

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique  
UNIVERSITÉ KASDI MERBAH OUARGLA  
FACULTÉ DE MATHÉMATIQUES ET SCIENCES DE LA  
MATIÈRE  
DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté En Vue De L'obtention Du

**DIPLÔME DE MASTER**

EN MATHÉMATIQUES

**Option** : Probabilité et Statistique

Par

**Guermi Hanane**

Intitulé

**Les modèles vectoriels autorégressifs VAR(p)**

Membres du jury

AMARA Abdelkader	M. C. A	UKMO	Président
MEDDI Fatima	M. C. B	UKMO	Examineur
ARBIA Hanane	M. C. B	UKMO	Rapporteur

2021-2022

## Résumé

Notre travail est basé sur la modélisation vectorielle autorégressive VAR d'ordre ( $p$ ). Nous avons modélisé deux séries temporelles avec des données réelles. Nous avons appliqué les critères du modèle d'apparence et nous l'estimons.

**Mots clés:** Modèles autorégressifs vectoriels, validation, le lag optimale, racine unitaire.

---

## Abstract

Our work is based on VAR autoregressive vector modeling of order ( $p$ ). We have modeled two time series with real data. We have applied the criteria of the appearance model and we estimate it.

**Keywords:** Vector autoregressive models, validation, optimal lag, unit root.

---

## ملخص

يعتمد عملنا على نمذجة الانحدار الذاتي VAR برتبة ( $p$ ). لقد قمنا بنمذجة سلسلتين زمنيتين لبيانات حقيقية. طبقنا معايير النموذج وقمنا بتقديره.

**الكلمات المفتاحية:** نماذج الانحدار الذاتي، الصحة، التأخر الأمثل، جذر الوحدة.

# Table des matières

Dédicace	iv
Remerciements	v
Liste des tableaux	vii
Table des figures	viii
Notations et abreviations	ix
Introduction	1
<b>1 Concepts de base des séries chronologiques</b>	<b>3</b>
1.1 Définition d'une série chronologique . . . . .	3
1.2 Les composantes d'une série chronologique . . . . .	4
1.2.1 La tendance . . . . .	4
1.2.2 La composante saisonnière . . . . .	4
1.2.3 La composante cyclique . . . . .	5
1.2.4 Les irrégularités . . . . .	5
1.3 Modélisation de la série . . . . .	6
1.3.1 Le modèle additif . . . . .	6
1.3.2 Le modèle multiplicatif . . . . .	7
1.3.3 Le modèle mixtes . . . . .	7

1.4	Choix du modèle . . . . .	8
1.5	Stationnarité . . . . .	8
1.6	Les processus aléatoires non stationnaires . . . . .	9
1.7	Bruit Blanc . . . . .	10
1.8	Autocovariance et autocorrelation . . . . .	11
1.9	Opérateurs de retard et de différenciation . . . . .	13
<b>2</b>	<b>Les modèles Vectoriels autorégressifs</b>	<b>16</b>
2.1	La représentation VAR(p) . . . . .	16
2.2	Stationnarité d'un modèle VAR . . . . .	19
2.3	Estimation des paramètres . . . . .	21
2.4	Détermination du nombre de retards . . . . .	22
2.5	Prévision . . . . .	22
2.5.1	Le cas d'un VAR(1) . . . . .	23
2.5.2	Le cas d'un VAR(p) . . . . .	24
2.6	La causalité . . . . .	25
2.6.1	Causalité au sens de Granger . . . . .	25
2.6.2	Causalité au sens de Sims . . . . .	28
2.7	La cointégration . . . . .	28
<b>3</b>	<b>Application</b>	<b>30</b>
3.1	Test de la stationnarité . . . . .	31
3.1.1	Test de la racine unitaire (ADF) . . . . .	33
3.2	Stationnarisation des séries . . . . .	39
3.3	Détermination de degré de retard . . . . .	47
3.4	Estimation du modèle . . . . .	48
3.5	La validation du modèle . . . . .	50
3.5.1	Test de la stationnarité du modèle . . . . .	50
3.6	Tests sur les résidus . . . . .	51
3.6.1	Test de la normalilté des résidus . . . . .	51
3.6.2	Autocorrélation test . . . . .	52
3.6.3	Test d'hétéroscédasticité . . . . .	53

3.7 La prévision . . . . .	54
<b>Conclusion</b>	<b>56</b>
<b>Bibliographie</b>	<b>57</b>

# Dédicace

*Je dédie ce modeste travail*

*A mes parents qui tiennent une place immense dans mon coeur.  
Med El Bachir et Fatima El Zahra*

*Tout ce que j'ai accompli dans ma vie, c'est grâce à votre soutien,  
votre amour et vos sacrifices votre amour et vos sacrifices. Merci  
infiniment. . .*

*A mes soeur Asma, Nawal, Kawther, Khalida , Nadjah, Sirine*

*A mon Cher frère Yasser*

*A mes grandes -mères et mon grand-père.*

*A ma belle- famille tantes et oncles*

*A mon fiancé Abdellbari*

*A mes amis Rayhana, Rima, Ilham*

*A tout ceux qui m'aiment, m'encouragent, m'ont aidé et qui m'aident  
toujours pour continuer sur la bonne voie.*

# Remerciements

*Je remercie Allah le tout puissant de m'avoir donné le courage, la volonté, la santé, et là Patience de mener à terme ce présent travail.*

*Mes plus vifs remerciements et ma profonde gratitude vont à Pr.ARBIA HANANE notre promoteur de mémoire pour sa grande patience, pour sa disponibilité, pour ses nombreux conseils, pour ses corrections et son appréciation au cours de l'élaboration de ce travail, notamment quand à la portée de la problématique. Elle a toujours été présent dans les moments où j'en avions besoin, elle je faudrait des pour le remercier.*

*Mes vifs remerciement vont également aux membres du jury pour l'intérêt qu'ils ont porté à ma recherche en acceptant d'examiner mon travail et de l'enrichir par leurs propositions.*

*Je tiens à remercier, particulièrement, Mr.Ibrahim Gouasse chef du  
Bureau de l'Administration et des Moyens au Centre National du  
Registre Commerce-Ouargla  
qui n'a jamais cessé de me soutenir, pour l'aide qu'il m'a procurée et  
pour ses  
précieux conseils.*

*Merci beaucoup au Centre  
la complexe Statistique de la wilaya Ouargla pour le temps et l'aide  
qu'ils m'ont apporté*

*Enfin, je ne saurais terminer cette partie sans exprimer notre  
gratitude à mes  
parents, mes grands parents, mes frères et mes soeurs qui j'ont  
toujours soutenue, encouragé et stimulée pendant mes études.*

*Merci à tous et à toutes.*



# Liste des tableaux

3.1	L'évolution des Taxe sur l'activité professionnels (TAP) selon le registre du commerce (Rm) de la wilaya de ouargla au cours de la période 2003-2019 . . . . .	31
3.2	Le degré de retard . . . . .	47
3.3	Prévision. . . . .	54

# Table des figures

1.1	Les composantes d'une série chronologique. . . . .	6
3.1	La série de Taxe sur l'activité professionnels (TAP) . .	32
3.2	La série de le registre du commerce(Rm) . . . . .	32
3.3	Représentation graphique de la série stationnaire (DTAP). 43	
3.4	Représentation graphique de la série stationnaire (DDRM). 47	
3.5	Les racines du polynôme caractéristique. . . . .	50
3.6	La prévision de TAP. . . . .	54
3.7	La prévision de RM. . . . .	55

## Notations et abriviations

<b>ADF</b>	Test de Dickey Fuller Augmenté
<b>AIC</b>	Critère d'information d'Akaike.
<b>ARIMA</b>	Autorégressif moyenne mobile intégré
<b>ARMA</b>	Autorégressif moyenne mobile
<b>BB</b>	Bruit blanc
<b>BBN</b>	Bruit blanc normale
<b>Cov</b>	Covariance
<b>DS</b>	Difference Stationary
<b>L</b>	Opérateur retard.
<b>MA</b>	Moyennes Mobiles (Moving Average).
<b>FAC</b>	Fonction d'autocorrélation
<b>MCO</b>	Moindres carrés ordinaires.
<b>TS</b>	Trend Stationary
<b>MAE</b>	Erreur absolue moyenne

**SC** Critère de Schwarz

**VAR** vecteur autorégressif

# Introduction

Dans l'analyse de séries temporelles multivariées, la plupart des modèles actuels sont utilisés à des fins de prévision, c'est-à-dire pour estimer les valeurs futures du système étudié à partir de l'historique des valeurs observées dans le passé, alors que les séries temporelles sont utilisées depuis longtemps et ils ont été utilisés dans divers domaines, par exemple : en médecine, en économétrie en astronomie, en météorologie et en théorie du signal.

Une série chronologique est un ensemble d'observations générées successivement dans le temps. L'avantage de toute série chronologique est que ses données sont disposées par rapport au temps et que les observations successives ne sont généralement pas indépendantes, c'est-à-dire Dépendez les uns des autres.

Le but de l'analyse des séries chronologiques est de connaître les changements qui se produisent dans le phénomène au cours d'une certaine période comparer les valeurs du phénomène entre elles car il est mesuré dans les mêmes unités et de la même manière à des dates différentes.

Les séries chronologiques comportent plusieurs modèles : AR, MA et ARIMA (cas univarié) et les modèle VAR (cas multivarié).

Dans ce travail nous présentons une famille de modèles autorégressifs de vecteurs VAR(p) [2]. Afin de traiter ces modèles, nous les appliquons à des données réelles.

Cette thèse est composée de trois chapitres. Dans le premier chapitre, nous avons abordé le concept de série temporelle, ses composantes, ses caractéristiques et propriétés (la stationnarité, le fonction d'autocovariance, et les fonctions d'autocorrélation simples et partielles,...).

Le deuxième chapitre traite les modèles Vectoriels autorégressifs (VAR) d'ordre (p). Un modèle VAR(p) est une généralisation des modèles AR(p) au cas multivarié. La dimension de ce processus représente le nombre de variables étudiées qui composent ce processus vectoriel et pour lequel on cherche à établir des relations linéaires entre les variables.

Le troisième chapitre représente une partie d'application, nous avons abordé la modélisation de deux séries temporelles avec des données réelles (Taxe sur l'activité professionnelle et le registrar du commerce) de la wilaya de Ouargla au cours de la période 2003-2019.

# Chapitre 1

## Concepts de base des séries chronologiques

Ce chapitre a pour objectif de commencer à familiariser le lecteur avec les principaux aspects des séries chronologiques.

### 1.1 Définition d'une série chronologique

Une série chronologiques (série temporelle) est une suite d'observations répétées d'un même phénomène à des dates différentes (par exemple la température moyenne journalière en un lieu donné, la consommation moyenne en électricité chaque mois, le prix du baril de pétrole chaque jour...). [9], [11].

Les dates sont souvent équidistantes (séries mensuelles, trimestrielles ou annuelles).

**Définition 1.1** Une série chronologique  $X_t$  est un ensemble d'observations indexée par le temps  $t$ .

## 1.2 Les composantes d'une série chronologique

On donne ici une présentation des trois aspects : la tendance, la composante saisonnière, la composante cyclique et la composante irrégulière. On peut décrire ainsi ces composantes : [5], [8]

### 1.2.1 La tendance

La tendance ou trend, notée  $Z_t$ , est l'évolution à long terme de la grandeur. La tendance traduit l'aspect général de la série : croissance de la consommation d'électricité, croissance du trafic aérien, diminution de la population rurale,....

La tendance a plusieurs forme :

1. **Tendance linéaire** :  $Z_t = a + b.t$ .
2. **Tendance quadratique** :  $Z_t = a + b.t + c.t^2$ .
3. **Tendance exponentiel** :  $Z_t = \exp(a + b.t)$ .
4. **Tendance logarithmique** :  $Z_t = a + b.\ln(t)$ .

### 1.2.2 La composante saisonnière

Le facteur saisonnier, noté  $S_t$ , se répète à intervalles de temps égaux avec une forme à peu près constante. Il peut être dû au rythme des saisons ou à des facteurs humains. Sa période est de 12 pour des séries mensuelles, de 4 pour des séries trimestrielles.

Si  $p$  désigne la période du mouvement saisonnier :

$$S_t = S_{t+p}$$



Le facteur saisonnier est donc totalement déterminé par  $p$  coefficients saisonniers :

$$S_t, \dots, S_j, \dots, S_p$$

### 1.2.3 La composante cyclique

Une autre composante parfois étudiée de manière spécifique a trait au phénomène cyclique ( $C_t$ ). Il s'agit d'un phénomène se répétant mais contrairement à la saisonnalité sur des durées qui ne sont pas fixes et généralement plus longues.

### 1.2.4 Les irrégularités

Cette composante, appelée aussi mouvement résiduel et notée  $\varepsilon_t$ , regroupe tout ce qui n'a pas été pris en compte par la tendance le facteur saisonnier.

Elle est la résultante de fluctuations irrégulières et imprévisibles dues à des facteurs perturbateurs non permanents.

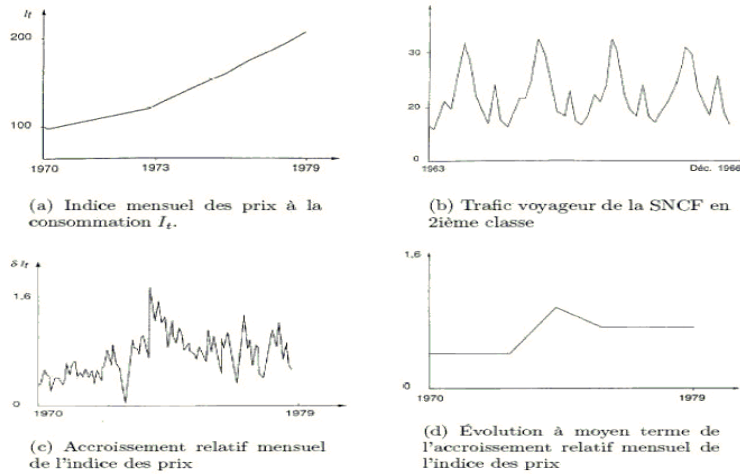


FIG. 1.1 – Les composantes d'une série chronologique.

## 1.3 Modélisation de la série

Un modèle est une image simplifiée de la réalité qui vise à traduire les mécanismes de fonctionnement du phénomène étudié et permet de mieux les comprendre. Un modèle peut être meilleur qu'un autre pour décrire la réalité. Nous présentons dans cette section une petite liste qui sert à résumer et classifier les différents modèles envisagés dans ce travail .

### 1.3.1 Le modèle additif

Nous considérons une série  $X_t$  admettant une décomposition additive

$$X_t = Z_t + S_t + \varepsilon_t, \quad t = 1, \dots, T.$$

Dans ce modèle, l'amplitude de la série reste constante au cours du temps. Ceci se traduit graphiquement par des fluctuations autour de

la tendance  $Z_t$  constantes au bruit près.

### 1.3.2 Le modèle multiplicatif

Nous considérons une série  $X_t$  admettant une décomposition multiplicative

$$X_t = Z_t \cdot S_t \cdot \varepsilon_t, \quad t = 1, \dots, T.$$

L'amplitude de la série n'est plus constante au cours du temps : elle varie au cours du temps proportionnellement à la tendance  $Z_t$  au bruit près.

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

### 1.3.3 Le modèle mixtes

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

$$X_t = Z_t S_t + \varepsilon_t, \quad t = 1, \dots, T.$$

## 1.4 Choix du modèle

Le choix d'un modèle ou d'un autre, l'incorporation ou non d'une composante, peuvent s'apprécier d'après le graphique de la série. [8]

**Graphiquement** : Constatation graphique du parallélisme ou non des droites des minima et des maxima :

- Droites parallèles : modèle additif.
- Droites non parallèles : modèle multiplicatif.

**Algébriquement** : On recherche le coefficient directeur de ces deux droites par une méthode analytique (deux points).

- Coefficients directeurs très proches : le modèle choisi sera additif.
- Coefficients directeurs très différents : le modèle choisi sera multiplicatif.

## 1.5 Stationnarité

Avant le traitement d'une série chronologique, il convient d'étudier les caractéristiques stochastiques. Si ces caractéristiques c'est-à-dire son espérance et sa variance se trouvent modifiées dans le temps, la série chronologique est considérée comme non stationnaire; dans le cas d'un processus stochastique invariant, la série temporelle est alors stationnaire.[4], [9]

De manière formalisée, le processus  $X_t$  est stationnaire si :

**Définition 1.2** *Un processus  $(X_t)$  est stationnaire au sens fort si pour tous  $t_1 < t_2 < \dots < t_n$  tel que  $t_i = 1, \dots, n$  si :*

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \stackrel{L}{=} (X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}) \forall k \in \mathbb{N}^*, \forall t_1, \dots, t_n \in \mathbb{Z}$$

**Définition 1.3**  $(X_t)$  est dit faiblement stationnaire si :

1. pour tout  $t \in \mathbb{Z}$ ,  $E[X_t^2] < \infty$
2. Il existe  $\mu \in \mathbb{R}$ , pour tout  $t \in \mathbb{Z}$ ,  $E(X_t) = \mu$ , la moyenne est constante et indépendante du temps  $t$
3. Il existe une fonction  $\gamma(h)$  telle que pour tous  $t \in \mathbb{Z}$ ,  $cov(X_t, X_{t-h})$  ne dépend pas de  $t$  et dans ce cas elle est notée :

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

## 1.6 Les processus aléatoires non stationnaires

La classe des processus non stationnaire est relativement vaste, il existe différents types de non stationnarité, on présente deux classes des processus non stationnaires : les processus *TS* et *DS* . [10], [11]

**Définition 1.4**  $(X_t)$  est un processus non stationnaire de type *TS* s'il peut s'écrire sous la forme

$$X_t = f(t) + z_t$$

où  $f(t)$  est une fonction polynômiale du temps, linéaire ou non linéaire et  $(z_t)$  est processus stochastique stationnaire.

Le processus  $(X_t)$  s'écrit comme la somme d'une fonction déterministe du temps et d'une composante stochastique stationnaire. Ce

processus ne satisfait plus la définition de la stationnarité du second ordre. On a en effet :

$$E(X_t) = f(t) + E(z_t)$$

Ce processus  $TS$  est non stationnaire car  $E(X_t)$  dépend du temps.

**Définition 1.5** *Un processus  $(X_t)$  est dit difference-stationary DS ou stationnaire en différence s'il peut s'écrire sous l'une des formes suivante :*

1.  $X_t = X_{t-1} + \varepsilon_t$ .
2.  $X_t = X_{t-1} + \beta + \varepsilon_t$  (avec  $\beta \neq 0$ )

La bonne manière de stationnariser une série  $TS$  consiste à estimer, en général par les moindres carrés ordinaires (MCO), l'expression de la tendance et à la retirer. Tandis que la stationnarisation des séries  $DS$  se fait par passage aux différences.

## 1.7 Bruit Blanc

**Définition 1.6 (Bruit blanc)** *Un bruit blanc est une suite de variables aléatoires  $(\varepsilon_t), t \in \mathbb{Z}$  centré  $E(\varepsilon_t) = 0$  et de variance constante  $\sigma^2$ . On note  $\varepsilon_t \sim BB(0, \sigma^2)$ .*

**Remarque 1.1** *Un bruit blanc est faiblement stationnaire (c'est à dire  $E(\varepsilon_t)$  et  $\text{var}(\varepsilon_t)$  ne dépendent pas de  $t$ ) et tel que  $\text{cov}(\varepsilon_t, \varepsilon_{t+h}) = 0$  si  $h \neq 0$ .*

**Définition 1.7 (Bruit blanc gaussien)** *Un processus  $((\varepsilon_t), t \in \mathbb{Z})$ , est un bruit blanc gaussien s'il s'agit d'un bruit blanc dont les variables aléatoires sont indépendantes et identiquement distribuées de loi normale  $N(0, \sigma^2)$ . On note  $\varepsilon_t \sim \text{BBN}(0, \sigma^2)$ .*

## 1.8 Autocovariance et autocorrelation

Dans cette section, nous présentons les fonctions d'autocorrélation et d'autocovariance qui sont des outils bien utiles à différentes étapes des études de séries temporelles. [1], [3], [4]

**Définition 1.8** *La fonction de covariance( ou d'autocovariance) d'un processus aléatoire  $(X_t, t \in \mathbb{Z})$ , notée  $\gamma(\cdot)$  est définie par :*

$$\forall t, s \in \mathbb{Z}, \gamma(t, s) = \text{cov}(X_t, X_s) = E[(X_t - E(X_t))(X_s - E(X_s))]$$

ou  $E(X_t) = E(X_s) = \mu$  est l'espérance du processus.

Pour un processus stationnaire, on peut réécrire la fonction de covariance comme une fonction d'une seule variable :

$$\forall h \in \mathbb{Z}, \gamma(h) = \text{cov}(X_t, X_{t+h}) = \gamma(-h)$$

**Remarque 1.2** *La fonction  $\gamma(\cdot)$  est paire.*

**Définition 1.9** *La fonction d'autocorrelation d'un processus stationnaire  $(X_t, t \in \mathbb{Z})$ , notée  $\rho(\cdot)$  est définie par :  $\forall h \in \mathbb{Z}$ ,*

$$\begin{aligned}
\rho(h) &= \frac{\text{cov}(X_t, X_{t+h})}{\sqrt{\text{var}(X_t) \text{var}(X_{t+h})}} \\
&= \frac{\text{cov}(X_t, X_{t+h})}{\text{var}(X_t)} \\
&= \frac{\gamma(h)}{\gamma(0)}
\end{aligned}$$

avec  $\gamma(0)$  est la variance du processus. Par définition, on a donc  $\rho(0) = 1$ .

**Remarque 1.3** La fonction  $\rho(\cdot)$  est paire ( $\rho(-h) = \rho(h)$ ).

L'équivalent empirique de la fonction d'autocorrélation, noté  $\hat{\rho}(h)$ , est obtenu à partir de l'estimateur suivant pour l'autocovariance  $\hat{\gamma}(h)$  à l'ordre  $h$  :

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X}), \forall h = 0, \dots, n-1, \quad n \geq 1$$

La matrice d'autocorrélation du vecteur  $(X_t, X_{t+1}, \dots, X_{t+h})$  est :



$$R(h) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(h-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(h-2) \\ \rho(2) & \rho(1) & 1 & \dots & \rho(h-3) \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho(h-1) & \rho(h-2) & \rho(h-3) & \dots & 1 \end{pmatrix}$$

**Définition 1.10** *La fonction d'autocorrelation partielle mesure la liaison (linéaire) entre  $X_n$  et  $X_1$  une fois retirés les liens transitant par les variables intermédiaires  $X_2, \dots, X_{n-1}$  :*

$$r(X_1, X_n) = \text{corr}(X_1, X_n / X_2, \dots, X_{n-1})$$

## 1.9 Opérateurs de retard et de différenciation

Ces opérateurs peuvent être utilisés afin de transformer un processus de moyenne non nulle en un processus moyenne de nulle. [4], [10]

**Définition 1.11 (Opérateur retard)** *Soit  $L$  l'opérateur retard,  $(X_t, t \in \mathbb{Z})$  une série de valeurs et  $k$  l'ordre du retard, l'opérateur qui fait passer de  $X_t$  à  $X_{t-k}$  :*

$$LX_t = X_{t-k}$$

on a

$$L^2 X_t = L(LX_t) = LX_{t-1} = X_{t-2}$$

Alors :

$$L^k X_t = X_{t-k}, \forall t \in \mathbb{Z}, \forall k \in \mathbb{N}$$

Si  $X_t = c, \forall t \in \mathbb{Z}$  alors :

$$L^k X_t = L^k c = c, \forall t \in \mathbb{Z}$$

car une constante ne peut varier.

**Définition 1.12 (opérateur de différenciation)** *L'opérateur de différence  $\Delta$  est défini par :*

$$\Delta X_t = X_t - X_{t-1}$$

Alors :

$$\Delta X_t = X_t - LX_{t-1} = (1 - L) X_t$$

*La différence seconde*

$$\Delta^2 X_t = \Delta(\Delta X_t) = (1 - L)^2 X_t = (1 - 2L + L^2) X_t = X_t - 2X_{t-1} + X_{t-2}$$

D'une façon générale :

$$\Delta^d X_t = \Delta (\Delta^{d-1} X_t) = (1 - L)^d X_t.$$

### Propriété 1.1

1.  $Lc = c$ , l'opérateur d'une constante  $c$  est une constante.
2.  $L^0 X_t = X_t$ .
3.  $(1 - L)^2 X_t \neq (1 - L^2) X_t$  en effet :

$$(1 - L)^2 X_t = (1 - 2L + L^2) X_t = X_t - 2X_{t-1} + X_{t-2} \neq X_t - X_{t-2}.$$

4.  $L^i (L^j X_t) = L^i X_{t-j} = X_{t-i-j}$ , de même  $L^i (L^j X_t) = L^{i+j} X_t = X_{t-i-j}$ .
5.  $(L^i + L^j) X_t = L^i X_t + L^j X_t = X_{t-i} + X_{t-j}$ .
6.  $L^{-i} X_t = X_{t+i}$ .
7. si  $|c| < 1$ . la somme infinie  $(1 + cL + c^2 L^2 + c^3 L^3 + \dots) X_t = \frac{X_t}{(1-cL)}$ .

L'opérateur retard simplifie grandement l'écriture des équations relatives aux séries. Il permet d'écrire une équation de récurrence comme un polynôme de l'opérateur retard appliqué à une série.

# Chapitre 2

## Les modèles Vectoriels autorégressifs

Un modèle Vectoriel autorégressif (VAR) d'ordre ( $p$ ), dit VAR( $p$ ), est une généralisation des modèles AR( $p$ ) au cas multivarié. La dimension de ce processus représente le nombre de variables étudiées qui composent ce processus vectoriel et pour lequel on cherche à établir des relations linéaires entre les variables.

### 2.1 La représentation VAR( $p$ )

Un processus vectoriel  $X_t$  de dimension  $n$ , admet une représentation VAR d'ordre  $p$ , notée  $VAR(p)$  si : [1]

$$X_t = A_0 + A_1X_{t-1} + A_2X_{t-2} + \cdots + A_pX_{t-p} + \varepsilon_t.$$

De façon équivalente

$$A(L) X_t = A_0 + \varepsilon_t,$$

où  $A_0$  de dimension  $n$  désigne un vecteur de constantes,  $A(L) = \sum_{i=0}^p A_i L^i$ , où les matrices  $A_i, \forall i = 0, \dots, p$  de dimension  $(n, n)$ , satisfont  $A_0 = I_n$  et  $A_p \neq 0$ , alors :

$$A(L) = I_n - \sum_{i=1}^p A_i L^i = I_n - A_1 L - A_2 L^2 - \dots - A_p L^p.$$

(c'est le polynôme matriciel de dimension  $(n, n)$ )

où :

$$X_t = \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{n,t} \end{pmatrix}, \quad A_i = \begin{pmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \cdot & \cdot & \cdot & a_{1n}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & \cdot & \cdot & \cdot & a_{2n}^{(i)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1}^{(i)} & a_{n2}^{(i)} & \cdot & \cdot & \cdot & a_{nn}^{(i)} \end{pmatrix}, \quad A_0 = \begin{pmatrix} a_1^{(0)} \\ a_2^{(0)} \\ \cdot \\ \cdot \\ a_n^{(0)} \end{pmatrix}$$

$$\varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \cdot \\ \cdot \\ \varepsilon_{n,t} \end{pmatrix} \text{ est un vecteur de bruits blancs.}$$

**Modèle bivarié avec représentation VAR(2) :**

$$X_t = A_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t.$$

$$\Leftrightarrow \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} a_1^{(0)} \\ a_2^{(0)} \end{pmatrix} + \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} \end{pmatrix} \begin{pmatrix} X_{1,t-2} \\ X_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}.$$

Alors :

$$\begin{cases} X_{1,t} = a_1^{(0)} + a_{11}^{(1)} X_{1,t-1} + a_{12}^{(1)} X_{2,t-1} + a_{11}^{(2)} X_{1,t-2} + a_{12}^{(2)} X_{2,t-2} + \varepsilon_{1,t} \\ X_{2,t} = a_2^{(0)} + a_{21}^{(1)} X_{1,t-1} + a_{22}^{(1)} X_{2,t-1} + a_{21}^{(2)} X_{1,t-2} + a_{22}^{(2)} X_{2,t-2} + \varepsilon_{2,t} \end{cases}$$

Pour :

$$X_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix} + \begin{pmatrix} 0.7 & 0.4 \\ 1 & 0.2 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} 0.1 & 0.87 \\ 0.25 & 0.98 \end{pmatrix} \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix},$$

on a :

$$\begin{cases} x_t = 1 + 0.7x_{t-1} + 0.4y_{t-1} + 0.1x_{t-2} + 0.87y_{t-2} + \varepsilon_{1,t} \\ y_t = 5 + x_{t-1} + 0.2y_{t-1} + 0.25x_{t-2} + 0.98y_{t-2} + \varepsilon_{2,t} \end{cases}$$

## 2.2 Stationnarité d'un modèle VAR

On généralise la notion de stationnarité faible ou de second ordre vue au chapitre 1 au cas vectoriel. [2]

**Définition 2.1** *Un processus vectoriel  $(X_t)$  de dimension  $n$  est dit stationnaire au second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si :*

1.  $\forall t \in \mathbb{Z}, E(X_t^2) < \infty.$
2.  $\forall t \in \mathbb{Z}, E(X_t) = \mu_{(n,1)}$  (indépendant de  $t$ ).
3.  $\forall (t, h) \in \mathbb{Z}^2, E[(X_{t+h} - \mu)(X_t - \mu)'] = \gamma(h)_{(n,n)}$  (indépendant de  $t$ ).

Lorsque l'on considère un processus  $VAR(p)$ , on peut démontrer que ces conditions de stationnarité reviennent à imposer des conditions sur les racines du déterminant du polynôme matriciel  $A(L)$ .

**Proposition 2.1** *Un processus vectoriel  $(X_t)$  de dimension  $n$  satisfait une représentation  $VAR(p)$  telle que  $\forall t \in \mathbb{Z}$*

$$A(L)X_t = (I - A_1X_{t-1} - \dots - A_pX_{t-p}) = A_0 + \varepsilon_t,$$

*est stationnaire si et seulement si les racines du déterminant du polynôme matriciel  $A(L)$  noté  $\lambda_i, i \in [1, n]$ , sont toutes supérieures à l'unité en module.*

$$\det[A(\lambda_i)] = |I_n \lambda_i^p - A_1 \lambda_i^{p-1} - A_2 \lambda_i^{p-2} - \dots - A_{p-1} \lambda_i - A_p| = 0.$$

$$|\lambda_i| < 1, \forall i \in [1, n].$$

**Exemple 2.1** On considère un processus bivarie  $(Y_t)$  admet une représentation  $VAR(1)$  tel que :

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

On pose  $Y_t (y_{1,t}, y_{2,t})$  et  $\varepsilon_t (\varepsilon_{1,t}, \varepsilon_{2,t})$ , cette représentation peut s'exprimer sous forme :

$$A(L)Y_t = A_0 + \varepsilon_t,$$

$$A_0 = (3, 1) \text{ et}$$

$$\begin{aligned} A(L) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.2 & 0.7 \\ 0.3 & 0.4 \end{pmatrix} L \\ &= \begin{pmatrix} 1 - 0.2L & -0.7L \\ -0.3L & 1 - 0.4L \end{pmatrix}. \end{aligned}$$

On calcule  $\det(A(L))$  :

$$\det(A(L)) = (1 - 0.2L)(1 - 0.4L) - (-0.3L * 0.7L).$$

$$\det[A(L)] = 0 \iff 1 + 0.8L^2 - 0.6L - 0.21L^2 = 0 \iff 1 - 0.6L - 0.13L^2 = 0,$$



on a donc deux racines  $\lambda_1 = 1.3$  et  $\lambda_2 = -5.91$  sont supérieures à 1 en module alors le processus  $Y_t$  est stationnaire.

**Définition 2.2 (Bruit blanc vectoriel)** Soit  $\varepsilon_t$ , un processus  $n$  dimensionnel qui s'écrit :

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{n,t} \end{bmatrix}, \forall t \in \mathbb{Z}.$$

Ce processus est appelé un bruit blanc  $n$  vectoriel faible si :

1.  $E(\varepsilon_t) = 0, \forall t \in \mathbb{Z}$ .
2.  $E[\varepsilon_t \varepsilon'_s] = 0, \forall t \neq s \in \mathbb{Z}$ .
3.  $E(\varepsilon_t \varepsilon'_{t-h}) = \begin{cases} \Omega, & h = 0 \\ 0, & h \neq 0 \end{cases}$
4.  $E[\varepsilon_t \varepsilon'_t] = \Sigma_\varepsilon, \forall t \in \mathbb{Z}$ .  
où  $\Sigma_\varepsilon$  est la matrice de variances-covariances constantes des  $n$  composantes du processus  $\varepsilon_t$ .

## 2.3 Estimation des paramètres

Les paramètres du processus VAR ne peuvent être estimés que sur des séries chronologiques stationnaires. Il existe deux méthodes d'estimation des paramètres d'un modèle vectoriel autorégressif : estimation de chaque équation d'un VAR par la méthode des moindres carrés et estimation par la méthode du maximum de vraisemblance. [4]

## 2.4 Détermination du nombre de retards

Pour déterminer le nombre de retards optimal pour un  $VAR(p)$ , on peut utiliser plusieurs méthodes. 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  $h$  et ( $h$  nombre de retards maximum pour la taille de l'échantillon considéré) pour chacun de ces modèles, on calcule les fonction  $AIC(p)$  et  $SC(p)$  de la façon suivante :

$$AIC(p) = \ln [\det | \Sigma_e |] + \frac{2k^2p}{T}.$$

$$SC(p) = \ln [\det | \Sigma_e |] + \frac{k^2p \ln(T)}{T}.$$

Avec :  $T$  est le nombre d'observation,  $k$  le nombre de variable du système,  $p$  le nombre de retards,  $\Sigma_e$  matrice des variances covariances des résidus du modèle. [4]

## 2.5 Prévision

Etant donné une série stationnaire, observée entre 1 et  $T$ , on cherche à faire de prévision à horizon  $h$ , et donc de prévoir  $X_{T+1}, \dots, X_{T+h}$ . Il s'agit de calculer les prévisions optimales du modèle  $VAR$  estimé, à savoir  $X_T(h)$  la prévision de  $X_{T+h}$  sachant l'ensemble d'information disponible en  $T$ , noté : [6]

$$\hat{X}_{T+h} = E(X_{T+h} / X_T, X_{T-1}, \dots, X_1).$$

### 2.5.1 Le cas d'un VAR(1)

Considérons le cas d'un modèle  $VAR(1)$  centré tel que  $\forall t$  :

$$X_t = A_0 + A_1 X_{t-1} + \varepsilon_t.$$

Supposons que l'on dispose d'une réalisation sur un échantillon de  $T$  réalisations  $(X_1, X_2, \dots, X_T)$ .

La prévision de la réalisation à la date  $T + 1$  de  $X_t$  est donnée par :

$$\begin{aligned}\widehat{X}_{T+1} &= E(X_{T+1}/X_T) \\ &= E(A_0 + A_1 X_T + \varepsilon_{T+1}/X_T) \\ &= A_0 + A_1 X_T + E(\varepsilon_{T+1}/X_T) \\ &= A_0 + A_1 X_T.\end{aligned}$$

À l'horizon  $T + 2$ , on a :

$$\begin{aligned}\widehat{X}_{T+2} &= E(X_{T+2}/X_T) \\ &= E(A_0 + A_1 X_{T+1} + \varepsilon_{T+2}/X_T) \\ &= A_0 + A_1 E(X_{T+1}/X_T) + E(\varepsilon_{T+2}/X_T) \\ &= A_0 + A_1 \widehat{X}_{T+1} \\ &= A_0 + A_1 (A_0 + A_1 X_T) \\ &= (I + A_1) A_0 + A_1^2 X_T.\end{aligned}$$

**Proposition 2.2** *De même façon à un horizon  $h$ , la prévision d'un VAR(1) est donné par :*

$$\begin{aligned}\widehat{X}_{T+h} &= E(X_{T+h}/X_T) \\ &= (I + A_1 + A_1^2 + \dots + A_1^{h-1}) A_0 + A_1^h X_T.\end{aligned}$$

*L'erreur de prévision s'écrit sous la forme :*

$$\begin{aligned}X_{T+h} - \widehat{X}_{T+h} &= X_{T+h} - E(X_{T+h}/X_T) \\ &= \sum_{i=0}^{h-1} A_1^i \varepsilon_{T+h-i}.\end{aligned}$$

**Remarque 2.1** *Par la définition des bruits blanc, cette erreur de prévision à une espérance nulle.*

## 2.5.2 Le cas d'un VAR(p)

Le modèle s'écrit :

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + \varepsilon_t.$$

La prévision optimale à la date  $T + 1$ , faite à la date  $T$  est

$$\widehat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1).$$

Donc

$$\widehat{X}_{T+1} = A_1 X_T + \dots + A_p X_{T-p}.$$

De façons analogue

$$X_{T+h} = A_1 X_{T+h-1} + \dots + A_p X_{T+h-p} + \varepsilon_{T+h}.$$

Donc

$$\widehat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1).$$

De façons réccursive

$$\widehat{X}_{T+h} = \begin{cases} A_1 \widehat{X}_{T+h-1} + \dots + A_{h-1} \widehat{X}_{T+1} + A_h X_T + \dots + A_p X_{T+h-p} & \text{pour } h \leq p. \\ A_1 \widehat{X}_{T+h-1} + \dots + A_p \widehat{X}_{T+h-p} & \text{pour } h > p. \end{cases}$$

## 2.6 La causalité

La causalité entre deux séries temporelles (chronologique) est généralement étudiée en termes d'amélioration de la prévision. [2], [6]

### 2.6.1 Causalité au sens de Granger

**Définition 2.3** *On dit que la variable  $y_t$  cause au sens de Granger la variable  $x_t$  si la prédictibilité de  $x_t$  est améliorée lorsque l'information relative à  $y_t$  est incorporée dans l'analyse. Soit le modèle VAR( $p$ ) pour lequel les variables  $x_t$  et  $y_t$  sont stationnaires :*

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_1^{(0)} \\ a_2^{(0)} \end{pmatrix} + \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ x_{t-2} \end{pmatrix} + \dots \\ + \begin{pmatrix} a_{11}^{(p)} & a_{12}^{(p)} \\ a_{21}^{(p)} & a_{22}^{(p)} \end{pmatrix} \begin{pmatrix} y_{t-p} \\ x_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}.$$

**Définition 2.4**  $x_t$  cause  $z_t$  au sens de Granger à la date  $t$  si le fait de tenir compte des informations passées de  $x_t$  améliore la prévision de  $z_t$  à tout horizon de temps  $h \in \mathbb{N}$ , Mathématiquement :

$$E(z_{t+h}/z_t, z_{t-1}, \dots, z_1) = E(z_{t+h}/z_t, \dots, z_1, x_t, \dots, x_1).$$

Selon la représentation moyenne mobile canonique, on a

$$X_t = B_0 + \sum_{i=0}^{\infty} B_i \varepsilon_{t-i} = B_0 + B(L) \varepsilon_t, \quad t \in \mathbb{Z} \text{ où } B(0) = I \text{ et } L \text{ est l'opérateur retard.}$$

En partitionnant  $y_t$ , on obtient

$$y_t = \begin{pmatrix} z_t \\ x_t \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} + \begin{pmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}, \quad t \in \mathbb{Z}$$

Avec

la matrice  $\begin{pmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{pmatrix}$  est la matrice  $B(L)$  partitionnée selon les sous-processus  $z_t$  et  $x_t$  où  $\varepsilon_{1,t}$  et  $\varepsilon_{2,t}$  sont les vecteurs de bruits blancs correspondant aux sous processus  $z_t$  et  $x_t$ .

La variable  $x_t$  ne cause pas  $z_t$  au sens de Granger si et seulement si

$$B_{12}^{(1)} = B_{12}^{(2)} = B_{12}^{(3)} = \dots = B_{12}^{(i)} \Leftrightarrow A_{12}^{(1)} = A_{12}^{(2)} = A_{12}^{(3)} = \dots = A_{12}^{(i)} = 0.$$

$$B_{12}^{(i)} = 0, \forall i = 1, 2, \dots \Leftrightarrow A_{12}^{(i)} = 0, \forall i = 1, 2, \dots, p.$$

Alors

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_1^{(0)} \\ a_2^{(0)} \end{pmatrix} + \begin{pmatrix} a_{11}^{(1)} & 0 \\ a_{21}^{(1)} & a_{22}^{(1)} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} a_{11}^{(2)} & 0 \\ a_{21}^{(2)} & a_{22}^{(2)} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \dots \\ + \begin{pmatrix} a_{11}^{(p)} & 0 \\ a_{21}^{(p)} & a_{22}^{(p)} \end{pmatrix} \begin{pmatrix} y_{t-p} \\ x_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}.$$

La représentation  $MA$  est commode pour calculer la variance de l'erreur de prevision. Mais pour tester la non causalité, on aura besoin de la forme  $VAR(p)$  estimable.

## 2.6.2 Causalité au sens de Sims

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

Ceci se traduit par la représentation suivante :

$$y_{1,t} = a_1^{(0)} + \sum_{i=1}^p a_{1,i}^{(1)} y_{1,t-i} + \sum_{i=1}^p a_{1,i}^{(2)} y_{2,t-i} + \sum_{i=1}^p b_i^{(2)} y_{2,t+i} + \varepsilon_{1,t}.$$

$$y_{2,t} = a_2^{(0)} + \sum_{i=1}^p a_{2,i}^{(1)} y_{1,t-i} + \sum_{i=1}^p a_{2,i}^{(2)} y_{2,t-i} + \sum_{i=1}^p b_i^{(1)} y_{1,t+i} + \varepsilon_{2,t}.$$

**Remarque 2.2** 1. La vectoriel  $y_{1,t}$  ne cause pas  $y_{2,t}$  si :  $b_1^{(2)} = b_2^{(2)} = \dots = b_p^{(2)} = 0$ .

2. La vectoriel  $y_{2,t}$  ne cause pas  $y_{1,t}$  si :  $b_1^{(1)} = b_2^{(1)} = \dots = b_p^{(1)} = 0$ .

## 2.7 La 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. [7]

**Définition 2.5** On considère un processus  $X_t = (x_{1,t}, x_{2,t}, \dots, x_{N,t})$  de dimension  $(N, 1)$  intégré d'ordre  $d$ . Les processus  $(x_{i,t}, t \in \mathbb{Z})$  sont dite cointégrées si et seulement s'il existe un vecteur  $\beta = (\beta_1, \beta_2, \dots, \beta_N) \in \mathbb{R}^N$  est un vecteur de cointégration tel que la combinaison linéaire  $\beta X_t$  est stationnaire ou intégré d'ordre 0. Le vecteur  $\beta$  correspond à un vecteur de cointegration.



**Définition 2.6**  $X_t$  et  $Y_t$  sont dites cointégrées si les deux conditions sont vérifiées :

1.  $X_t$  et  $Y_t$  deux séries intégrées de même ordre  $d$ , ( $X_t \sim I(d)$ ,  $Y_t \sim I(d)$ ).
2. Il existe une combinaison linéaire de ces séries qui sont intégrées d'ordre strictement inférieur à  $d$  c'est-à-dire :

$$\begin{cases} X_t \rightarrow I(d) \\ Y_t \rightarrow I(d) \end{cases} \text{ Alors } \alpha X_t + \beta Y_t \sim I(d-b) \text{ avec } d \geq 0, b \geq 0.$$

Pour l'intégration, on note  $X_t \sim I(d)$ . Pour la cointégration, on note  $(X_t, Y_t) \sim CI(d, b)$ .

**Remarque 2.3** Soit une série  $X_t$  stationnaire et une série  $Y_t$  intégrée d'ordre 1 alors  $(X_t + Y_t)$  est intégrée d'ordre 1.

# Chapitre 3

## Application

Le tableau suivant présente l'évolution des Taxe sur l'activité professionnels (TAP) selon le registre du commerce (Rm) de la wilaya de ouargla au cours de la période 2003-2019.

séries\Années	2003	2004	2005	2006
<b>TAP</b>	1503035107	2487970876	2900253410	3841286135
<b>RM</b>	15803	16988	17909	20022
séries\Années	2007	2008	2009	2010
<b>TAP</b>	4646585742	5337426523	3693701907	8012377885
<b>RM</b>	20792	24268	24677	25491
séries\Années	2011	2012	2013	2014
<b>TAP</b>	10143913363	8971307003	7879447843	8448856467
<b>RM</b>	27188	28395	29971	31319
séries\Années	2015	2016	2017	2018
<b>TAP</b>	8983658681	10009054273	980770333	10315165840
<b>RM</b>	37443	38791	37620	36256
séries\Années	2019			
<b>TAP</b>	11715462843			
<b>RM</b>	41816			

TAB. 3.1 – L'évolution des Taxe sur l'activité professionnels (TAP) selon le registre du commerce (Rm) de la wilaya de ouargla au cours de la période 2003-2019

### 3.1 Test de la stationnarité

Nous commençons par identifier le graphique de chacune des deux séries.

Les figures 1 et 2 montrent la non stationnarité des deux séries.

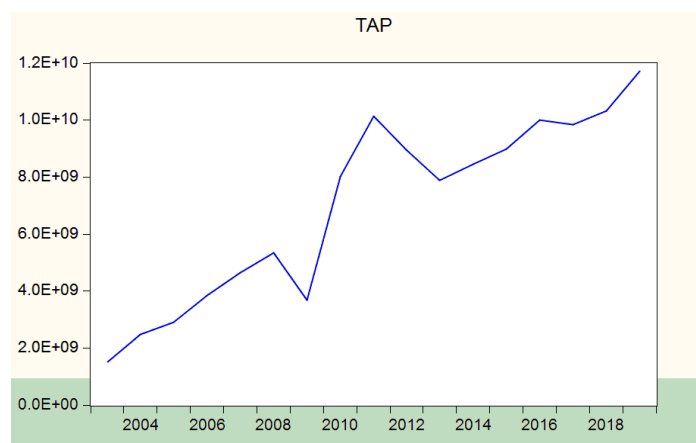


FIG. 3.1 – La série de Taxe sur l'activité professionnels (TAP)

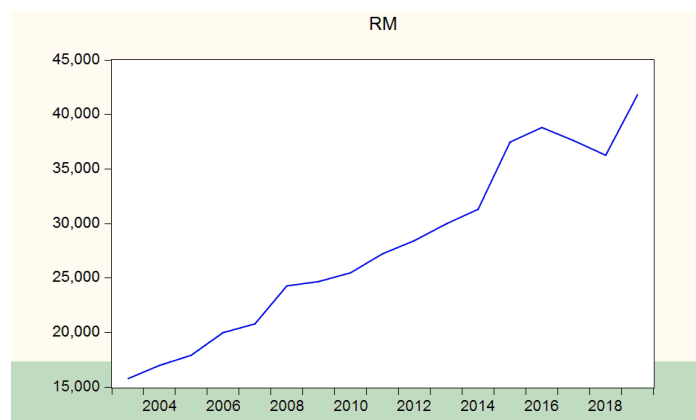


FIG. 3.2 – La série de le registre du commerce(Rm)

### 3.1.1 Test de la racine unitaire (ADF)

1) La série Taxe sur l'activité professionnels (TAP) :

Modèle sans tendance et avec constante :

Null Hypothesis: TAP has a unit root  
 Exogenous: Constant  
 Lag Length: 0 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.118884	0.6810
Test critical values:		
1% level	-3.920350	
5% level	-3.065585	
10% level	-2.673459	

\*MacKinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 16

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(TAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:23  
 Sample (adjusted): 2004 2019  
 Included observations: 16 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TAP(-1)	-0.128116	0.114503	-1.118884	0.2820
C	1.50E+09	8.40E+08	1.780435	0.0967
R-squared	0.082082	Mean dependent var		6.38E+08
Adjusted R-squared	0.016516	S.D. dependent var		1.39E+09
S.E. of regression	1.38E+09	Akaike info criterion		45.04208
Sum squared resid	2.66E+19	Schwarz criterion		45.13865
Log likelihood	-358.3366	Hannan-Quinn criter.		45.04702
F-statistic	1.251901	Durbin-Watson stat		2.092249
Prob(F-statistic)	0.282021			

La série a une racine unitaire ( $p\text{-value} > 0.05$ ) et la probabilité critique affectées à la constante est supérieurs à 0.05.

## Modèle avec tendance et avec constante :

Null Hypothesis: TAP has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 0 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.646971	0.2672
Test critical values:		
1% level	-4.667883	
5% level	-3.733200	
10% level	-3.310349	

\*MacKinnon (1996) one-sided p-values.

Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 16

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(TAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:25  
 Sample (adjusted): 2004 2019  
 Included observations: 16 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TAP(-1)	-0.686677	0.259420	-2.646971	0.0201
C	1.87E+09	7.50E+08	2.500887	0.0265
@TREND("2003")	3.95E+08	1.69E+08	2.332509	0.0364
R-squared	0.352899	Mean dependent var		6.38E+08
Adjusted R-squared	0.253344	S.D. dependent var		1.39E+09
S.E. of regression	1.20E+09	Akaike info criterion		44.81747
Sum squared resid	1.87E+19	Schwarz criterion		44.96233
Log likelihood	-355.5398	Hannan-Quinn criter.		44.82489
F-statistic	3.544792	Durbin-Watson stat		1.785502
Prob(F-statistic)	0.059064			

La série a une racine unitaire (la p-value > 0.05), la tendance et la constante sont significative (la probabilité critique < 0.05).

## Modèle sans tendance et sans constante :

Null Hypothesis: TAP has a unit root  
 Exogenous: None  
 Lag Length: 0 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
<b>Augmented Dickey-Fuller test statistic</b>	<b>1.150415</b>	<b>0.9276</b>
Test critical values: 1% level	-2.717511	
5% level	-1.964418	
10% level	-1.605603	

\*MacKinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 16

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(TAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:27  
 Sample (adjusted): 2004 2019  
 Included observations: 16 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TAP(-1)	0.057809	0.050251	1.150415	0.2680
R-squared	-0.125758	Mean dependent var		6.38E+08
Adjusted R-squared	-0.125758	S.D. dependent var		1.39E+09
S.E. of regression	1.47E+09	Akaike info criterion		45.12118
Sum squared resid	3.26E+19	Schwarz criterion		45.16947
Log likelihood	-359.9694	Hannan-Quinn criter.		45.12365
Durbin-Watson stat	2.046247			

La série a une racine unitaire (p-value>0.05).

D'après ce qu'on a vu précédemment, la série (TAP) est non stationnaire.

## 2) La série le registre du commerce (Rm)

### Modèle sans tendance et avec constante :

Null Hypothesis: RM has a unit root  
 Exogenous: Constant  
 Lag Length: 3 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.844280	0.9904
Test critical values: 1% level	-4.057910	
5% level	-3.119910	
10% level	-2.701103	

\*MacKinnon (1996) one-sided p-values.

Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 13

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(RM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:32  
 Sample (adjusted): 2007 2019  
 Included observations: 13 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RM(-1)	0.073888	0.087516	0.844280	0.4230
D(RM(-1))	-0.354822	0.286236	-1.239611	0.2502
D(RM(-2))	-0.759687	0.314819	-2.413091	0.0423
D(RM(-3))	-0.953502	0.389581	-2.447506	0.0401
C	2896.508	2458.370	1.178223	0.2726
R-squared	0.563941	Mean dependent var		1676.462
Adjusted R-squared	0.345911	S.D. dependent var		2218.707
S.E. of regression	1794.396	Akaike info criterion		18.10645
Sum squared resid	25758862	Schwarz criterion		18.32374
Log likelihood	-112.6919	Hannan-Quinn criter.		18.06179
F-statistic	2.586531	Durbin-Watson stat		2.182279
Prob(F-statistic)	0.117716			



## Modèle avec tendance et avec constante :

Null Hypothesis: RM has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 1 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.815012	0.0457
Test critical values:		
1% level	-4.728363	
5% level	-3.759743	
10% level	-3.324976	

\*Mackinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 15

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(RM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:34  
 Sample (adjusted): 2005 2019  
 Included observations: 15 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RM(-1)	-1.339072	0.351001	-3.815012	0.0029
D(RM(-1))	0.627178	0.313251	2.002157	0.0705
C	18044.00	4417.585	4.084585	0.0018
@TREND("2003")	2221.537	581.5119	3.820277	0.0028
R-squared	0.582911	Mean dependent var		1655.200
Adjusted R-squared	0.469159	S.D. dependent var		2067.200
S.E. of regression	1506.138	Akaike info criterion		17.69566
Sum squared resid	24952968	Schwarz criterion		17.88448
Log likelihood	-128.7175	Hannan-Quinn criter.		17.69365
F-statistic	5.124412	Durbin-Watson stat		2.146847
Prob(F-statistic)	0.018490			

## Modèle sans tendance et sans constante :

Null Hypothesis: RM has a unit root  
 Exogenous: None  
 Lag Length: 3 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	4.010551	0.9996
Test critical values: 1% level	-2.754993	
5% level	-1.970978	
10% level	-1.603693	

\*Mackinnon (1996) one-sided p-values.

Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 13

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(RM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:36  
 Sample (adjusted): 2007 2019  
 Included observations: 13 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RM(-1)	0.165373	0.041234	4.010551	0.0031
D(RM(-1))	-0.289366	0.286785	-1.008997	0.3393
D(RM(-2))	-0.755532	0.321517	-2.349896	0.0433
D(RM(-3))	-0.959760	0.397857	-2.412321	0.0391
R-squared	0.488273	Mean dependent var		1676.462
Adjusted R-squared	0.317697	S.D. dependent var		2218.707
S.E. of regression	1832.688	Akaike info criterion		18.11261
Sum squared resid	30228701	Schwarz criterion		18.28645
Log likelihood	-113.7320	Hannan-Quinn criter.		18.07688
Durbin-Watson stat	2.158442			

De la même manière que nous analysons les résultats précédents pour la première série, nous constatons que la série est non stationnaire.

Nous concluons que l'instationnarité des deux série est due au fait qu'elle contient une racine unitaire.

## 3.2 Stationnarisation des séries

Pour enlever la racine unitaire, nous allons stationnariser la série **TAP** en utilisant la première différence, ce qui conduit donc à la série **dTAP** :

$$dtap = tap - tap(-1).$$

On utilise aussi le test de racine unitaire (test ADF) :

1)

Null Hypothesis: DTAP has a unit root  
 Exogenous: Constant  
 Lag Length: 1 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.083801	0.0087
Test critical values:		
1% level	-4.004425	
5% level	-3.098896	
10% level	-2.690439	

\*Mackinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 14

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DTAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:40  
 Sample (adjusted): 2006 2019  
 Included observations: 14 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DTAP(-1)	-1.624465	0.397783	-4.083801	0.0018
D(DTAP(-1))	0.476499	0.267879	1.778789	0.1029
C	9.97E+08	4.43E+08	2.247549	0.0461
R-squared	0.646478	Mean dependent var		70572462
Adjusted R-squared	0.582201	S.D. dependent var		2.19E+09
S.E. of regression	1.42E+09	Akaike info criterion		45.17113
Sum squared resid	2.21E+19	Schwarz criterion		45.30807
Log likelihood	-313.1979	Hannan-Quinn criter.		45.15845
F-statistic	10.05772	Durbin-Watson stat		2.219057
Prob(F-statistic)	0.003283			

La série ne contient pas de racine unitaire car la p-value < 0.05.

2)

Null Hypothesis: DTAP has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 1 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.985729	0.0368
Test critical values: 1% level	-4.800080	
5% level	-3.791172	
10% level	-3.342253	

\*Mackinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 14

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DTAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:44  
 Sample (adjusted): 2006 2019  
 Included observations: 14 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DTAP(-1)	-1.670031	0.419003	-3.985729	0.0026
D(DTAP(-1))	0.501050	0.280218	1.788075	0.1041
C	1.55E+09	1.09E+09	1.418730	0.1864
@TREND("2003")	-55062007	99082965	-0.555716	0.5906
R-squared	0.657068	Mean dependent var		70572462
Adjusted R-squared	0.554189	S.D. dependent var		2.19E+09
S.E. of regression	1.47E+09	Akaike info criterion		45.28357
Sum squared resid	2.15E+19	Schwarz criterion		45.46616
Log likelihood	-312.9850	Hannan-Quinn criter.		45.26667
F-statistic	6.386771	Durbin-Watson stat		2.261845
Prob(F-statistic)	0.010845			

La série ne contient pas de racine unitaire (la p-value<0.05).

3)

Null Hypothesis: DTAP has a unit root  
 Exogenous: None  
 Lag Length: 0 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.421158	0.0021
Test critical values: 1% level	-2.728252	
5% level	-1.966270	
10% level	-1.605026	

\*Mackinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 15

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DTAP)  
 Method: Least Squares  
 Date: 05/26/22 Time: 19:47  
 Sample (adjusted): 2005 2019  
 Included observations: 15 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DTAP(-1)	-0.925236	0.270445	-3.421158	0.0041
R-squared	0.455245	Mean dependent var		27690749
Adjusted R-squared	0.455245	S.D. dependent var		2.12E+09
S.E. of regression	1.57E+09	Akaike info criterion		45.24544
Sum squared resid	3.43E+19	Schwarz criterion		45.29264
Log likelihood	-338.3408	Hannan-Quinn criter.		45.24493
Durbin-Watson stat	1.927068			

La série ne contient pas de racine unitaire ( $p\text{-value} < 0.05$ ). Donc la série est stationnaire selon les trois test de *Dickey Fuller*.

### Graphique de la série stationnaire :

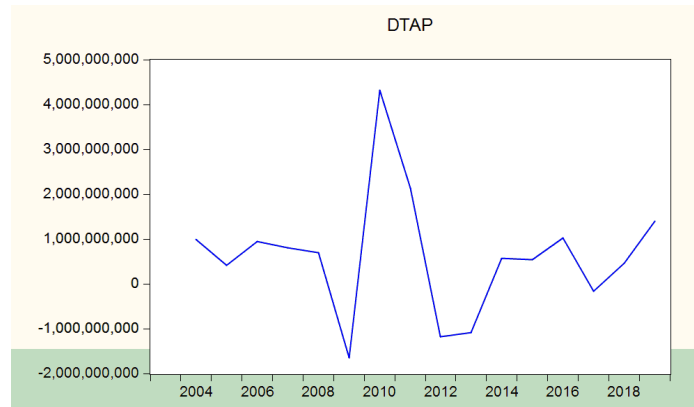


FIG. 3.3 – Représentation graphique de la série stationnaire (DTAP).

On remarque que la série fluctue autour d'une moyenne constante, donc la série est stationnaire.

Pour la stationnarité de la deuxième série ( $rm$ ), nous avons utilisé la méthode des différences secondes car la série reste non stationnaire aux différences premières, ce qui conduit donc à la série **ddrm** :

$$ddrm = drm - drm(-1).$$

On utilise aussi le test de racine unitaire (test ADF).

1)

Null Hypothesis: DDRM has a unit root  
 Exogenous: Constant  
 Lag Length: 2 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.881491	0.0030
Test critical values:		
1% level	-4.121990	
5% level	-3.144920	
10% level	-2.713751	

\*MacKinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 12

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DDRM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 20:09  
 Sample (adjusted): 2008 2019  
 Included observations: 12 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DDRM(-1)	-3.549652	0.727166	-4.881491	0.0012
D(DDRM(-1))	1.836888	0.552482	3.324791	0.0105
D(DDRM(-2))	0.955568	0.320847	2.978270	0.0176
C	51.97957	694.5891	0.074835	0.9422
R-squared	0.810611	Mean dependent var		688.9167
Adjusted R-squared	0.739591	S.D. dependent var		4639.025
S.E. of regression	2367.310	Akaike info criterion		18.63810
Sum squared resid	44833244	Schwarz criterion		18.79973
Log likelihood	-107.8286	Hannan-Quinn criter.		18.57825
F-statistic	11.41372	Durbin-Watson stat		2.115259
Prob(F-statistic)	0.002916			

La série ne contient pas de racine unitaire car la p-value  $< 0.05$ .



2)

Null Hypothesis: DDRM has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 2 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.557945	0.0186
Test critical values: 1% level	-4.992279	
5% level	-3.875302	
10% level	-3.388330	

\*Mackinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 12

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DDRM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 20:14  
 Sample (adjusted): 2008 2019  
 Included observations: 12 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DDRM(-1)	-3.731513	0.818683	-4.557945	0.0026
D(DDRM(-1))	1.962229	0.614185	3.194847	0.0152
D(DDRM(-2))	1.002108	0.343882	2.914100	0.0225
C	1412.056	2412.392	0.585334	0.5767
@TREND("2003")	-132.3389	223.8904	-0.591088	0.5730
R-squared	0.819615	Mean dependent var		688.9167
Adjusted R-squared	0.716537	S.D. dependent var		4639.025
S.E. of regression	2469.873	Akaike info criterion		18.75606
Sum squared resid	42701903	Schwarz criterion		18.95810
Log likelihood	-107.5363	Hannan-Quinn criter.		18.68125
F-statistic	7.951456	Durbin-Watson stat		2.116423
Prob(F-statistic)	0.009644			

Notre série ne contient pas de racine unitaire (la p-value < 0.05).

3)

Null Hypothesis: DDRM has a unit root  
 Exogenous: None  
 Lag Length: 2 (Automatic - based on SIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.252504	0.0001
Test critical values:		
1% level	-2.771926	
5% level	-1.974028	
10% level	-1.602922	

\*MacKinnon (1996) one-sided p-values.  
 Warning: Probabilities and critical values calculated for 20 observations  
 and may not be accurate for a sample size of 12

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(DDRM)  
 Method: Least Squares  
 Date: 05/26/22 Time: 20:18  
 Sample (adjusted): 2008 2019  
 Included observations: 12 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DDRM(-1)	-3.558143	0.677418	-5.252504	0.0005
D(DDRM(-1))	1.841593	0.517683	3.557375	0.0061
D(DDRM(-2))	0.956917	0.302125	3.167293	0.0114
R-squared	0.810479	Mean dependent var		688.9167
Adjusted R-squared	0.768363	S.D. dependent var		4639.025
S.E. of regression	2232.702	Akaike info criterion		18.47213
Sum squared resid	44864629	Schwarz criterion		18.59336
Log likelihood	-107.8328	Hannan-Quinn criter.		18.42725
Durbin-Watson stat	2.107202			

La p-value < 0.05, donc la série ne contient pas de racine unitaire.

D'après les résultats précédents, nous constatons que la série est stationnaire.

La représentation graphique de la série stationnarisée est donnée sur la figure ci-dessous.

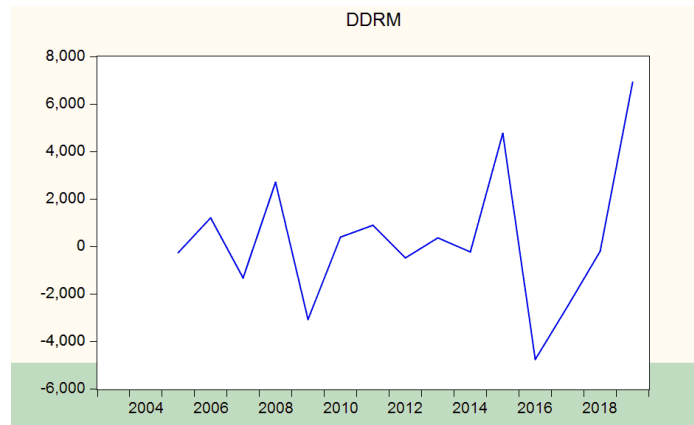


FIG. 3.4 – Représentation graphique de la série stationnaire (DDRM).

### 3.3 Détermination de degré de retard

Lag	LogL	LR	FPE	AIC	SC	HQ
1	-383.2418	1.137091	5.23e+25	64.87363	65.11609	64.78387
2	-380.7999	2.848845	7.40e+25	65.13332	65.53741	64.98371
3	-373.5522	6.039797	5.41e+25	64.59203	65.15775	64.38258

TAB. 3.2 – Le degré de retard

Nous remarquons à partir des résultats donnés que le degré de retard optimal est 1(c'est moins critérien)(Pour le degré 3, le modèle n'est pas valide).

## 3.4 Estimation du modèle

Comme la série a le degré de retard 1, nous estimons maintenant le modèle par la méthode Lag Order Selection (OLS), qui est la méthode la plus approuvée dans cette étude :

Vector Autoregression Estimates  
Date: 05/26/22 Time: 22:01  
Sample (adjusted): 2006 2019  
Included observations: 14 after adjustments  
Standard errors in ( ) & t-statistics in [ ]

	DTAP	DDRM
DTAP(-1)	-0.046508 (0.30177) [-0.15412]	1.17E-07 (6.1E-07) [0.19314]
DDRM(-1)	-153284.8 (187548.) [-0.81731]	-0.390334 (0.37797) [-1.03271]
C	6.28E+08 (4.5E+08) [1.38096]	194.6189 (916.122) [0.21244]
R-squared	0.066552	0.088468
Adj. R-squared	-0.103166	-0.077265
Sum sq. resids	2.69E+19	1.09E+08
S.E. equation	1.56E+09	3149.898
F-statistic	0.392131	0.533800
Log likelihood	-314.5549	-130.9488
Akaike AIC	45.36499	19.13554
Schwarz SC	45.50193	19.27248
Mean dependent	6.30E+08	331.3571
S.D. dependent	1.49E+09	3034.836
Determinant resid covariance (dof adj.)		2.32E+25
Determinant resid covariance		1.43E+25
Log likelihood		-445.2003
Akaike information criterion		64.45719
Schwarz criterion		64.73107
Number of coefficients		6

Alors notre modèle est :

$$\begin{aligned}
X_t &= A_0 + A_1 X_{t-1} + \varepsilon_t. \\
&= \begin{pmatrix} 6.28 * 10^8 \\ 194.6189 \end{pmatrix} + \begin{pmatrix} -0.046508 & 1.17 * 10^{-7} \\ -153284.8 & -0.390334 \end{pmatrix} X_{t-1} + \varepsilon_t.
\end{aligned}$$

## 3.5 La validation du modèle

### 3.5.1 Test de la stationnarité du modèle

Pour vérifier la stationnarité du modèle, nous utilisons le test des racines multiples, et si toutes les racines sont inférieures à une et se trouvent à l'intérieur du cercle, le modèle est stationnaire.

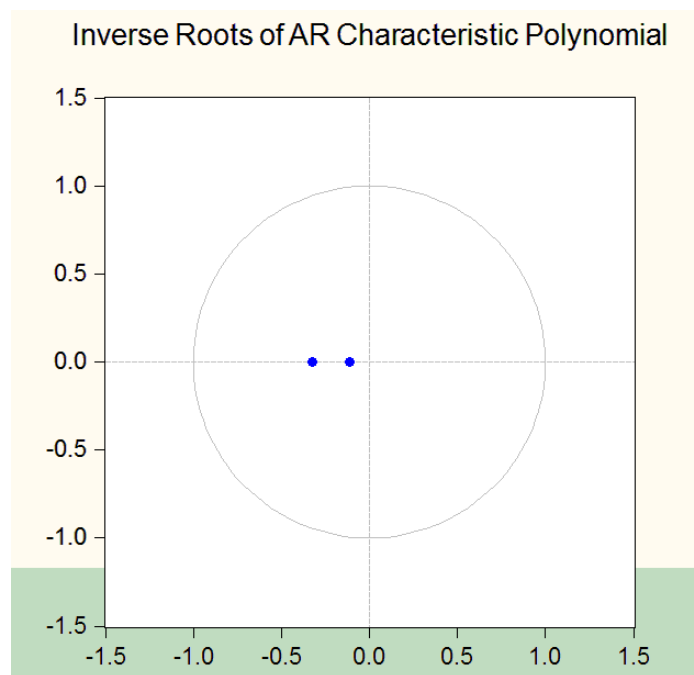


FIG. 3.5 – Les racines du polynôme caractéristique.

D'après ce graphe, la série est stationnaire.

## 3.6 Tests sur les résidus

### 3.6.1 Test de la normalité des résidus

Ce test est utilisé pour tester si les résidus sont de loi normale ou non.

VAR Residual Normality Tests  
Orthogonalization: Cholesky (Lutkepohl)  
Null Hypothesis: Residuals are multivariate normal  
Date: 05/26/22 Time: 22:35  
Sample: 2003 2019  
Included observations: 14

Component	Skewness	Chi-sq	df	Prob.*
1	0.512721	0.613394	1	0.4335
2	0.741740	1.283751	1	0.2572
Joint		1.897145	2	0.3873

Component	Kurtosis	Chi-sq	df	Prob.
1	2.776580	0.029118	1	0.8645
2	3.149420	0.013024	1	0.9091
Joint		0.042142	2	0.9791

Component	Jarque-Bera	df	Prob.
1	0.642512	2	0.7252
2	1.296775	2	0.5229
Joint	1.939287	4	0.7469

\*Approximate p-values do not account for coefficient estimation

Donc les probabilités de chaque résidu sont supérieures à (la p-value > 0.05), alors le résidu suit une distribution normale.

### 3.6.2 Autocorrélation test

Dans ce test, il faut savoir si les résidus sont autocorrélés ou non.

VAR Residual Portmanteau Tests for Autocorrelations  
 Null Hypothesis: No residual autocorrelations up to lag h  
 Date: 05/27/22 Time: 14:31  
 Sample: 2003 2019  
 Included observations: 14

Lags	Q-Stat	Prob.*	Adj Q-Stat	Prob.*	df
1	0.184009	---	0.190250	---	---
2	6.784325	0.1477	7.888526	0.0954	4

\*Test is valid only for lags larger than the VAR lag order.  
 df is degrees of freedom for (approximate) chi square distribution

VAR Residual Serial Correlation LM Tests  
 Date: 05/27/22 Time: 12:27  
 Sample: 2003 2019  
 Included observations: 14

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	3.410167	4	0.4917	0.888534	(4, 16.0)	0.4931
2	7.547345	4	0.1096	2.235482	(4, 16.0)	0.1109

Null hypothesis: No serial correlation at lags 1 to h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	3.410167	4	0.4917	0.888534	(4, 16.0)	0.4931
2	11.59189	8	0.1704	1.748709	(8, 12.0)	0.1846

\*Edgeworth expansion corrected likelihood ratio statistic.

D'après les résultats, nous remarquons que la p-value > 0.05 de chacun des deux tests, alors il n'y a pas d'autocorrélation des résidus.



### 3.6.3 Test d'hétéroscédasticité

Ce test fonctionne sur la connaissance de la stabilité de l'homogénéité de la variance et il est considéré comme l'une des étapes les plus importantes dans l'étude du modèle VAR :

VAR Residual Heteroskedasticity Tests (Levels and Squares)  
 Date: 05/27/22 Time: 11:06  
 Sample: 2003 2019  
 Included observations: 14

---

Joint test:

Chi-sq	df	Prob.
13.41396	12	0.3397

---

Individual components:

Dependent	R-squared	F(4,9)	Prob.	Chi-sq(4)	Prob.
res1*res1	0.533621	2.574403	0.1099	7.470695	0.1130
res2*res2	0.204829	0.579578	0.6851	2.867599	0.5802
res2*res1	0.383282	1.398342	0.3095	5.365942	0.2518

VAR Residual Heteroskedasticity Tests (Includes Cross Terms)  
 Date: 05/27/22 Time: 11:16  
 Sample: 2003 2019  
 Included observations: 14

---

Joint test:

Chi-sq	df	Prob.
18.00754	15	0.2623

---

Individual components:

Dependent	R-squared	F(5,8)	Prob.	Chi-sq(5)	Prob.
res1*res1	0.798983	6.359512	0.0113	11.18576	0.0478
res2*res2	0.226458	0.468407	0.7904	3.170409	0.6737
res2*res1	0.476655	1.457255	0.3024	6.673166	0.2461

On remarque que (la  $p\text{-value} > 0.05$ ) dans chacun des deux tests, alors l'homogénéité de la variance des résidus est stable.

### 3.7 La prévision

Nous obtenons les prévisions suivantes pour 5 années :

Année /La prévision	La prévision de TAP	La prévision de RM
<b>2020</b>	$1.23 * 10^{10}$	43247
<b>2021</b>	$1.28 * 10^{10}$	44680
<b>2022</b>	$1.33 * 10^{10}$	46104
<b>2023</b>	$1.38 * 10^{10}$	47511
<b>2024</b>	$1.42 * 10^{10}$	48901

TAB. 3.3 – Prévision.

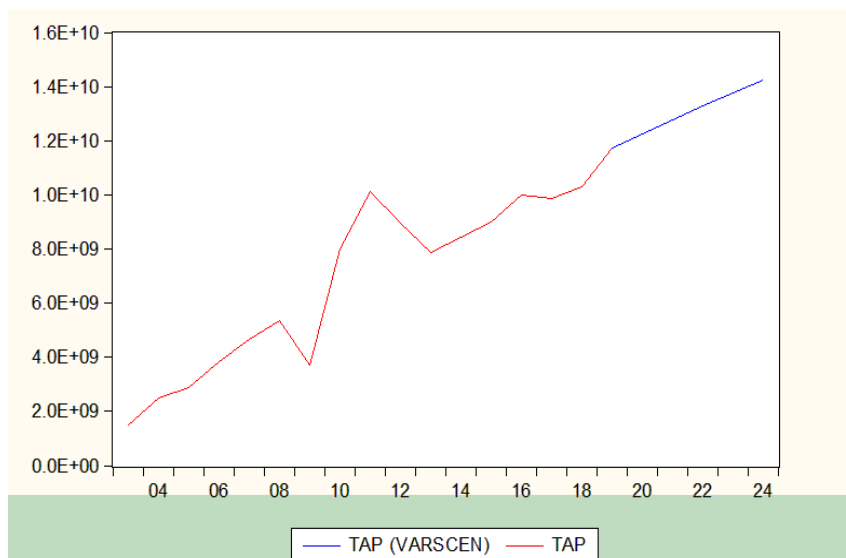


FIG. 3.6 – La prévision de TAP.

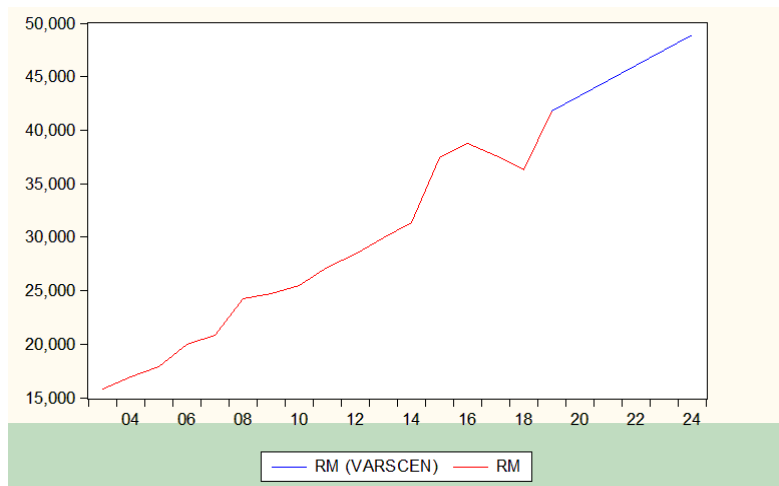


FIG. 3.7 – La prévision de RM.

# Conclusion

Dans ce mémoire, nous avons vu que le modèle VAR est un outil de modélisation de séries multivariées lorsque les variables à l'étude sont toutes pertinentes.

De la première partie du travail, on a effectué une revue des principales notions concernant les séries chronologiques.

Un avantage important dans la modélisation des séries chronologiques est de prévoir les valeurs futures. Dans ce cadre, nous avons vu le modèle vectoriels autorégressifs VAR(p).

Empiriquement, l'application de ce modèle nécessite la démarche suivante : (Le logiciel qui nous sert de pratique est : Eviews et pour modéliser notre série, on utilise les procédures : Déterminer le lag ou décalage optimal, Estimer le VAR et tester la validation des résultats obtenus (tests sur les résidus)).

Nous concluons que l'estimation et la validité du modèle pour les séries temporelles multivariées dépend de la réalisation de tous ses calibrages.

# Bibliographie

- [1] Agung, I. G. N. (2011). Time series data analysis using EViews.
- [2] Bourbonnais, R., (2015). Économétrie : Cours et exercices corrigés, 9 Edition, DUNOD, paris.
- [3] Cryer, J. D., & Chan, K. S. (2008). Time series analysis : with applications in R , Second Edition. New York : Springer.
- [4] Enders, W., (2015). Applied econometric time series. fourth edition.
- [5] Goldfarb, B., & Pardoux, C. (2011). Introduction à la méthode statistique, 6<sup>o</sup> édition. Dunod, paris.
- [6] Hafida, M. G., & Youcef, M.(??). Modèles linéaires et non linéaires pour des séries de taux de change et estimation non paramétrique de densité.
- [7] HAMLILI, A., (2018). Test of causality between the economics indexes in algeria (Doctoral dissertation).

- [8] Lagnoux, A., (2015). Renforcement Statistique Séries chronologiques. Université de Toulouse Le Mirail.
- [9] Prenat, M., Keribin, C., & Rossignol, R. (2010). Séries chronologiques. Université Paris-Sud.
- [10] Raoult, J. P., (2011). Note de lecture :«Séries temporelles avec R-Méthodes et cas». Statistique et Enseignement.
- [11] Rubenthaler, S., (2019). Séries chronologiques (avec R)(Cours et exercices).