux données macroéconomiques canadiennes 90

**Tableau 3.8.** test de causalité

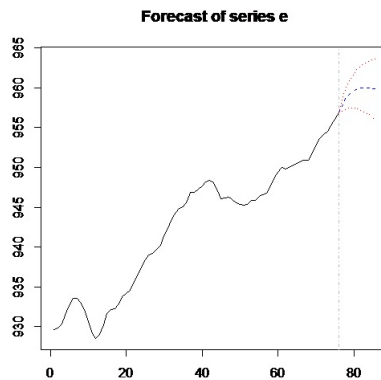
Hypothèse nulle :	F-Test	Probabilité
prod rw ne cause pas e U	17.191	<b>1.27e-12</b>
prod e ne cause pas rw U	16.365	<b>4.64e-12</b>
prod U ne cause pas e rw	14.194	<b>1.463e-10</b>
U e ne cause pas prod rw	6.0691	<b>0.0001071</b>
rw e ne cause pas prod U	11.105	<b>2.247e-08</b>
U rw ne cause pas prod e	8.3711	<b>2.161e-06</b>

Notons que les valeurs indiquées en gras sont significatives au seuil de 10%. L'hypothèse nulle d'absence de causalité de Granger du salaire réel et de la productivité du travail à l'emploi et au chômage doit être rejetée; alors que l'hypothèse nulle de causalité non-instantanée ne peut pas être rejetée. Ce résultat est économiquement plausible, compte tenu des frictions observées sur les marchés du travail.

#### Prévision

- **Code R 3.7.** Prévision d'objets de classe varest

```
p= predict(varsimest,n.ahead =8)
plot(p,names ='e')
```



**Figure 3.14 :** Emploi : Prévisions à 10 pas avec un niveau de confiance à 95%.

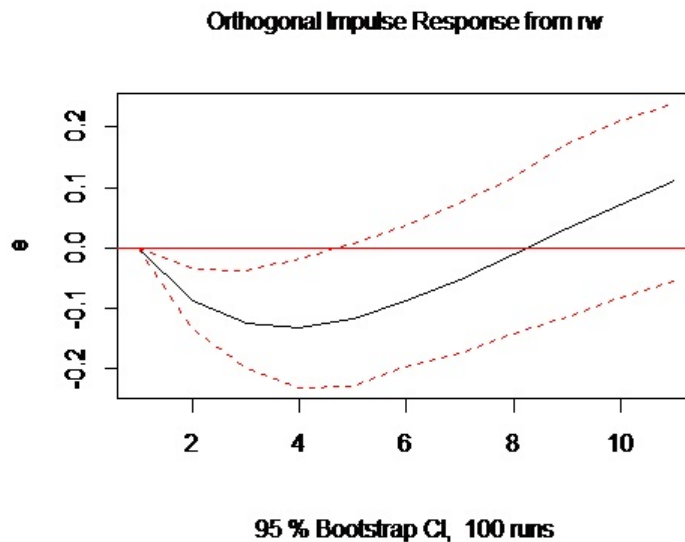
### 3.4.3 Application aux données macroéconomiques canadiennes 91

#### Analyse de réponse impulsionnelle (Analyse du choc)

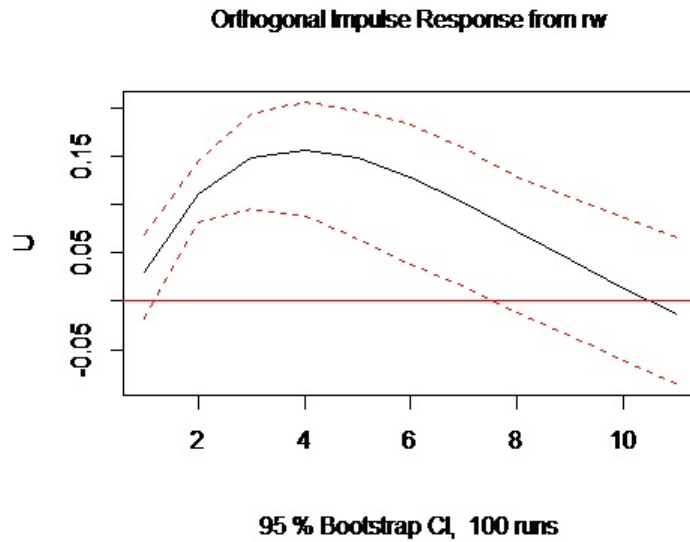
Dans le code R 3.8, la méthode `irf` pour les objets avec l'attribut de classe `svecest` est utilisée et l'argument `boot = TRUE` a été utilisé, de sorte que les bandes de confiance autour des trajectoires de réponse impulsionnelle peuvent être calculées. Le sortie de IRA est illustré à la figure (3.15), (3.16).

#### • Code R 3.8. Analyse de réponse impulsionnelle

```
irf.rw.eU=irf(varsimest, impulse ='rw',response = c('e','U'),
              boot = TRUE)
plot(irf.rw.eU)
irf.rw.rwprod=irf(varsimest, impulse ='rw',response = c('prod',
              'rw'),boot = TRUE)
plot(irf.rw.rwprod)
```

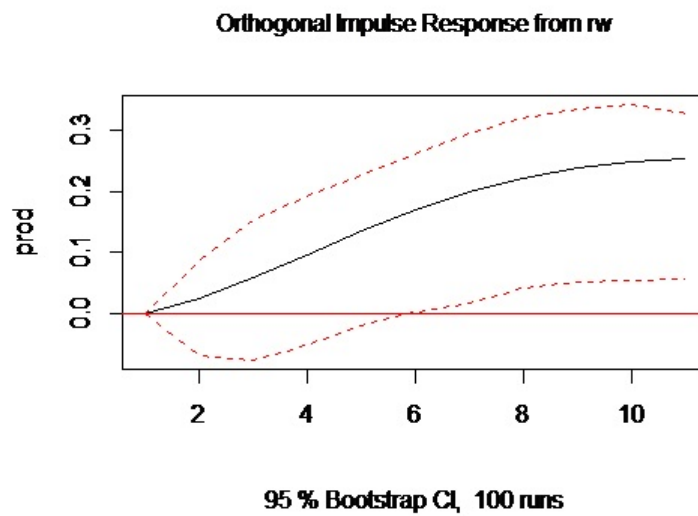


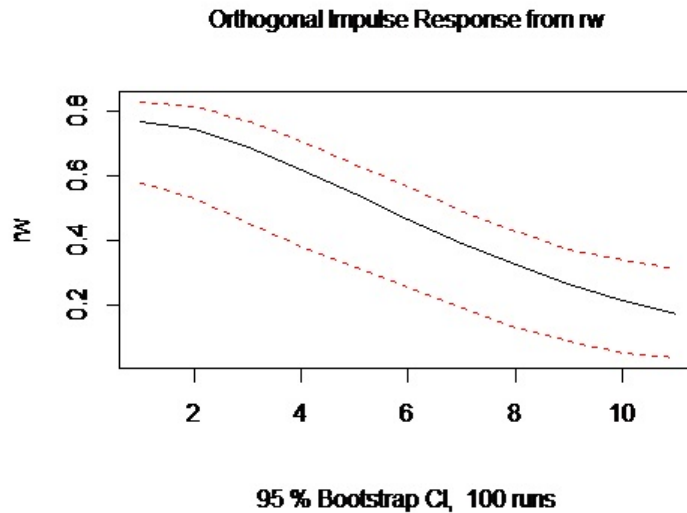
### 3.4.3 Application aux données macroéconomiques canadiennes 92



**Figure 3.15** : Réponse impulsionnelle de U et e à rw.

Même si un choc positif sur les salaires réels (rw) influence l'emploi(e) et le chômage (U) dans les directions que l'on suppose a priori, il ne le fait que de manière significative pour le chômage pendant la période allant de la troisième à la sixième période.





**Figure 3.16** : Réponse impulsionnelle de salaire réel à lui même et PIB

Même si un choc positif sur les salaires réels influence à lui même et le PIB (prod) dans les directions que l'on suppose a priori.

### Décomposition de la variance de l'erreur de prévision

L'étude, basée sur les fonctions de réponse impulsionnelle, peut être complétée par une analyse de la décomposition de la variance de l'erreur de prévision. L'objectif est de calculer la contribution de chacune des innovations à la variance de l'erreur. De façon générale, on écrit la variance de l'erreur de prévision à un horizon  $h$  (ici  $h$  va de 1 à 48) en fonction de la variance de l'erreur attribuée à chacune des variables. On effectue ensuite le rapport entre chacune de ces variances et la variance totale pour obtenir son poids relatif en pourcentage. .

- **Code R 3.9. Décomposition de la variance de l'erreur de prévision du chômage au Canada**

```
fevd.U=fevd(varsimest,n.ahead=10)$U
fevd.U
plot(fevd.U,addbars=2)
```

### 3.4.3 Application aux données macroéconomiques canadiennes 94

Tableau 3.9. Décomposition de la variance de  $U$

Période	e	prod	rw	U
1	0.6277022	0.0008180297	0.008310711	0.36316910
2	0.7049036	0.0202743646	0.067549850	0.20727223
3	0.6723315	0.0602459971	0.117438100	0.14998440
4	0.6063217	0.1140388472	0.142455231	0.13718417
5	0.5397753	0.1763222722	0.148847909	0.13505456
6	0.4805563	0.2431268442	0.144280521	0.13203637
7	0.4292286	0.3108211579	0.134109655	0.12584054
8	0.3850094	0.3759894277	0.121826154	0.11717502
9	0.3471673	0.4357055302	0.109681951	0.10744518
10	<b>0.3151703</b>	<b>0.4878123270</b>	<b>0.099065170</b>	<b>0.09795217</b>

La variance de l'erreur de prévision de  $U$  est due à  $prod$  pour 48%, à ses propres innovations pour 9%, à celles de  $e$  pour 31%, à celles de  $rw$  pour 9%.

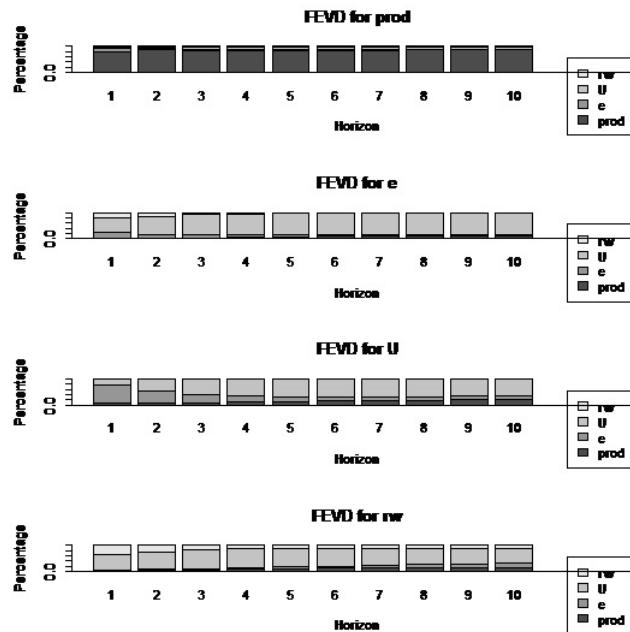


Figure 3.17 : FEVD pour Canada

### 3.4.4 Application aux données macroéconomiques Algériennes

#### Précisions

Les séries chronologiques originales sont publiées sur le site : <http://perspective.usherbrooke.ca/bilan/tend/DZA/fr/NY.GDP.MKTP.CD.html>. La gamme de l'échantillon est du 1er trimestre 1994 au 4ème trimestre 2016. Les séries suivantes ont été utilisées dans la construction de la série prévue à Algérie :

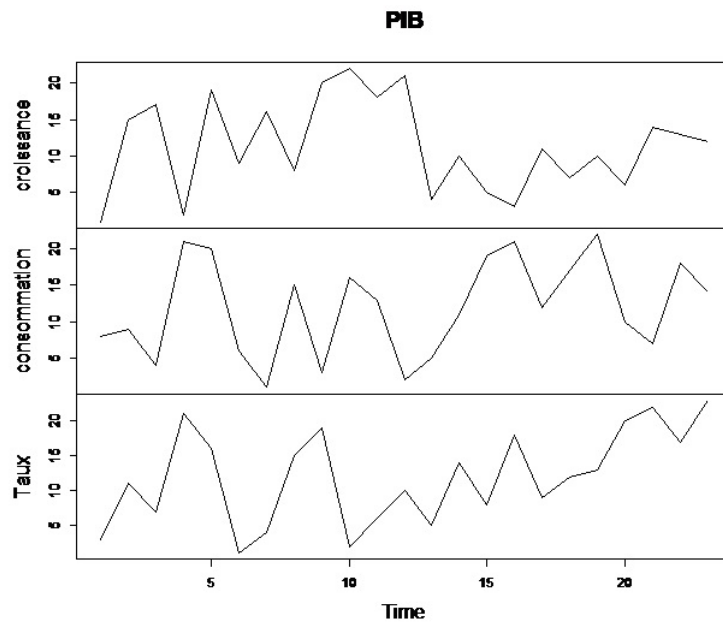
- TIR : Taux d'intérêt réel (%).
- IPC : Indice des prix à la consommation.
- PIB : Croissance annuelle du PIB (%).

#### La méthodologie utilisée :

Une analyse préliminaire des données est effectuée en affichant les statistiques récapitulatives des séries concernées ainsi que les graphiques des séries temporelles correspondantes (voir figure 3.18).

##### • Code R 4.1.

```
ALG= read.table(file.choose())
ALG=as.ts(ALG)
ALG
ALG=window(ALG,start=1994,end=2011,freq=1)
ALG
r=ts(ALG,start=2012,end=2016)
r
colnames(ALG)=c('TIR','PIB','ICP')
acf(ALG)
pacf(ALG)
plot(ALG)
```



**Figure 3.18** : plot des séries chronologiques des donnée Algérie.

Les séries semblent non stationnaires

### Test de stationnarité

Pour chaque variable, on commence par le test ADF, on étudie la significativité de la tendance et si la variable est non stationnaire, on passe au test ADF en première différence.

#### • Code R 4.2. Test de stationnarité

```
adf.test(ALG[, 'ICP'])
```

Augmented Dickey-Fuller Test

data: ALG[, 'ICP']

Dickey-Fuller = -3.899, Lag order = 2, p-value = 0.02864

alternative hypothesis: stationary

```
adf.test(ALG[, 'TIR'])
```

Augmented Dickey-Fuller Test



### 3.4.4 Application aux données macroéconomiques Algériennes 97

```
data: ALG[, 'TIR']
Dickey-Fuller = -1.2674, Lag order = 2, p-value = 0.8515
alternative hypothesis: stationary

adf.test(ALG[, 'PIB'])
```

#### Augmented Dickey-Fuller Test

```
data: ALG[, 'PIB']
Dickey-Fuller = -2.4133, Lag order = 2, p-value = 0.4149
alternative hypothesis: stationary
```

La valeur du test `adf.test` sur la série ICP est inférieure à la valeur critique à 5%, alors la série ICP est stationnaire.

les valeur du test `adf.test` sur la série TIR et PIB est supérieure à la valeur critique à 5%. Les séries TIR et PIB sont non stationnaire.

#### Sélection de l'ordre de décalage

La première étape de cette analyse, consiste à déterminer le nombre de retards du modèle VAR(p). Pour ce faire, nous estimons un certain nombre de processus autorégressifs et nous retenons celui qui minimise conjointement le critères d'Akaike.

#### • Code R 4.3. Déterminer un décalage approprié

```
infocrit=VARselect(varsimest,lag.max = 3,type ='both')
infocrit
```

Selon l'AIC et HQ et SC, le nombre de retard optimal est  $p = 3$ , le critère indique SC une longueur de retard de  $p = 1$ .

#### Estimation une modèle VAR

Dans une étape suivante, la donnée ALG est estimé avec la fonction `VAR()` et en tant que régresseurs déterministes, une constante est incluse.

• Code R 4.4. Estimation

```
varsimest= VAR(ALG,p =1,type ='both',season = NULL,  
              exogen = NULL)  
varsimest
```

VAR Estimation Results:

=====

Estimated coefficients for equation TIR:

=====

Call:

TIR = TIR.l1 + PIB.l1 + ICP.l1 + const + trend

TIR.l1	PIB.l1	ICP.l1	const	trend
0.06134191	0.19410383	0.11293142	13.13921085	-0.49344137

Estimated coefficients for equation PIB:

=====

Call:

PIB = TIR.l1 + PIB.l1 + ICP.l1 + const + trend

TIR.l1	PIB.l1	ICP.l1	const	trend
-0.1688532	-0.1033001	0.2492970	9.0626395	0.3056309

Estimated coefficients for equation ICP:

=====

Call:

ICP = TIR.l1 + PIB.l1 + ICP.l1 + const + trend

TIR.l1	PIB.l1	ICP.l1	const	trend
-0.40038752	-0.03383018	-0.11919600	17.05679369	-0.02971214

On a par exemple :

### 3.4.4 Application aux données macroéconomiques Algériennes 99

---

$$TIR_t = -0.1688532 TIR_{t-1} - 0.1033001 PIB_{t-1} + 0.2492970 ICP_{t-1} + 9.0626395 + 0.3056309 t$$

#### Test de diagnostic

- Code R 4.5. Test de diagnostic

```
‡ test
‡ tester la normalité
var2c.norm = normality.test(varsimest,
  multivariate.only = TRUE)
var2c.norm

$JB

      JB-Test (multivariate)

data:  Residuals of VAR object varsimest
Chi-squared = 5.4196, df = 6, p-value = 0.4912

$Skewness

      Skewness only (multivariate)

data:  Residuals of VAR object varsimest
Chi-squared = 2.4023, df = 3, p-value = 0.4932

$Kurtosis

      Kurtosis only (multivariate)

data:  Residuals of VAR object varsimest
Chi-squared = 3.0173, df = 3, p-value = 0.389

‡ Portmanteau-Test
```

### 3.4.4 Application aux données macroéconomiques Algériennes 100

```
var2c.serial= serial.test(varsimest,lags.pt =16,  
                          type = 'PT.asymptotic')  
var2c.serial
```

Portmanteau Test (asymptotic)

data: Residuals of VAR object varsimest  
Chi-squared = 75.052, df = 135, p-value = 1

```
plot(var2c.serial,names = 'TIR')  
plot(var2c.serial,names = 'ICP')  
plot(var2c.serial,names = 'PIB')
```

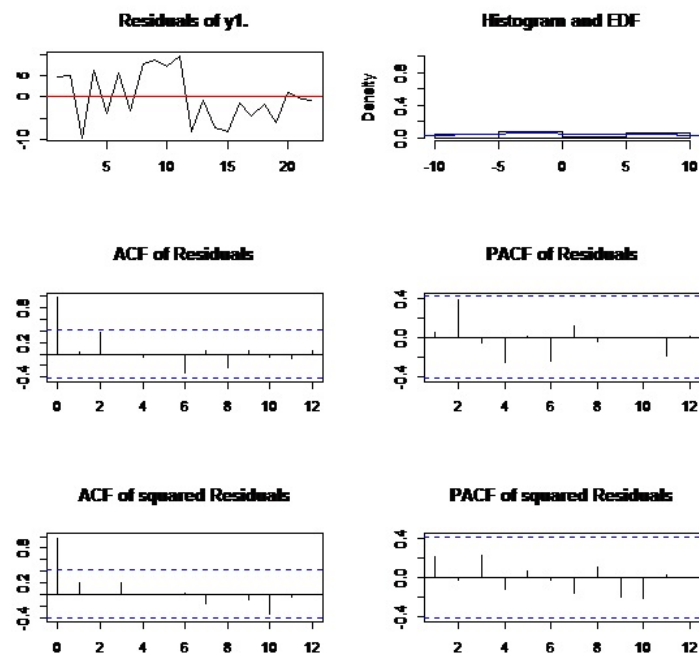


Figure 3.19 : diagnostic des résidus de TIR.

### 3.4.4 Application aux données macroéconomiques Algériennes 101

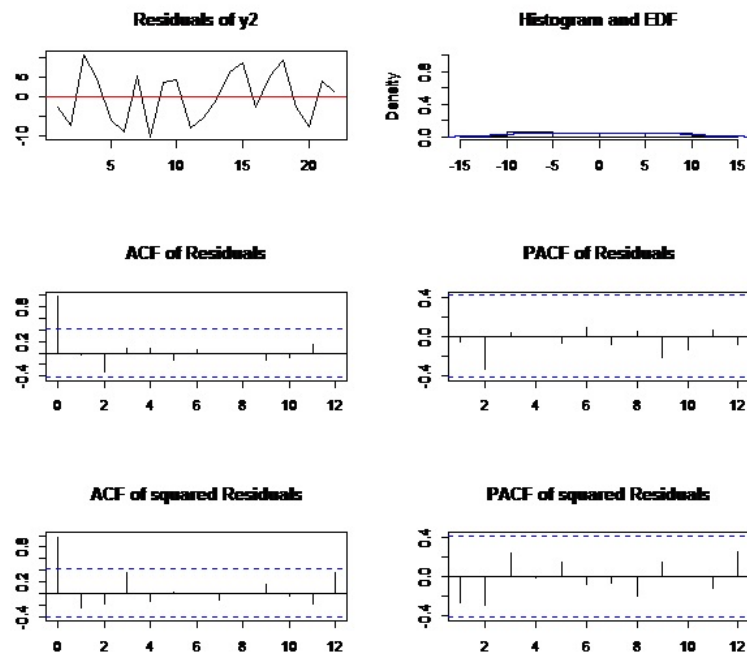


Figure 3.20 ; diagnostic des résidus de ICP.

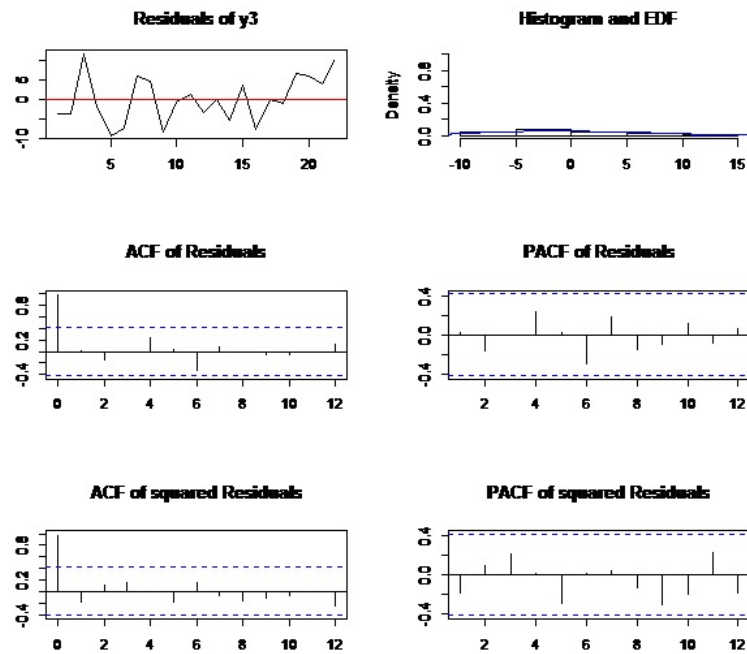


Figure 3.21 : diagnostic des résidus de PIB.

### 3.4.4 Application aux données macroéconomiques Algériennes 102

#### Tests de causalité

Nous nous proposons d'illustrer la notion de causalité au sens de Granger en procédant à un test de non causalité. Les résultats obtenus, pour un nombre de retards  $p$  égal à 1

##### • Code R 4.6. Tests de causalité

```
‡ Granger et causalité instantanée
var.causal=causality(varsimest,cause =c('TIR','ICP'))
var.causal=causality(varsimest,cause =c('TIR','PIB'))
var.causal=causality(varsimest,cause =c('PIB','ICP'))
```

**Tableau 3.10.** *test de causalité*

Hypothèse nulle :	F-Test	Probabilité
TIR ICP ne cause pas PIB	0.24219	0.7861
TIR PIB ne cause pas ICP	0.44936	0.6413
PIB ICP ne cause pas TIR	1.7933	0.1799

Notons que les valeurs indiquées sont significatives au seuil de 10%.

On constate que l'hypothèse nulle selon laquelle le taux d'intérêt réel et indice des prix à la consommation ne cause pas la croissance annuelle du PIB est rejetée au seuil de 10% (la probabilité associée est de  $(0.7861 > 0.10)$ ). En faisant le même travail pour les autres variables, on constate la même chose

#### Prévision

##### • Code R 4.7. Prévision

```
pp= predict(svec,n.ahead =8)
pp
‡ Plot de prévision pour taux d'intérêt réel
plot(pp,names ='TIR')
‡ Fanchart pour indice des prix à la consommation (fanchart)
```

### 3.4.4 Application aux données macroéconomiques Algériennes 103

```
fanchart(pp,names='ICP')  
‡ Fanchart pour croissance annuelle du PIB(fanchart)  
fanchart(pp,names ='PIB')
```

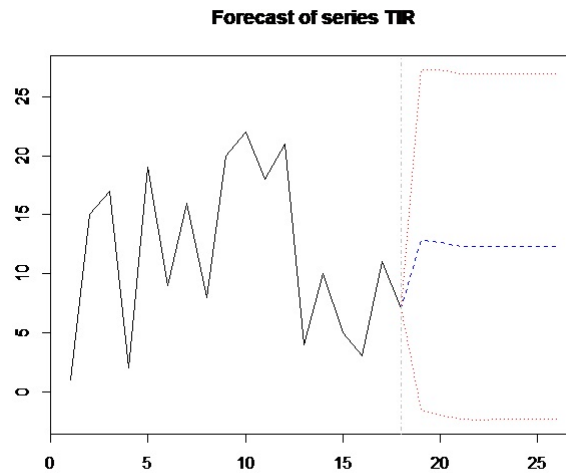


Figure 3.22 : Prévision du taux d'intérêt réel du processus VAR(1).

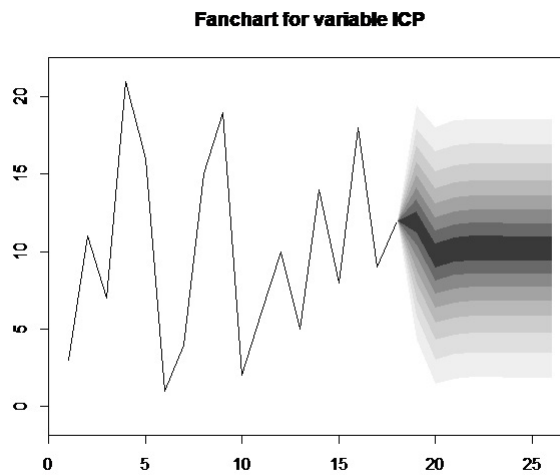
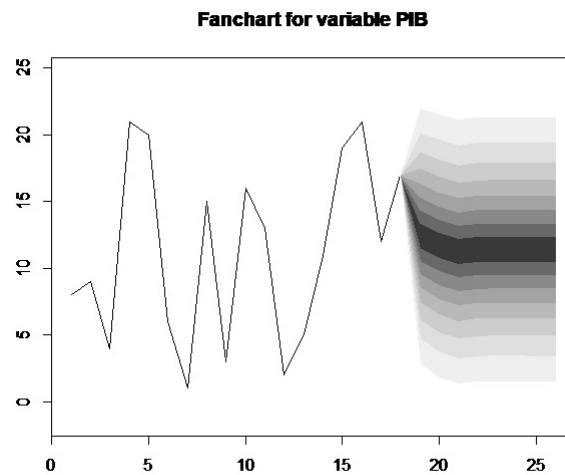


Figure 3.23 : Prévision de l'indice des prix à la consommation du processus VAR(1) .

### 3.4.4 Application aux données macroéconomiques Algériennes 104



**Figure 3.24** : Prévision de croissance annuelle du PIB de processus VAR(1) .

#### Analyse de réponse impulsionnelle (analyse du choc)

la méthode `irf` pour les objets avec l'attribut de classe `svecest` est utilisée et l'argument `boot = TRUE` a été utilisé, de sorte que les bandes de confiance autour des trajectoires de réponse impulsionnelle peuvent être calculées

#### • Code R 4.8. Analyse de réponse impulsionnelle

```
irf.TIR=irf(varsimest,impulse='TIR',response=
  c('PIB'),boot = TRUE)
irf.TIR
plot(irf.TIR)
irf.ICP=irf(varsimest,impulse='ICP',response=
  c('PIB'),boot = TRUE)
irf.ICP
plot(irf.ICP)
irf.PIB=irf(varsimest,impulse='PIB',response=
  c('PIB'),boot = TRUE)
irf.PIB
plot(irf.PIB)
```



### 3.4.4 Application aux données macroéconomiques Algériennes 105

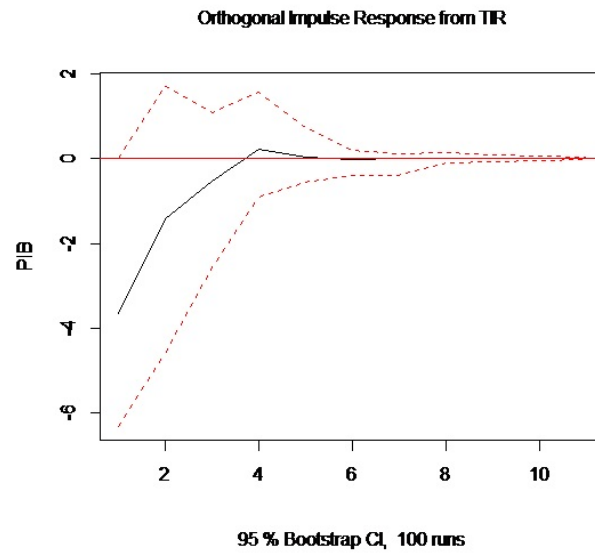


Figure 3.25 : Réponse impulsionnelle du TIR à la PIB .

La réponse de PIB à un choc sur TIR est négative et instantanée figure (3.25).

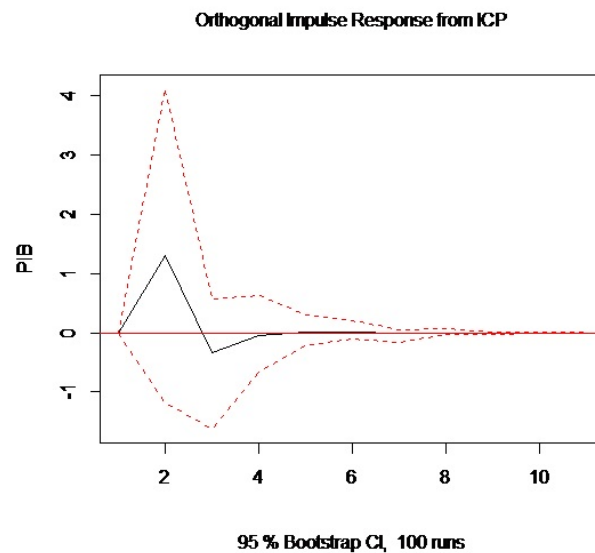
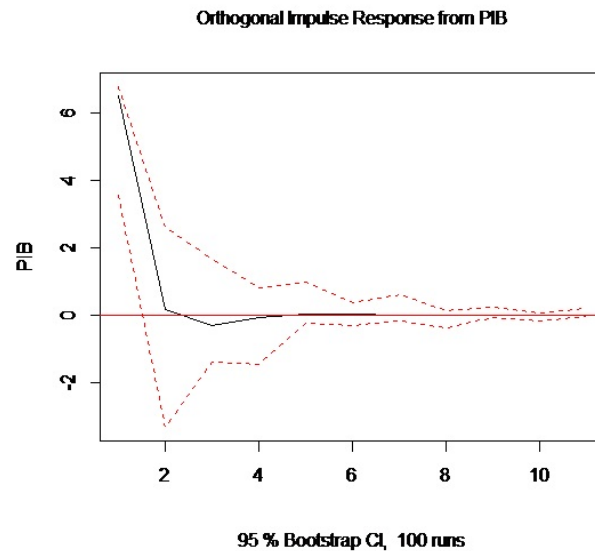


Figure 3.26 : Réponses impulsionnelles du ICP à PIB.

### 3.4.4 Application aux données macroéconomiques Algériennes 106

En effet, un choc positif sur ICP se traduit par un effet positif sur PIB pendant la première année. Cet effet disparaît ensuite et rapidement avant de trouver son niveau de court terme (figure 3.26)



**Figure 3.27** : Réponses impulsionnelles de PIB à lui même.

La réponse de PIB à un choc sur lui même est positive (figure 3.27).

#### Décomposition de la variance de l'erreur de prévision

- **Code R 4.9. Décomposition de la variance de l'erreur de prévision**

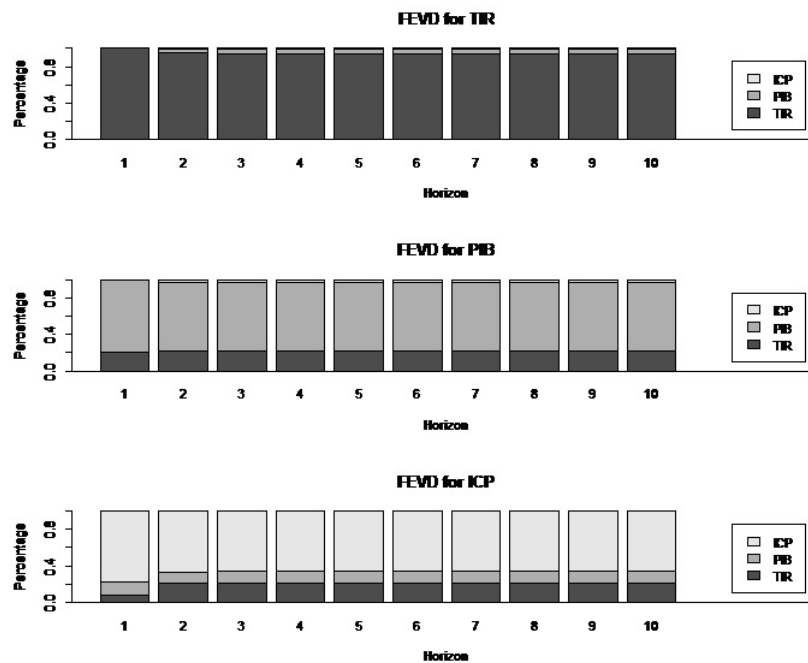
```
fevd.U=fevd(varsimest,n.ahead=10)$ICP
fevd.U
plot(fevd(varsimest,n.ahead=10),addbars=2)
```

### 3.4.4 Application aux données macroéconomiques Algériennes 107

**Tableau 3.11.** *Décomposition de la variance de ICP*

Période	TIR	PIB	ICP
1	0.08265066	0.1372674	0.7800819
2	0.21227758	0.1212311	0.6664913
3	0.21517234	0.1265882	0.6582395
4	0.21568058	0.1265990	0.6577204
5	0.21566097	0.1266401	0.6576989
6	0.21567396	0.1266380	0.6576881
7	0.21567374	0.1266387	0.6576876
8	0.21567384	0.1266387	0.6576875
9	0.21567384	0.1266387	0.6576875
10	<b>0.21567384</b>	<b>0.1266387</b>	<b>0.6576875</b>

La variance de l'erreur de prévision de l'indice des prix à la consommation est au taux d'intérêt réel pour 21%, à ses propres innovations pour 65%, à l'indice des prix à la consommation pour 12%.



**Figure 3.28 :** FEVD pour processus VAR(1)

## 3.5 Conclusion

La représentation VAR, a connu un nombre considérable de développements et d'applications au cours des dernières décennies qui lui ont assuré un succès croissant. Tout d'abord, la méthodologie a été perfectionnée à l'aide de nouvelles méthodes d'identification des chocs et par le recours aux méthodes bayésiennes. Par ailleurs, les modèles vectoriels à correction d'erreur (VECM), très populaires ces dernières années, reposent sur la représentation VAR. Les processus VAR se révèlent particulièrement utiles pour étudier l'impact des chocs macroéconomiques - ainsi que leurs canaux de transmission - ce qui explique qu'un très grand nombre de travaux empiriques y ait eu recours pour étudier l'effet des politiques économiques. Ce mémoire m'a permis d'apprendre de nouveaux concepts et objectifs (causalité et analyse de choc) de l'étude d'une série chronologiques et de connaître les modèles VAR de près, Ce travail m'a permis aussi d'assembler mes connaissances et d'améliorer mes compétence en matière de programmation par le langage R.

# Bibliographie

- [1] Akaike, H. (1969) : Fitting autoregressive models for prediction, annals of the Institute of Statistical Mathematics, 21 : 243-247.
- [2] Bernhard, Pfaff : Analysis of Integrated Series with R and Cointegrated Time Series with R, 61476 Kronberg im Taunus, Germany.
- [3] Brockwell, P. J. et Davis, R. A. (2002) : Introduction to Time Series and Forecasting, 2e édition. Springer : New York.
- [4] Bourbonnais, R : 'Économétrie', Edition Dunod, ISBN.
- [5] Gouriéroux, C., Monfort, A. (1995) : Séries temporelles et modèles dynamiques. Economica.
- [6] Granger, C. (1981) : Some properties of time series data and their use in econometric model specification, Journal of Econometrics 16, 121- 130.
- [7] Hafner, C. M. and Herwartz, H. (2009) : Testing for linear vector autoregressive dynamics under multivariate generalized autoregressive heteroskedasticity, Statistica Neerlandica, 63 : 294-323.
- [8] Hamilton, J. (1994) : Time Series Analysis, Princeton University Press, Princeton.
- [9] Hannan, E. J. (1970) : Multiple Time Series, Wiley : New York.
- [10] Hendry, D. (1995) : Dynamic Econometrics, Oxford University Press, Oxford.
- [11] Jean-Baptiste Gossé, L'apport de la représentation VAR de Christopher A. Sims à la science économique, CFAP, University of Cambridge [jg548@cam.ac.uk](mailto:jg548@cam.ac.uk), Cyriac Guillaumin, CREG, Université Grenoble Alpes, [cyriac.guillaumin@upmf-grenoble.fr](mailto:cyriac.guillaumin@upmf-grenoble.fr).

- 
- [12] Lütkepohl, H., Krätzig, M., Phillips, P., Ghysels, E. , Smith, R. (2004) : Applied Time Series Econometrics (Themes in Modern Econometrics), Cambridge University Press.
- [13] Lütkepohl, H. (2005) : New Introduction to Multiple Time Series Analysis, Springer-Verlag : Berlin.
- [14] Lütkepohl, H. (2006) : New Introduction to Multiple Time Series Analysis, Springer, New York.
- [15] Quinn, B. (1980) : Order determination for a multivariate autoregression, Journal of the Royal Statistical Society, B42 : 182-185.
- [16] Shumway, R. H. et Stoffer, D. S. (2006) : Time Series Analysis and Its Applications With R Examples, Springer : New York.
- [17] Sims, C. (1980) : Macroeconomics and Reality, Econometrica, 48, 1–48.
- [18] Stock, J.H. and M.W. Watson (2001) : Vector Autoregressions, Journal of Economic Perspectives, 15, 101-115.
- [19] Wei, W. W. (1994) : Time Series Analysis, Univariate and Multivariate Methods, Addison-Wesley Company.