

République Algérienne Démocratique et Populaire.
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.
Université Mouloud Mammeri, Tizi-Ouzou.
Faculté des Sciences.
Département de Mathématiques.



Spécialité : Mathématiques
Option : Probabilités et statistique
Mémoire de fin de cycle intitulé :

Modèles de Cointégration en Séries Chronologiques et Applications

Réalisé par :

MESMAGOU Fatma

Encadré par :

M^r FELLAG Hocine

Soutenu devant le jury d'examen composé de :

M^{me} ATIL Lynda,
M^r FELLAG Hocine,
M^{me} BELKACEM Cherifa,

Maître de conférence A, UMMTO,
Professeur, UMMTO,
Maître de conférence B, UMMTO,

Présidente.
Rapporteur.
Examinatrice.

Soutenu le : 03/07/2025

Dédicaces

Je dédie ce modeste travail :

À l'âme douce et bienveillante de ma grand-mère, dont l'amour silencieux et la sagesse profonde ont toujours été un refuge et une source d'inspiration, je dédie ce travail avec tout mon respect et ma gratitude.

À mes chers parents, piliers de ma vie et source inépuisable d'inspiration, votre amour, votre soutien indéfectible et vos sacrifices ont façonné la personne que je suis aujourd'hui. Je vous suis profondément reconnaissante pour l'éducation précieuse que vous m'avez donnée, qui guide chacun de mes pas. Vous demeurez à jamais gravés dans mon cœur avec une affection infinie.

À mes adorables sœurs : Sarah et Tinhinane, rayons de joie et d'encouragements constants, merci pour votre présence lumineuse, vos conseils avisés et votre confiance. Je vous souhaite un avenir éclatant, riche en bonheur et en réussites.

À mes chers frères : Yanis et Amar, compagnons de chaque instant, merci pour votre soutien, vos sourires et votre présence rassurante. Vous êtes ma force et mon inspiration, et je vous souhaite un avenir rempli de bonheur et de réussite.

À cette personne précieuse qui a su m'épauler dans les moments de doute, m'apporter force et réconfort, et croire en moi quand j'en avais le plus besoin. Ta présence discrète mais constante a été un véritable pilier tout au long de ce parcours, et je te dédie ce travail avec toute ma gratitude.

À mes amies Mélissa, kenza et Dalila, merci pour votre soutien indéfectible, vos rires partagés et votre présence précieuse qui ont illuminé ce parcours. Votre amitié est un trésor qui m'a porté dans les moments difficiles et célébré avec moi chaque victoire.

À mes chères camarades de la promotion Master 2 Probabilités et Statistique, (Soumeya, Lila, Rachida, Thanina, Hanane, Amel.K, Amel.L, Lisa, Lydia, Karima), vous avez été bien plus que des collègues d'études : vous êtes devenues une véritable famille, source de soutien, de partage et d'inspiration. Grâce à vous, cette aventure a été riche en apprentissages, en rires et en moments inoubliables. J'ai aimé chaque instant passé à vos côtés, et c'est avec une profonde gratitude que je vous dédie ce travail, fruit de notre parcours commun.

À toutes les personnes qui, de près ou de loin, ont contribué à cette étape de ma vie : votre soutien, vos encouragements et votre bienveillance ont été essentiels et m'ont permis de persévérer, recevez ici l'expression de ma reconnaissance la plus sincère.

Remerciements

Avant toute chose, je rends grâce au Bon Dieu, source de toute sagesse, de force et de lumière, qui m'a guidée et soutenue tout au long de ce parcours. Sans sa bienveillance et sa protection, rien de tout cela n'aurait été possible.

Ma profonde gratitude va à mon enseignant et promoteur, Monsieur FELLAG Hocine, qui a été bien plus qu'un simple guide. Sa confiance, sa patience et son soutien constant ont été des forces précieuses tout au long de mon parcours. Je lui suis sincèrement reconnaissante pour tout ce qu'il m'a apporté.

Je tiens à remercier profondément Madame ATIL et Madame BELKACEM, membres du jury, grâce à leur rigueur, leurs précieux conseils et leur soutien bienveillant, ce travail a pu prendre forme. Leur engagement et leur confiance ont été des appuis essentiels tout au long de cette étape.

Je tiens à exprimer ma profonde reconnaissance à l'ensemble des enseignants du département de mathématiques. Leur enseignement rigoureux, leur disponibilité et leur passion pour la discipline ont été des sources précieuses de savoir et de motivation tout au long de mon parcours.

Enfin, un très grand merci à toute ma famille qui m'a gratifié son amour et fourni les motivations. Je leur adresse toute ma gratitude du fond du cœur.

Table des matières

Introduction Générale

1	Notions générales sur les séries chronologiques	1
1.1	Introduction	1
1.2	Définitions et propriétés	1
1.2.1	Série chronologique	1
1.2.2	Composantes d'une série chronologique	2
1.2.3	Processus stochastique	3
1.2.4	Processus Bruit Blanc	3
1.2.5	Opérateur retard	4
1.2.6	Opérateur différence	4
1.2.7	La stationnarité	4
1.2.8	La non stationnarité	6
1.2.9	Les autocorrélations et la densité spectrale	7
1.2.10	Processus autorégressif d'ordre p	9
1.3	Tests de la racine unitaire	11
1.3.1	Test de Dickey-Fuller (DF)	12
1.3.2	Test de Dickey-Fuller augmenté (ADF)	15
1.4	Autres tests	20
1.4.1	Test de Phillips-Perron (PP)	20
1.4.2	Test d'Elliott, Rothenberg et Stock (ERS)	22
1.4.3	Test de Kwiatkowski, Phillips, Schmidt et Shin(KPSS)	24
2	cointégration linéaire et non linéaire	25
2.1	Introduction	25
2.2	Les tests de cointégration	27
2.2.1	Test d'Engle et Granger (1987)	27
2.2.2	Exemple R pour la méthode d'Engle et Granger (1987)	28
2.2.3	Test de Johansen (1988)	30

2.2.4	Test de Phillips-Ouliaris	31
2.3	Modèle à correction d'erreur	32
2.4	Autres approches multivariées	33
2.4.1	Cointégration linéaire basée sur l'approche ARDL	33
2.4.2	Modèles ARDL	34
2.4.3	Test de Cointégration aux Bornes (Bounds-test)	36
2.5	Cointégration non linéaire	40
2.6	Modeles à modèles à correction d'erreur non linéaires (NECM)	41
2.6.1	Tests de cointégration non linéaire	42
2.6.2	Test basé sur les rangs (Range Unit Root (RUR))	42
2.6.3	Test RCC (Record Counting Cointegration)	43
2.6.4	Test basé sur les statistiques d'ordre induit (IOC)	43
2.7	Approche Paramétrique	43
2.7.1	Méthode d'estimation	44
2.8	Approche Non Paramétrique	44
2.8.1	Fonction de transfert	45
2.9	Review sur la cointégration non linéaire	45
3	Simulations et applications	50
2.1	Introduction	50
3.2	Une application financière de la cointégration linéaire	51
3.3	Une application de la cointégration non linéaire : Tendances linéaire versus une tendance quadratique	53
3.3.1	Processus gaussien	54
3.3.2	Étude de Monte Carlo	57
3.3.3	La démarche par processus gaussien	61
3.3.4	Application au pont de référence Z24	63
3.4	Arbitrage statistique et trading financier	64
3.4.1	Simulation d'un fonds de couverture avec Modèle AR(1)	65
	Conclusion générale	68

Introduction Générale

L'analyse des séries temporelles constitue un pilier fondamental de l'économétrie moderne, trouvant des applications dans des domaines aussi variés que la finance, la macroéconomie ou l'ingénierie. Cette méthodologie permet non seulement de comprendre les dynamiques sous-jacentes des données chronologiques, mais aussi d'élaborer des prévisions robustes et d'identifier des relations structurelles entre variables. Dans ce mémoire, nous nous intéressons particulièrement à trois aspects complémentaires de l'analyse des séries temporelles : les tests de stationnarité, la cointégration sous ses formes linéaire et non linéaire, et leurs applications pratiques. L'importance croissante de ces méthodes dans l'analyse économique contemporaine s'explique par leur capacité à modéliser des phénomènes dynamiques complexes tout en fournissant des outils statistiques rigoureux pour tester des hypothèses théoriques.

La question de la stationnarité se pose en préambule à toute analyse sérieuse de séries temporelles. Une série stationnaire voit ses propriétés statistiques - moyenne, variance et autocovariances - invariantes dans le temps, ce qui facilite considérablement sa modélisation. Cependant, de nombreuses séries économiques et financières présentent des tendances ou des ruptures qui remettent en cause cette hypothèse. Les tests de racine unitaire, tels que ceux de Dickey-Fuller augmenté (ADF) ou Phillips-Perron (PP), permettent de détecter ces non-stationnarités et de déterminer si une différenciation est nécessaire. Le choix du test approprié n'est pas neutre, car chacun présente des spécificités en termes de puissance et de robustesse face à différents types de non-stationnarité. Par exemple, le test ADF montre une meilleure performance pour les séries avec tendance déterministe, tandis que le test PP est plus robuste face à l'hétéroscédasticité. Ces considérations techniques revêtent une importance cruciale dans les applications empiriques où le mauvais choix de test peut conduire à des conclusions erronées.

Au-delà de la stationnarité individuelle des séries, la cointégration offre un cadre théorique puissant pour analyser les relations entre variables non stationnaires. Le concept, introduit par Granger (1981) puis formalisé par Engle et Granger (1987), permet d'identifier des combinaisons linéaires stationnaires entre séries intégrées de même ordre. Cette propriété revêt une importance particulière en économétrie, car elle correspond souvent à des relations d'équilibre de long terme entre variables économiques. Les tests traditionnels de cointégration, comme ceux d'Engle-Granger ou de Johansen, reposent cependant sur l'hypothèse forte de linéarité, ce qui peut s'avérer restrictif dans de nombreuses situations réelles. La reconnaissance de cette limitation a motivé le développement d'approches alternatives capables de capturer des dynamiques plus complexes.

En effet, les relations économiques présentent fréquemment des non-linéarités : ajustements asymétriques, effets de seuil, changements de régime ou dynamiques dépendantes de l'état du système. Ces caractéristiques ont motivé le développement de méthodes de cointégration non linéaire, capables de capturer des relations plus complexes entre va-

riables. Les travaux pionniers de Granger et Hallman (1991) sur les transformations non linéaires de séries intégrées, ou ceux de Dufrénot et Mignon (2012) sur les modèles à correction d'erreur non linéaires, ont ouvert la voie à une littérature riche et en plein essor. Ces approches étendent considérablement le champ d'application de la cointégration, tout en posant des défis méthodologiques substantiels en termes d'estimation et de test. Parmi les développements récents les plus prometteurs, on peut citer les modèles à seuil qui permettent de capturer des asymétries dans les mécanismes d'ajustement, ou les approches markoviennes qui modélisent les changements de régime dans les relations de long terme.

L'utilité de ces concepts théoriques se révèle pleinement dans leurs applications pratiques. En finance, les stratégies d'arbitrage statistique s'appuient souvent sur des relations de cointégration pour identifier des opportunités de trading. Ces stratégies exploitent la propriété de retour à la moyenne des séries cointégrées pour générer des profits lorsque les écarts par rapport à l'équilibre deviennent trop importants. En macroéconomie, la cointégration permet de modéliser des relations d'équilibre entre variables comme la consommation et le revenu, offrant ainsi un cadre rigoureux pour tester des théories économiques fondamentales. Plus récemment, des domaines comme le monitoring structural (SHM) ont adopté ces méthodes pour distinguer les effets environnementaux des véritables dommages matériels dans les infrastructures. Ces applications concrètes témoignent de la vitalité et de la pertinence opérationnelle de l'analyse des séries temporelles dans des contextes variés.

Ce mémoire se structure autour de trois chapitres qui développent successivement ces différentes dimensions. Le premier chapitre établit les fondements théoriques des tests de stationnarité. Nous y présentons les définitions essentielles des processus stochastiques, en insistant particulièrement sur les concepts de stationnarité faible et forte. La question cruciale des tests de racine unitaire est abordée à travers une présentation des principales méthodes : Dickey-Fuller (DF), Dickey-Fuller augmenté (ADF), Phillips-Perron (PP), Elliott-Rothemberg-Stock (ERS) et KPSS. Cette progression reflète notre volonté de concilier rigueur théorique et pertinence pratique, tout en mettant en lumière les développements les plus récents de la littérature. Le deuxième chapitre approfondit le concept de cointégration, depuis ses fondements linéaires jusqu'aux développements non linéaires les plus récents. Nous commençons par exposer la théorie traditionnelle de la cointégration linéaire, telle que formulée par Engle et Granger (1987), en mettant l'accent sur sa relation avec les modèles à correction d'erreur. Les tests classiques (Engle-Granger, Johansen) sont présentés avec leurs forces et limites. La seconde partie du chapitre est consacrée aux extensions non linéaires, qui permettent de modéliser des relations plus complexes entre variables économiques. Nous y discutons notamment des travaux de Dufrénot et Mignon (2012) sur les modèles à correction d'erreur non linéaires. Le troisième chapitre illustre concrètement les aspects pratiques de la cointégration linéaire et non linéaire à travers quelques études de simulations et applications basées sur des données réelles. Nous débutons par une étude de cas sur la relation entre la livre sterling et l'euro avant le Brexit, qui permet de montrer l'apport de la cointégration linéaire dans l'analyse des parités de change. Cette application est complétée par deux analyses mettant en œuvre des méthodes non linéaires : la première, inspirée des travaux de Cross et Worden (2011), concerne la surveillance de l'intégrité structurale (SHM) et montre comment la cointégration peut distinguer les effets environnementaux des dommages matériels. Le cas du pont Z24 sert d'exemple concret pour cette application technique. La seconde analyse explore une stratégie d'arbitrage statistique en finance, où nous modélisons un portefeuille suivant un processus AR(1) pour évaluer une règle de trading basée sur le retour à

la moyenne.

Ces différentes applications démontrent la polyvalence des outils de cointégration, qu'ils soient linéaires ou non linéaires, pour identifier des relations stables entre variables dans des contextes variés. Elles soulignent également l'importance du choix méthodologique approprié en fonction des caractéristiques des données étudiées et des questions de recherche posées. Le passage progressif des fondements théoriques aux cas pratiques permet d'appréhender tant la rigueur conceptuelle que la valeur opérationnelle de ces méthodes.

L'apport principal de ce travail réside dans sa double perspective, à la fois théorique et appliquée, offrant au lecteur une compréhension complète des enjeux contemporains de l'analyse des séries temporelles. Les développements récents sur la cointégration non linéaire y occupent une place centrale, reflétant leur importance croissante dans la littérature économétrique. Les applications choisies couvrent des domaines variés (finance, technologie, macroéconomie) pour mieux illustrer la généralité des méthodes présentées.

En conclusion, cette étude montre comment l'analyse des séries temporelles, et particulièrement les méthodes de cointégration, permettent de modéliser des relations complexes tout en fournissant des outils opérationnels pour la prévision et la prise de décision. Les perspectives de recherche incluent l'intégration de méthodes d'apprentissage automatique pour l'identification des relations non linéaires, ainsi que l'extension de ces approches à des données de plus haute fréquence, particulièrement pertinentes en finance de marché.

Chapitre 1

Notions générales sur les séries chronologiques

1.1 Introduction

L'analyse d'une série temporelle consiste à étudier la stationnarité, et l'analyse de cette dernière fait appel à des tests formels s'appuyant soit sur la représentation graphique de la série temporelle analysée, ou bien sur l'analyse de son corrélogramme qui permet de faire une hypothèse sur cette stationnarité, soit en mobilisant des tests statistiques qu'on appelle des tests de racine unitaire .

Ce premier chapitre établit les fondements théoriques et méthodologiques nécessaires à l'analyse des séries temporelles. Nous rappelons d'abord les caractéristiques clés des séries chronologiques, en insistant sur les concepts de stationnarité et de non-stationnarité. La seconde partie est consacrée à une étude comparative des cinq principaux tests de racine unitaire : le test de Dickey-Fuller (DF), son extension augmentée (ADF), le test Phillips-Perron (PP), la procédure Elliott-Rothenberg-Stock (ERS) et le test KPSS. Chacun de ces tests présente des spécificités en termes d'hypothèses, de puissance et de robustesse aux différentes formes de non-stationnarité.

1.2 Définitions et propriétés

1.2.1 Série chronologique

Définition 1.2.1 . *On appelle série chronologique, chronique ou encore série temporelle une suite $\{x_1, x_2, \dots, x_n\}$ d'observations enregistrées à un temps spécifique t . Le nombre n est appelé la longueur de la série. Cette dernière peut être journalière, mensuelle, trimestrielle ou annuelle.*

Mathématiquement, une série chronologique est définie comme étant une trajectoire d'un processus aléatoire noté $\{X_t, t \in \mathbb{T}\}$, où \mathbb{T} est appelé espace de temps qui peut être discret ou continu ([43]).

Remarque 1.2.1 *Les dates d'observations sont généralement ordonnées de manière régulière dans le temps. On distingue deux types de séries chronologiques :*

- *Série continue : C'est une série où l'observation est faite d'une manière continue dans le temps. A titre d'exemple : résultat d'un électrocardiogramme.*

- *Série discrète* : C'est une série ou l'observation est faite sur des intervalles de temps fixés a priori. Dans ce cas $\mathbb{T} = \mathbb{Z}$ et les dates d'observations sont plus souvent équidistantes .

1.2.2 Composantes d'une série chronologique

Une série X_t est composée des quatre parties comme suit :

- a) **La tendance** : ou trend notée f_t .Il s'agit d'un terme de la série qui traduit l'évolution à moyen terme du phénomène. Elle sera estimée sous forme paramétrique (linéaire, polynomiale, logarithmique, exponentielle,... etc) .

- **Tendance linéaire** : La tendance la plus simple est dite linéaire .Dans ce cas on peut estimer les paramètres par la méthode des moindres carrés c'est une régression linéaire simple

$$f_t = a + bt$$

- **Tendance quadratique** : La forme de la tendance quadratique est

$$f_t = a + bt + ct^2$$

- **Tendance polynomiale d'ordre q** : On peut ajuster la série par un polynôme d'ordre q. Les paramètres peuvent être estimés par la méthode des moindres carrés .Dans ce cas on a une régression linéaire simple avec q variables explicatives

$$f_t = b_0 + b_1t + b_2t^2 + \dots + b_qt^q$$

- **Tendance logistique** : la fonction logistique permet de modéliser des processus ne pouvant dépasser une certaine valeur c

$$f_t = \frac{c}{1 + be^{-at}}$$

Par exemple la figure 1.1 ci-dessous montre l'évolution du fret aérien des ADP (Aéroports de Paris) entre janvier 82 et décembre 2007 (en millier de tonnes, source INSEE (Institut national de la statistique et des études économiques)). Elle illustre bien les limites d'un ajustement de la tendance par une fonction. Alors que le trend semblait linéaire pendant 10 ans, celui-ci s'est modifié et apparaît maintenant comme une évolution exponentielle (courbe rouge)... jusqu'au jour où ce modèle sera à son tour invalidé!

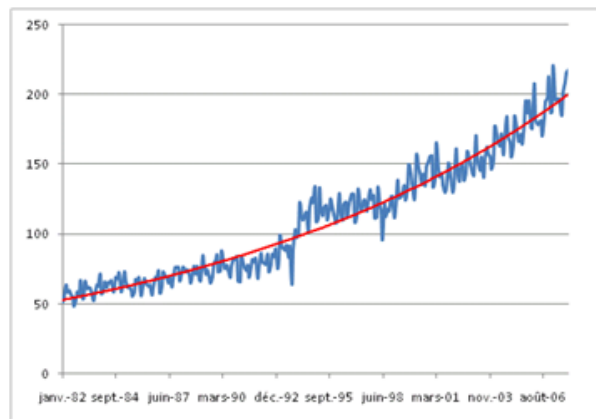


FIGURE 1.1 – L'évolution du fret aérien des ADP

- b) **La composante saisonnière** : Elle représente des effets périodiques de période connue p qui se reproduisent de façon plus ou moins identique d'une période à l'autre. Elle est notée S_t et elle est généralement supposée périodique d'une période p .

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

- c) **La composante résiduelle** : bruit ou résidu notée ϵ_t , correspond à des fluctuations irrégulières. En général de faible intensité mais de nature aléatoire.
- d) **La composante cyclique** : notée C_t , c'est une succession de mouvement persistant des variations de mouvement ascendant (période de prospérité) et de mouvement descendant (période de dépression). Pour de longues séries, un mouvement cyclique peut se superposer à la tendance.

Remarque 1.2.2 *Des phénomènes accidentels (grèves, conditions météorologiques exceptionnelles, crash financier) peuvent notamment intervenir dans les composantes d'une série chronologique.*

1.2.3 Processus stochastique

On appelle processus stochastique, ou processus aléatoire, un phénomène dont l'évolution dans le temps est imprévisible, c'est-à-dire qu'il est impossible, même en connaissant la position de la variable au temps t , de prédire avec exactitude sa position au temps $t + \Delta t$

1.2.4 Processus Bruit Blanc

On appelle bruit toute variation imprévisible d'une quantité dans le temps. Le bruit blanc est le plus aléatoire des bruits : dans ce cas, il n'existe aucune corrélation entre les accroissements successifs de la quantité. Autrement dit, la fonction d'autocorrélation (à définir plus loin) est nulle.

Définition 1.2.2 *On dit que le processus $\{\epsilon_t; t \in \mathbb{Z}\}$ est un bruit blanc faible (processus de hasard pur) s'il vérifie :*

1. $\mathbb{E}(\epsilon_t) = 0, \quad \forall t \in \mathbb{Z}$
2. $\mathbb{E}(\epsilon_t^2) = \sigma^2, \quad \forall t \in \mathbb{Z}$
3. $\text{cov}(\epsilon_t, \epsilon_s) = \mathbb{E}(\epsilon_t \epsilon_s) = 0, \quad \forall t \neq s \in \mathbb{Z}$

Si de plus les variables ϵ_t et ϵ_s sont indépendantes, on parle d'un bruit blanc fort.

1.2.5 Opérateur retard

Définition 1.2.3 On appelle opérateur retard L (Lag) ou B (Backward) l'opérateur qui associe à un processus $(X_t)_{t \in \mathbb{Z}}$ tel que

$$L : X_t \rightarrow L(X_t) = LX_t = X_{t-1}. \quad (1.2.1)$$

Cet opérateur est

- Linéaire
- inversible : son inverse est $L^{-1} = F$ tel que $FX_t = X_{t+1}$ et est appelé l'opérateur avance De plus on a :

$$L^p X_t = X_{t-p}$$

Soit le polynôme d'opérateurs L noté

$$A(L) = a_0 L^0 + a_1 L^1 + \dots + a_p L^p = \sum_{k=0}^p a_k L^k \quad (1.2.2)$$

et

$$\left(\sum_{k=0}^p a_k L^k \right) X_t = \sum_{k=0}^p a_k X_{t-k} \quad (1.2.3)$$

1.2.6 Opérateur différence

On définit l'opérateur Δ_d de différence d'ordre d comme l'opérateur linéaire tel que :

$$\Delta_d Y_t = Y_t - Y_{t-d} = (1 - B^d) Y_t$$

On peut aussi prendre l'opérateur d'ordre un et l'appliquer plusieurs fois :

$$\Delta^2(Y_t) = \Delta(\Delta(Y_t)) = \Delta(Y_t - Y_{t-1}) = (1 - B)(Y_t - Y_{t-1}) = Y_t - 2Y_{t-1} + Y_{t-2}$$

Ces opérateurs peuvent être utilisés pour transformer un processus de moyenne non nulle en un processus de moyenne nulle. On peut aussi s'en servir pour enlever la composante saisonnière de la série.

1.2.7 La stationnarité

La notion stationnarité joue un rôle capital dans la théorie des processus aléatoires surtout en analyse des séries chronologiques. Généralement on considère deux types de stationnarités, la stationnarité stricte et la stationnarité faible.

Définition 1.2.4 (Stationnarité stricte) Un processus $\{X_t, t \in \mathbb{Z}\}$ est dit strictement stationnaire si la distribution conjointe de la suite $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ est la même que celle de la suite $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ pour tout $n \in \mathbb{N}$ et tout $h \in \mathbb{Z}$. Cela peut être formalisé comme suit :

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = F_{t_1+h, t_2+h, \dots, t_n+h}(x_1, x_2, \dots, x_n) \quad (1.2.4)$$

On peut aussi dire que la structure probabiliste d'un processus strictement stationnaire est invariante en effectuant un changement d'origine.

Remarque 1.2.3 *En pratique, il est souvent très difficile d'assurer la stationnarité stricte, ce qui explique pourquoi on se contente généralement d'évaluer ou de supposer la stationnarité faible.*

Définition 1.2.5 (Stationnarité faible) *Une série temporelle faiblement stationnaire, est un processus à variance finie tel que :*

1. $E(X_t) = \mu$ où μ est une constante indépendante de t .
2. $cov(X_t, X_{t-h}) = \gamma(h)$ (ne dépend que de h) ([59]).

Dorénavant, nous utiliserons le terme stationnaire pour signifier une stationnarité faible ; si un processus est stationnaire au sens strict, nous utiliserons le terme strictement stationnaire.

Exemple 1.2.1 • *Par définition un processus bruit blanc faible $(\varepsilon_t) \sim bb(0, \sigma_\varepsilon^2)$ est faiblement stationnaire.*

- *Un processus linéaire $X_t = \sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}$, où $\sum_{i=-\infty}^{+\infty} \psi_i^2 < +\infty$ est stationnaire du second ordre car :*

1.

$$E(X_t) = E\left(\sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}\right) = \sum_{i=-\infty}^{+\infty} \psi_i E(\varepsilon_{t-i}) = 0 < +\infty.$$

2.

$$\begin{aligned} E(X_t^2) &= Var(X_t) = E\left[\left(\sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i}\right)^2\right] \\ &= E\left(\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-j}\right) \\ &= E\left(\sum_{i=-\infty, i=j}^{+\infty} \psi_i^2 \varepsilon_{t-i}^2 + \sum_{i=-\infty, i \neq j}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-j}\right) \\ &= E\left(\sum_{i=-\infty}^{+\infty} \psi_i^2 \varepsilon_{t-i}^2\right) = \sum_{i=-\infty}^{+\infty} \psi_i^2 E(\varepsilon_{t-i}^2) = \sigma_\varepsilon^2 \sum_{i=-\infty}^{+\infty} \psi_i^2 < +\infty \end{aligned}$$

3.

$$\begin{aligned}
 \text{Cov}(X_t, X_{t+h}) &= E(X_t X_{t+h}) - E(X_t)E(X_{t+h}) = E(X_t X_{t+h}) \\
 &= E \left[\left(\sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i} \right) \left(\sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t+h-j} \right) \right] \\
 &= E \left[\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t+h-j} \right] \\
 &= E \left[\sum_{\substack{i=-\infty \\ i=j-h}}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-(j-h)} + \sum_{\substack{i=-\infty \\ i \neq j-h}}^{+\infty} \psi_i \psi_j \varepsilon_{t-i} \varepsilon_{t-(j-h)} \right] \\
 &= E \left[\sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h} \varepsilon_{t-i}^2 \right] = \sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h} E(\varepsilon_{t-i}^2) \\
 &= \sigma_\varepsilon^2 \sum_{i=-\infty}^{+\infty} \psi_i \psi_{i+h}.
 \end{aligned}$$

1.2.8 La non stationnarité

Un processus dont l'espérance et la variance ne se stabilisent pas simultanément au cours du temps est qualifié de non stationnaire. On peut distinguer deux phénomènes à l'origine de la non stationnarité d'un processus, selon sa nature déterministe ou stochastique.

a) La non stationnarité déterministe

Définition 1.2.6 On dit que le processus $(X_t)_{t \in \mathbb{N}}$ est caractérisé par une non stationnarité déterministe, ou encore que le processus $(X_t)_{t \in \mathbb{N}}$ est TS (Trend stationary) s'il peut s'écrire :

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

où f est une fonction qui dépend du temps et Y_t est un processus stationnaire.

b) La non stationnarité stochastique

Définition 1.2.7 On dit que le processus X_t est caractérisé par une non stationnarité stochastique, ou encore que le processus X_t est DS si le processus différencié une fois $(1 - L)X_t$ est stationnaire.

On parle aussi de processus intégré d'ordre 1 : $(1 - L)X_t = Z_t$ est stationnaire ce qui implique $X_t = X_{t-1} + Z_t$.

De manière générale, on dit que le processus X_t est un processus intégré d'ordre d si le processus différencié d fois $(1 - L)^d X_t$ est stationnaire.

1.2.9 Les autocorrélations et la densité spectrale

Fonction d'autocovariance

Définition 1.2.8 Soit $\{X_t, t \in \mathbb{Z}\}$ un processus tel que $V(X_t) < +\infty$ pour tout $t \in \mathbb{Z}$. La fonction d'autocovariance $\gamma_X(r, s)$ de X_t est définie par :

$$\gamma_X(r, s) = \text{cov}(X_s, X_r) = E[(X_r - E(X_r))(X_s - E(X_s))], \quad r, s \in \mathbb{Z} \quad (1.2.5)$$

Dans le cas d'un processus stationnaire, la fonction d'autocovariance sera une fonction d'une seule variable :

$$\gamma_X(h) = \gamma_X(h, 0) = \text{cov}(X_t, X_{t+h}), \quad \forall t, h \in \mathbb{Z}$$

Propriétés

La fonction d'autocovariance d'un processus stationnaire vérifie :

1. $\forall h \in \mathbb{Z}, \gamma(-h) = \gamma(h)$ elle est paire.
2. $\forall n \in \mathbb{N}, \forall (a_i)_{1 \leq i \leq n} \in \mathbb{R}^n, \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(t_i - t_j) \geq 0$ elle est positive .
3. $\gamma(0) = \text{Var}(X_t)$
4. $|\gamma(h)| \leq \gamma(0) \quad \forall h \in \mathbb{Z}$

Remarque 1.2.4 La fonction d'autocovariance peut être normalisée et la nouvelle fonction obtenue est appelée fonction d'autocorrélation.

Fonction d'autocorrélation

La fonction d'autocorrélation mesure la corrélation entre les termes d'une série chronologique. De façon général la corrélation est une version standardisée de la covariance , spécifiquement la corrélation entre deux variables X_1 et X_2 est donnée par

$$\text{cor}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sigma_{X_1} \sigma_{X_2}}$$

où σ_{X_1} et σ_{X_2} sont les écarts types de X_1 et X_2 .

Définition 1.2.9 On appelle fonction d'autocorrélation d'un processus stationnaire $\{X_t; t \in \mathbb{Z}\}$ la fonction ρ définie de \mathbb{Z} dans \mathbb{R} par :

$$\forall h \in \mathbb{Z}, \rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

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

Propriétés

La fonction d'autocorrélation d'un processus stationnaire vérifie les propriétés suivantes :

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

Remarque 1.2.5 La fonction $\rho(h)$ est l'expression du lien linéaire entre X_t et X_{t+h} qui veut dire le lien entre le présent et le passé d'ordre h .

Fonction d'autocorrélation partielle

Définition 1.2.10 *L'autocorrélation partielle désigne la corrélation entre X_t et X_{t-h} obtenue lorsque l'influence des variables X_{t-h-i} (avec $0 < i < h$) a été retirée. On note $\rho(h)$ et ϕ_{hh} les fonctions respectivement d'autocorrélation et d'autocorrélation partielle de X_t au retard h .*

Soit ρ_h la matrice symétrique formée des $(h-1)$ premières autocorrélations de $\{X_t, t \in T\}$.

$$\rho_h = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{h-1} \\ \rho_1 & 1 & & & \rho_{h-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{h-3} \\ \vdots & & & \ddots & \\ \rho_{h-1} & \rho_{h-2} & & & 1 \end{pmatrix} \quad \rho_h^* = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_1 \\ \rho_1 & 1 & & & \rho_2 \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_3 \\ \vdots & & & \ddots & \\ \rho_{h-1} & \rho_{h-2} & & & \rho_n \end{pmatrix}$$

où $\phi_{hh} = \frac{|\rho_h^*|}{|\rho_h|}$

$|\rho_h^*|$ = le déterminant de la matrice (ρ_h) dans laquelle on remplace la dernière colonne par le vecteur $[\rho_1, \dots, \rho_h]$

Densité spectrale

Définition 1.2.11 *La densité spectrale d'un processus aléatoire stationnaire est donnée par :*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) \exp(-i\lambda h) \quad \forall \lambda \in \mathbb{R},$$

où $\gamma(h)$ est la fonction d'autocovariance du processus.

propriété

La densité spectrale est une fonction paire, positive, continue et périodique de période 2π .

Remarque 1.2.6 *La fonction d'autocovariance $\gamma(h)$ s'écrit :*

$$\begin{aligned} \gamma(h) &= \text{Cov}(X_t, X_{t-h}) \\ &= \int_{-\pi}^{\pi} f(\lambda) \cos(\lambda h) d\lambda \\ &= \int_{-\pi}^{\pi} f(\lambda) \exp(i\lambda h) d\lambda. \end{aligned}$$

Exemple 1.2.2 *Un bruit blanc (ε_t) est caractérisé par :*

$$\begin{cases} \gamma(0) = \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, & \text{pour } h = 0 \\ \gamma(h) = 0, & \text{pour } h \neq 0 \end{cases}$$

alors sa densité spectrale est donnée par :

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}.$$

1.2.10 Processus autorégressif d'ordre p

Les processus autorégressifs ont été introduits par Yule en 1927([22]).

Définition 1.2.12 *Un processus (X_t) est dit autorégressif d'ordre p ($AR(p)$) si l'observation présente X_t est générée par une moyenne pondérée des observations passées jusqu'à la p -ième période, sous la forme suivante :*

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad \forall t \in \mathbb{Z} \quad (1.2.6)$$

où :

- $\phi_1, \phi_2, \dots, \phi_p$ sont les coefficients autorégressifs.
- $\{\varepsilon_t\}$ est un bruit blanc .

Remarque 1.2.7 *Un processus autorégressif d'ordre p ($AR(p)$) peut s'écrire de la manière suivante :*

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

En utilisant l'opérateur retard L , où $LX_t = X_{t-1}$, on peut réécrire cette équation sous la forme :

$$X_t = \phi_1 LX_t + \phi_2 L^2 X_t + \dots + \phi_p L^p X_t + \varepsilon_t$$

En factorisant X_t , on obtient :

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) X_t = \varepsilon_t$$

Cela peut être noté sous la forme d'un polynôme retard :

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

où $\Phi(L)$ est le polynôme retard défini par :

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

Théorème 1.2.1 *Un $AR(p)$ existe et est stationnaire au sens faible si toutes les racines z_i de l'équation caractéristique :*

$$1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$$

vérifient la condition $|z_i| > 1 \quad \forall i$ (en dehors du cercle unité)

Exemple 1.2.3

$$X_t = 0.6X_{t-1} - 0.09X_{t-2} + \varepsilon_t \sim AR(2)$$

Alors

$$X_t - 0.6X_{t-1} + 0.09X_{t-2} = \varepsilon_t$$

$$(1 - 0.6L + 0.09L^2) X_t = \varepsilon_t$$

donc

$$\Phi(z) = 1 - 0.6z + 0.09z^2$$

$$\Phi(z) = (1 - 0.3z)^2$$

$$z_1 = z_2 = \frac{1}{0.3}$$

$$|z_1| = |z_2| = \left| \frac{1}{0.3} \right| > 1,$$

alors X_t est stationnaire.

Remarque 1.2.8 Une racine unitaire se produit lorsque le polynôme caractéristique a une racine $z = 1$, ce qui signifie que :

$$1 - \phi_1 - \phi_2 - \dots - \phi_p = 0 \quad \Rightarrow \quad \sum_{i=1}^p \phi_i = 1$$

C'est une condition nécessaire mais pas suffisante pour avoir une racine unitaire. Pour éviter une racine unitaire, il faut que cette somme ne soit pas exactement égale à 1, mais cela ne garantit pas que toutes les racines soient hors du cercle unité.

La fonction auto-covariance et auto-corrélation d'un AR(p)

Trouvons la fonction d'autocovariance pour un AR(p)

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

on a

$$\Phi(L)X_t = \varepsilon_t \implies X_t = \Phi^{-1}(L)\varepsilon_t$$

Donc

$$E(X_t) = E(\Phi^{-1}(L))E(\varepsilon_t) = 0$$

Fonction d'auto-covariance

$$\begin{aligned} \gamma_X(h) &= \text{Cov}(X_t, X_{t-h}) \\ &= E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) \\ &= E(X_t X_{t-h}) \quad \text{pour } h \geq 0 \end{aligned}$$

pour $h = 0$

$$\begin{aligned} \gamma_X(0) &= \text{Cov}(X_t, X_t) \\ &= E(X_t^2) - \underbrace{E(X_t)}_{=0} \\ &= E(X_t X_t) \\ &= \phi_1 E(X_t X_{t-1}) + \phi_2 E(X_t X_{t-2}) + \dots + \phi_p E(X_t X_{t-p}) + E(X_t \varepsilon_t) \\ &= \sum_{i=1}^p \phi_i \gamma_X(i) + \sigma_\varepsilon^2 \end{aligned}$$

Pour $h \geq 1$

$$\begin{aligned}
 X_t X_{t-h} &= \phi_1 X_{t-1} X_{t-h} + \cdots + \phi_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h} \\
 \gamma_X(h) &= \sum_{i=1}^p \phi_i \gamma_X(h-i) + E(\varepsilon_t X_{t-h}) \\
 &= \sum_{i=1}^p \phi_i \gamma_X(h-i) \quad (\text{car } E(\varepsilon_t X_{t-h}) = 0 \text{ pour } h \geq 1) \\
 \gamma_X(h) &= \begin{cases} \sum_{i=1}^p \phi_i \gamma_X(i) + \sigma_\varepsilon^2 & \text{si } h = 0 \\ \sum_{i=1}^p \phi_i \gamma_X(h-i) & \text{si } h \geq 1 \end{cases} \quad (1.2.7)
 \end{aligned}$$

Le système (1.2.7) s'appelle le système de Yule-Walker.

La fonction d'autocorrélation

$$\begin{aligned}
 \rho_X(h) &= \frac{\gamma_X(h)}{\gamma_X(0)} \\
 &= \sum_{i=1}^p \alpha_i \frac{\gamma_X(h-i)}{\gamma_X(0)} \\
 &= \sum_{i=1}^p \alpha_i \rho_X(h-i), \quad \forall h \geq 1.
 \end{aligned}$$

1.3 Tests de la racine unitaire

La stationnarité est étroitement liée à la localisation des racines de l'équation caractéristique du polynôme autorégressif (AR).

Si toutes les racines se trouvent à l'extérieur du disque unitaire c'est-à-dire si leur module est strictement supérieur à 1 ($|z_i| > 1$), le processus est stationnaire, en revanche si une ou plusieurs racines sont à l'intérieur du disque unitaire, le processus devient non stationnaire.

Les racines sur le disque unitaire (module égal à 1) indiquent souvent la présence d'une racine unitaire, caractéristique des processus intégrés comme les marches aléatoires, où la variance augmente avec le temps.

Pour détecter ces cas, des tests de racine unitaire, tels que le test de Dickey-Fuller ou le test de Phillips-Perron (PP), sont utilisés. Dans ce qui suit nous allons présenter le test classique de Dickey et Fuller dans les deux cas suivant :

- **Cas 1** : Le test de Dickey et Fuller simple concerne un AR(1).
- **Cas 2** : Le test de Dickey et Fuller augmenté concerne un AR(p) avec ($p \geq 2$)

1.3.1 Test de Dickey-Fuller (DF)

Le test de Dickey-Fuller (DF) a été introduit en 1979 ([13]) dans sa version première, il est fondé sur l'estimation par la méthode des moindres carrés ordinaires, il permet de détecter la présence de la racine unitaire dans un processus autorégressif du premier ordre AR(1). La construction de ce test est basée sur trois modèles autorégressifs du premier ordre dont les erreurs sont identiquement et indépendamment distribuées (iid).

- Modèle [1] : $X_t = \Phi X_{t-1} + \varepsilon_t$ modèle autorégressif d'ordre 1 .
- Modèle [2] : $X_t = c + \Phi X_{t-1} + \varepsilon_t$ modèle autorégressif avec constante .
- Modèle [3] : $X_t = c + \beta t + \Phi X_{t-1} + \varepsilon_t$ modèle autorégressif avec constante et tendance .

Où $\varepsilon_t \sim iid(0, \sigma^2)$ et $|\Phi| \leq 1$. Ceci correspond au modèle de base (modèle (1)) utilisé dans les tests de Dickey-Fuller (DF). Il est bien établi que dans le cas d'un processus non stationnaire, $\phi = 1$ (alors X_t est I(1)). À l'inverse, l'hypothèse de stationnarité implique $|\phi| < 1$ (X_t est I(0)).

L'approche DF repose sur le test d'hypothèses suivant :

$$\left\{ \begin{array}{l} H_0 : \Phi = 1 \text{ processus non stationnaire, les formes de non stationnarité sont :} \\ \quad [1] X_t = \Phi X_{t-1} + \varepsilon_t \\ \quad [2] X_t = \Phi X_{t-1} + c + \varepsilon_t \\ \quad [3] X_t = \Phi X_{t-1} + \beta t + c + \varepsilon_t \text{ où } \Phi = 1 \text{ et } \varepsilon_t \sim iid(0, \sigma^2) \\ H_1 : |\Phi| < 1. \end{array} \right.$$

On soustrait X_{t-1} des deux côtés on obtient :

$$\left\{ \begin{array}{l} H_0 : \Phi = 1 \text{ processus non stationnaire, les formes de non stationnarité sont :} \\ \quad [1] \Delta X_t = (\Phi - 1)X_{t-1} + \varepsilon_t \\ \quad [2] \Delta X_t = (\Phi - 1)X_{t-1} + c + \varepsilon_t \\ \quad [3] \Delta X_t = (\Phi - 1)X_{t-1} + \beta t + c + \varepsilon_t \text{ où } (\Phi - 1) = 0 \text{ et } \varepsilon_t \sim iid(0, \sigma^2) \\ H_1 : |\Phi| < 1 \end{array} \right.$$

La statistique de test

Dickey et Fuller ont proposé deux statistiques de test, la première est la statistique de Student usuelle mais qui a des valeurs critiques différentes de la valeur critique d'une loi asymptotique normale, que l'on va noter τ_Φ , : Sous H_0 vraie, la statistique de test pour l'estimateur de Φ est donnée par :

$$\tau_\Phi = \frac{\hat{\Phi} - 1}{\hat{\sigma}_\Phi} \tag{1.3.1}$$

Où

- $\hat{\sigma}_\Phi$ désigne un estimateur de l'écart-type de l'estimateur des moindres carrés ordinaires (OLS) de Φ .
- $\hat{\Phi}$ est l'estimateur de Φ .

La seconde statistique est basée sur le fait que $Z = T(\hat{\Phi} - 1)$ converge en distribution vers une fonctionnelle de Browniens qui ne dépend pas du paramètre de nuisance σ^2 . L'écriture de cette statistique sous $H_0 : \Phi = 1$ est :

$$Z = \frac{T^{-1} \sum X_{t-1} \varepsilon_t}{T^{-1} \sum X_{t-1}^2} \xrightarrow{L} \frac{0.5(W^2(1) - 1)}{\int_0^1 W^2(r) dr} \quad (1.3.2)$$

Avec

- $W(r)$ est un mouvement Brownien standard.
- T est la taille de l'échantillon à simuler.
- Et la distribution de τ sous H_0 est :

$$\tau_{\Phi} \xrightarrow{L} \frac{0.5(W^2(1) - 1)}{\sqrt{(\int_0^1 W^2(r) dr)}} \quad (1.3.3)$$

Remarque 1.3.1 *la statistique du test de Dickey-Fuller est applicable pour les trois modèles principaux du test ,et les valeurs critiques tabulées sont données par Fuller (1976)([23]), ainsi que par Dickey et Fuller (1981)([14]) .*

Définition 1.3.1 *Un processus stochastique $\{W_t, t \geq 0\}$ est un mouvement Brownien (ou processus de Wiener) de volatilité σ^2 si :*

1. $W_0 = 0$.
2. W_t suit une loi normale de moyenne 0 et de variance $\sigma^2 t$ ($N(0, \sigma^2 t)$) .
3. W_t est un processus à accroissement stationnaire et indépendants.

Remarque 1.3.2 *Si $\{W_t, t \geq 0\}$ est de moyenne 0 et de variance $\sigma^2 = 1$ ($N(0, \sqrt{t})$) on dit alors que c'est un mouvement Brownien standard .*

- La Table 1.1 présente les valeurs critiques pour les tests de Dickey-Fuller sans tendance, distinguant entre le modèle sans constante (D_{ρ}) et celui avec constante (D_{τ}).

TABLE 1.1 – Tests de Dickey-Fuller avec et sans constante

Valeur critique à	1%	5%	10%
Test D_{ρ}			
$T = 25$	-17.2	-12.5	-10.2
$T = 50$	-18.9	-13.3	-10.7
$T = 100$	-19.8	-13.7	-11.0
$T = 250$	-20.3	-14.0	-11.2
$T = \infty$	-20.7	-14.1	-11.3
Test D_{τ}			
$T = 25$	-3.75	-3.00	-2.63
$T = 50$	-3.58	-2.93	-2.60
$T = 100$	-3.51	-2.89	-2.58
$T = 250$	-3.46	-2.88	-2.57
$T = \infty$	-3.43	-2.86	-2.57

Chapitre 1 : Notions générales sur les séries chronologiques

- La Table 1.2 présente les valeurs critiques pour les tests de Dickey-Fuller incluant à la fois une constante et une tendance linéaire.

TABLE 1.2 – Tests de Dickey-Fuller avec constante et tendance

Valeur critique à	1%	5%	10%
Test D_ρ			
$T = 25$	-22.5	-17.9	-15.6
$T = 50$	-25.7	-19.8	-16.8
$T = 100$	-27.4	-20.7	-17.5
$T = 250$	-28.4	-21.3	-18.0
$T = \infty$	-29.5	-21.8	-18.3
Test D_τ			
$T = 25$	-4.38	-3.60	-3.24
$T = 50$	-4.15	-3.50	-3.18
$T = 100$	-4.04	-3.45	-3.15
$T = 250$	-3.99	-3.43	-3.13
$T = \infty$	-3.96	-3.41	-3.12

L'application du test de Dickey-Fuller simple se fait comme suit :

Étape 1

Estimation du modèle (3) :

$$X_t = \Phi X_{t-1} + \beta t + c + \varepsilon_t$$

- Pour le modèle (3), deux hypothèses nulles sont considérées. La première est : :
 - $H_0 : \Phi = 1$ et $c = \beta = 0$
 - $H_0 : \Phi = 1$ et $\beta = 0$
- Tests individuels :
- Test d'hypothèse $H_0 : \beta = 0$ contre $H_1 : \beta \neq 0$ (test de Student)
 - Si $\tau < k_1$ ou $\tau > k_2$: on rejette H_0 , on dit alors que X_t est un processus potentiellement DS
 - Sinon on passe à étape 2
- Test d'hypothèse $H_0 : \Phi = 1$ contre $H_1 : |\Phi| < 1$:
 - Si $\tau_{\hat{\Phi}} < \tau_{\text{tabulée}}$: on rejette H_0 (processus stationnaire)
 - Si $\tau_{\hat{\Phi}} > \tau_{\text{tabulée}}$: on accepte H_0 (processus TS)

Étape 2

Estimation du modèle (2) :

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

- Test d'hypothèse $H_0 : c = 0$ contre $H_1 : c \neq 0$:

- Si $\tau < \tau_{\text{tabulée}}$ ou $\tau > \tau_{\text{tabulée}}$: on rejette H_0 : test de racine unitaire sur Φ
- Sinon on passe à l'étape 3
- Test d'hypothèse $H_0 : \Phi = 1$ contre $H_1 : |\Phi| < 1$:
 - Si $\tau_{\hat{\Phi}} > \tau_{\text{tabulée}}$ on rejette H_0 donc le processus est non stationnaire (différenciation)
 - Si $\tau_{\hat{\Phi}} < \tau_{\text{tabulée}}$ on alors que X_t est stationnaire

Étape 3

Estimation du modèle (1) :

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

- Test d'hypothèse $H_0 : \Phi = 1$ contre $H_1 : |\Phi| < 1$:
 - Si $\tau_{\hat{\Phi}} < \tau_{\text{tabulée}}$ on rejette H_0 donc le processus est non stationnaire
 - Si $t_{\hat{\Phi}} < t_{\text{tabulé}}$: stationnaire

1.3.2 Test de Dickey-Fuller augmenté (ADF)

Le test de Dickey-Fuller augmenté (ADF) a été développé pour pallier une limitation majeure du test Dickey-Fuller original (1979). Alors que ce dernier ne s'appliquait qu'à des processus AR(1), l'ADF élargit ce cadre en incorporant des termes retardés des différences de la série, permettant ainsi de traiter des processus AR(p) plus généraux. Cette extension vise spécifiquement à corriger le problème d'autocorrélation des résidus, fréquent dans les données économiques réelles, tout en conservant la capacité à détecter la présence de racine unitaire. L'ADF représente ainsi une avancée méthodologique cruciale pour l'analyse des séries temporelles non stationnaires.

Modèles du Test ADF

Le test ADF inclut trois modèles principaux, similaires à ceux du test original de Dickey-Fuller mais avec des termes autorégressifs supplémentaires pour capturer l'autocorrélation des erreurs.

1) Modèle avec constante et tendance

$$\Delta X_t = \rho X_{t-1} + \alpha + \beta t + \sum_{j=1}^p \delta_j \Delta X_{t-j} + \epsilon_t \quad (1.3.4)$$

2) Modèle avec constante

$$\Delta X_t = \rho X_{t-1} + \alpha + \sum_{j=1}^p \delta_j \Delta X_{t-j} + \epsilon_t \quad (1.3.5)$$

3) Modèle sans constante ni tendance

$$\Delta X_t = \rho X_{t-1} + \sum_{j=1}^p \delta_j \Delta X_{t-j} + \epsilon_t \quad (1.3.6)$$

Démarche de la Différenciation

Considérons un processus autorégressif d'ordre p :

$$X_t = \alpha + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \epsilon_t$$

On peut définir :

$$\nu_t = \Phi_2 X_{t-2} + \Phi_3 X_{t-3} + \dots + \Phi_p X_{t-p} + \epsilon_t$$

En remplaçant dans l'équation précédente :

$$X_t = \alpha + \Phi_1 X_{t-1} + \nu_t$$

Pour illustrer la démarche avec un processus AR(3) :

$$X_t = \alpha + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \Phi_3 X_{t-3} + \epsilon_t$$

Pour les termes X_{t-2} et X_{t-3} , on écrit :

$$\begin{cases} X_{t-2} = X_{t-1} - \Delta X_{t-1} \\ X_{t-3} = X_{t-2} - \Delta X_{t-2} = X_{t-1} - \Delta X_{t-1} - \Delta X_{t-2} \end{cases}$$

En remplaçant dans le modèle AR(3), on obtient :

$$\begin{aligned} X_t &= \alpha + \Phi_1 X_{t-1} \\ &\quad + \Phi_2 (X_{t-1} - \Delta X_{t-1}) \\ &\quad + \Phi_3 (X_{t-1} - \Delta X_{t-1} - \Delta X_{t-2}) + \epsilon_t \end{aligned}$$

Après simplification :

$$X_t = \alpha + (\Phi_1 + \Phi_2 + \Phi_3) X_{t-1} - (\Phi_2 + \Phi_3) \Delta X_{t-1} - \Phi_3 \Delta X_{t-2} + \epsilon_t$$

En soustrayant X_{t-1} des deux côtés :

$$\Delta X_t = \alpha + \underbrace{(\Phi_1 + \Phi_2 + \Phi_3 - 1)}_{\rho} X_{t-1} - \underbrace{(\Phi_2 + \Phi_3)}_{\delta_1} \Delta X_{t-1} - \underbrace{\Phi_3}_{\delta_2} \Delta X_{t-2} + \epsilon_t$$

On obtient la forme standard du test ADF :

$$\Delta X_t = \alpha + \rho X_{t-1} + \sum_{j=1}^2 \delta_j \Delta X_{t-j} + \epsilon_t$$

avec :

$$\rho = \Phi_1 + \Phi_2 + \Phi_3 - 1, \quad \delta_1 = -(\Phi_2 + \Phi_3), \quad \delta_2 = -\Phi_3$$

Remarque 1.3.3 Pour un processus AR(p), la condition pour qu'une racine unitaire existe est que la somme des coefficients autorégressifs soit égale à 1, c'est-à-dire $\sum_{i=1}^p \Phi_i = 1$. Dans le cas d'un AR(3), cela se simplifie en $\Phi_1 + \Phi_2 + \Phi_3 = 1$. En posant $\rho = \Phi_1 + \Phi_2 + \Phi_3 - 1$, on peut tester si $\rho = 0$, ce qui équivaut à vérifier si $\Phi_1 + \Phi_2 + \Phi_3 = 1$, indiquant la présence d'une racine unitaire. Si $\rho < 0$, alors $\Phi_1 + \Phi_2 + \Phi_3 < 1$, ce qui suggère que la série est stationnaire.

Hypothèses du Test ADF

Les hypothèses du test ADF sont les suivantes :

$$\begin{cases} H_0 : \rho = 0 & \text{(La série temporelle contient une racine unitaire.)} \\ H_1 : \rho < 0 & \text{(La série temporelle ne contient pas de racine unitaire.)} \end{cases}$$

Pour tester H_0 , Dickey et Fuller ont proposé une statistique de test qui est :

$$t_n = \frac{\hat{\rho}}{\sigma_{\hat{\rho}}}$$

Où

- $\hat{\rho}$ C'est l'estimateur de ρ .
- $\hat{\sigma}_{\hat{\rho}}$ C'est l'estimateur de l'écart-type de $\hat{\rho}$.

Propriétés Asymptotiques

Nous énonçons d'abord les conditions que nous imposons sur la structure de dépendance du processus stationnaire d'ordre deux sous-jacent U_t qui conduit à la marche aléatoire sous l'hypothèse nulle. En supposant que U_t est purement non déterministe, c'est-à-dire qu'il possède une densité spectrale dont le logarithme est intégrable, alors la représentation de Word donne :

$$U_t = \sum_{j=1}^{\infty} \beta_j \varepsilon_{t-j} + \varepsilon_t$$

où $\sum_{j=1}^{\infty} \beta_j^2 < \infty$ et ε_t est un processus sans corrélation avec une moyenne nulle et une variance finie $0 < \text{Var}(\varepsilon_t) = \sigma^2 < \infty$.

Propriété 1

La propriété 1 impose deux conditions supplémentaires :

- (i) La fonction d'autocovariance $\gamma_U(h) = \text{Cov}(U_t, U_{t+h})$, $h \in \mathbb{Z}$, du processus $U = \{U_t, t \in \mathbb{Z}\}$ satisfait $\sum_{h \in \mathbb{Z}} |\gamma_U(h)| < \infty$, et la densité spectrale f_U de U est strictement positive, c'est-à-dire $f_U(\omega) > 0$ pour toutes les fréquences ω .
- (ii) $E(\varepsilon_t) < \infty$ et le processus $\{\varepsilon_t\}$ satisfait la condition de dépendance faible suivante : $\sum_{n=1}^{\infty} n \|P_1(\varepsilon_t)\| < \infty$ où $P_t(Y) = E(Y|\mathcal{F}_t) - E(Y|\mathcal{F}_{t-1})$ et $\mathcal{F}_s = (\dots, \varepsilon_{s-1}, \varepsilon_s)$.

Propriété 2

Le processus X satisfait l'une (et une seule) des deux conditions suivantes :

- (i) (Cas de racine unitaire :) $X_0 = 0$ et $X_t = X_{t-1} + U_t$ pour $t = 1, 2, \dots$
- (ii) (Cas stationnaire :) $X_t = U_t$ pour $t = 0, 1, 2, \dots$
Où $\{U_t\}$ est le processus stationnaire d'ordre deux spécifié dans l'Hypothèse 1.

L'hypothèse $X_0 = 0$ simplifie la notation et n'affecte pas les résultats asymptotiques de cet article pour le cas de racine unitaire. Elle peut être remplacée par d'autres hypothèses concernant la valeur initiale X_0 , à condition que cette variable aléatoire reste bornée en probabilité.

L'Hypothèse 2 stipule simplement que X_t est soit un processus stationnaire satisfaisant l'Hypothèse 1, soit qu'il est obtenu par intégration d'un tel processus stationnaire.

Remarque 1.3.4 *Le test suit une distribution t asymétrique où la plupart des valeurs sont négatives ([40]). Une fois la valeur t calculée, elle est comparée aux valeurs critiques simulées fournies par Fuller (1976) ([23]). L'hypothèse nulle est rejetée pour des petites valeurs, comme on peut le voir dans les hypothèses où l'alternative est que ρ est inférieure à zéro.*

Remarque 1.3.5 *Pour un processus où $p = 1$, le test de Dickey-Fuller est un cas particulier du test de Dickey-Fuller augmenté. Étant donné qu'il est d'ordre un, les termes différenciés retardés des périodes précédentes dans les équations (1.3.5) et (1.3.6), $\sum \Phi_i \Delta X_{t-j}$, disparaissent. Les équations deviennent alors :*

$$\Delta Y_t = \rho X_{t-1} + \varepsilon_t$$

et

$$\Delta X_t = \alpha + \rho X_{t-1} + \varepsilon_t.$$

Ce sont les modèles de régression proposés par Dickey et Fuller en 1979.

Exemple d'application du test ADF en R

Dans l'exemple qui suit, nous utilisons le package `urca` en R pour appliquer le test ADF sur la série temporelle "gnp.r" du jeu de données `nporg`. Le test est configuré pour inclure une tendance linéaire (`type = "trend"`), et le nombre de retards est sélectionné en fonction du critère BIC (Critère d'Information Bayésien) (`selectlags = "BIC"`).

Le package `urca` inclut un jeu de données appelé `nporg`, qui contient des séries macroéconomiques américaines (comme le PIB, prix, taux d'intérêt, etc.). Ici, nous analysons la série `gnp.r` (PIB réel) extraite de ce jeu de données.

Le BIC (Bayesian Information Criterion, ou Critère d'Information Bayésien) est un critère statistique utilisé pour sélectionner le meilleur modèle parmi plusieurs options, en équilibrant la qualité d'ajustement et la complexité.

La formule du BIC est :

$$\text{BIC} = \ln(n) \cdot k - 2 \ln(\hat{L})$$

où :

- n = nombre d'observations,
- k = nombre de paramètres (ex. : retards + tendance),
- \hat{L} = vraisemblance du modèle.

Dans la fonction `ur.df()` (test ADF), l'option `selectlags = "BIC"` permet de :

Chapitre 1 : Notions générales sur les séries chronologiques

- Choisir automatiquement le nombre optimal de retards (lags) à inclure dans le modèle.
- Éviter les problèmes d'autocorrélation dans les résidus, tout en limitant la surparamétrisation (trop de retards inutiles).
- Plus le BIC est bas, meilleur est le modèle.
- Le BIC pénalise plus fortement les modèles complexes que le critère AIC (Akaike Information Criterion), favorisant ainsi des modèles plus simples.
- Il est utilisé pour éviter le biais causé par un nombre inapproprié de retards

```
# Exemple de programme R
> require("urca")
> data(nporg)
> gnp <- na.omit(nporg[, "gnp.r"])
> summary(ur.df(y = gnp, lags = 4, type = "trend",
+ selectlags = "BIC"))

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-47.149  -9.212   0.819  11.031  23.924

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.89983     4.55369  -0.417  0.67821
z.lag.1      -0.05322     0.03592  -1.481  0.14441
tt           0.74962     0.36373   2.061  0.04423 *
z.diff.lag   0.39082     0.13449   2.906  0.00533 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.19 on 53 degrees of freedom
Multiple R-squared:  0.2727,    Adjusted R-squared:  0.2316
F-statistic: 6.625 on 3 and 53 DF,  p-value: 0.0006958

Value of test-statistic is: -1.4814 3.8049 2.7942

Critical values for test statistics:
```

	1pct	5pct	10pct
tau3	-4.04	-3.45	-3.15
phi2	6.50	4.88	4.16
phi3	8.73	6.49	5.47

Interprétation des résultats du test ADF

Analyse de la statistique de test

Les résultats du test de Dickey-Fuller augmenté révèlent une statistique de test de $-1,4814$. Comparée aux valeurs critiques de référence :

- Valeur critique à 1% : $-4,04$
- Valeur critique à 5% : $-3,45$
- Valeur critique à 10% : $-3,15$

Test d'hypothèse

La règle de décision conduit aux observations suivantes :

- La statistique observée ($-1,4814$) est supérieure à la valeur critique au seuil de 5% ($-3,45$)
- Par conséquent, on ne peut rejeter l'hypothèse nulle (H_0) de présence de racine unitaire

Conclusion sur la stationnarité

Cette analyse permet de conclure que :

- La série `gnp.r` présente une non-stationnarité au seuil de signification de 5%
- Cette conclusion est confirmée par l'examen des différents seuils critiques (1%, 5% et 10%)

Analyse complémentaire

L'examen des coefficients du modèle révèle par ailleurs que :

- La tendance linéaire (`tt`) est statistiquement significative ($p = 0,044$)
- Le terme retardé de la série différenciée (`z.diff.lag`) présente également une significativité ($p = 0,005$)

1.4 Autres tests

1.4.1 Test de Phillips-Perron (PP)

Phillips et Perron (1988) ([52]) ont proposé une méthode non paramétrique pour corriger l'autocorrélation des résidus, alternative au test de Dickey-Fuller. Leur approche modifie les statistiques de Dickey-Fuller via une correction de la variance des résidus, sans ajouter de termes retards ([53]).

Hypothèses du test

Les hypothèses du test de Phillips-Perron sont les suivantes :

$$\begin{cases} H_0 : \rho = 1 & \text{(la série contient une racine unitaire)} \\ H_1 : \rho < 1 & \text{(la série temporelle est stationnaire)} \end{cases}$$

Modèles du test

Trois modèles sont généralement utilisés pour tester l'hypothèse nulle d'une racine unitaire :

1) Modèle sans constante ni tendance : $Y_t = \rho Y_{t-1} + \varepsilon_t$

La statistique de test utilisée est :

$$Z(\hat{t}_\rho) = \left(\frac{\hat{\sigma}_c}{\hat{\sigma}_{Tl}} \right) \hat{t}_\rho - \frac{1}{2} \frac{\hat{\sigma}_{Tl}^2 - \hat{\sigma}_c^2}{\hat{\sigma}_{Tl} \left(T^{-2} \sum_{i=2}^T Y_{i-1}^2 \right)^{1/2}}$$

où $\hat{\sigma}_c^2$ est l'estimateur de la variance de court terme de ε_t

$$\hat{s}^2 = \frac{1}{T} \sum_{t=2}^T \varepsilon_t^2 \quad (1.4.1)$$

et $\hat{\sigma}_{Tl}^2$ est l'estimateur de la variance de long terme :

$$\hat{\sigma}_{Tl}^2 = \hat{\sigma}_c^2 + \frac{2}{T} \sum_{k=1}^l \left(1 - \frac{k}{l+1} \right) \sum_{i=k+1}^T \hat{\varepsilon}_i \hat{\varepsilon}_{i-k} \quad (1.4.2)$$

2) Modèle avec constante : $Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$

la statistique de test utilisée est :

$$Z(\tilde{t}_\rho) = \left(\frac{\tilde{\sigma}_c}{\tilde{\sigma}_{Tl}} \right) \tilde{t}_\rho - \frac{1}{2} \frac{\tilde{\sigma}_{Tl}^2 - \tilde{\sigma}_c^2}{\tilde{\sigma}_{Tl} \left(T^{-2} \sum_{i=2}^T (Y_{i-1} - \bar{Y}_{-1})^2 \right)^{1/2}}$$

$$Z(\tilde{t}_\alpha) = \left(\frac{\tilde{\sigma}_c}{\tilde{\sigma}_{Tl}} \right) \tilde{t}_\alpha + \frac{1}{2} \frac{\sqrt{T}(\tilde{\sigma}_{Tl}^2 - \tilde{\sigma}_c^2)}{\tilde{\sigma}_{Tl} \left(T^{-2} \sum_{i=2}^T (Y_{i-1} - \bar{Y}_{-1})^2 \right)^{1/2}} \cdot \frac{\sum_{i=2}^T Y_{i-1}}{\sqrt{\sum_{i=2}^T Y_{i-1}^2}}$$

avec :

- $\tilde{\sigma}_{Tl}$ et $\tilde{\sigma}_c$ donnés par des équations similaires à (1.4.1) et (1.4.2)
- $\bar{Y}_{-1} = (T-1)^{-1} \sum_{t=2}^T Y_{t-1}$
- $l = \text{int} \left[4 \left(\frac{T}{100} \right)^{1/4} \right]$

3) Modèle avec constante et tendance : $Y_t = \alpha + \beta t + \rho Y_{t-1} + \varepsilon_t$

La statistique de test utilisée est :

$$Z(t_\rho^*) = \left(\frac{\sigma_c^*}{\sigma_{Tl}^*} \right) t_\rho^* - \frac{T^3}{4\sqrt{3}} \frac{(\sigma_{Tl}^*)^2 - (\sigma_c^*)^2}{\sigma_{Tl}^* \sqrt{D_{xx}}}$$

$$Z(t_\alpha^*) = \left(\frac{\sigma_c^*}{\sigma_{Tl}^*} \right) t_\alpha^* - \frac{T^4 \sqrt{T} ((\sigma_{Tl}^*)^2 - (\sigma_c^*)^2) \sum_{i=2}^T Y_{i-1}}{4 \sqrt{3(\sigma_{Tl}^*)^2 D_{xx} \left(12D_{xx} + T^3 \left(\sum_{i=2}^T Y_{i-1} \right)^2 \right)}}$$

$$Z(t_\beta^*) = \left(\frac{s^*}{\sigma_T^*} \right) t_\beta^* = \frac{T \sqrt{T} (\sigma_{Tl}^{*2} - s^{*2}) \left(T \sum_{t=2}^T Y_{t-1}^2 - 2 \sum_{t=2}^T Y_{t-1} \sum_{t=2}^T Y_{t-1} \right)}{\left[4\sigma_{Tl}^{*2} D_{xx} \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{-1})^2 \right]^{1/2}}$$

où D_{xx} est le déterminant du produit scalaire de la matrice de données avec elle-même, soit :

$$D_{xx} = \frac{T^2(T^2 - 1)}{12} \sum_{t=2}^T Y_{t-1}^2 - T \left(\sum_{t=2}^T Y_{t-1} \right)^2 + \frac{T(T+1)}{2} \sum_{t=2}^T Y_{t-1} - \frac{T(T+1)(2T+1)}{6} \left(\sum_{t=2}^T Y_{t-1} \right)^2$$

et où σ_{Tl}^{*2} et s^{*2} sont donnés respectivement par (1.4.1) et (1.4.2) en remplaçant ϵ_t par ϵ_t^* .

En général, on prend $l = O(T^{1/4})$. Schwert (1989) suggère de calculer les statistiques de Phillips-Perron pour deux valeurs de l :

$$l = \left\lceil 4 \left(\frac{T}{100} \right)^{1/4} \right\rceil$$

où $[x]$ désigne la partie entière de x , c'est-à-dire le plus grand entier inférieur ou égal à x .

Pour conclure, notez que si $l = 0$, les tests de Phillips-Perron se ramènent aux procédures simples de Dickey-Fuller.

Remarque 1.4.1 Une caractéristique particulièrement intéressante des statistiques transformées que Phillips et Perron proposent (Perron 1988, p. 308-309) est que leurs distributions asymptotiques sont identiques à celles dérivées par Dickey et Fuller. Ceci implique que la procédure de test de Phillips et Perron peut être utilisée en se référant aux valeurs critiques asymptotiques tabulées par Dickey et Fuller.

Correction non paramétrique

La correction non paramétrique de Phillips-Perron utilise une estimation de la variance à long terme ($\hat{\sigma}_{Tl}^2$) pour ajuster les statistiques de test. Cela permet de capturer l'autocorrélation des résidus sans ajouter de termes retard supplémentaires.

Remarque 1.4.2 Les tests de Dickey-Fuller apparaissent donc comme des cas particuliers des tests de Phillips-Perron si $\hat{s}^2 = \hat{\sigma}_{Tl}^2$ ([16])

1.4.2 Test d'Elliott, Rothenberg et Stock (ERS)

Le test de Elliott, Rothenberg et Stock (ERS), souvent désigné sous le nom de DF-GLS (Dickey-Fuller Generalized Least Squares), est une variante améliorée du test de Dickey-Fuller pour détecter la présence d'une racine unitaire dans une série temporelle. Ce test a été proposé en 1996 par Elliott, Rothenberg et Stock.

Modèle de base

Considérons une série chronologique $\{Y_t, t \in \mathbb{Z}\}$ définie par :

$$\begin{cases} Y_t = d_t + \mu_t \\ \mu_t = \alpha\mu_{t-1} + \nu_t \end{cases}$$

où d_t est une composante déterministe (une constante ou une tendance polynomiale) et ν_t est un processus stationnaire non observé à moyenne nulle.

Hypothèses du test

Les hypothèses du test sont les suivantes :

$$\begin{cases} H_0 : \alpha = 1 \\ H_1 : |\alpha| < 1 \end{cases}$$

Elliott a supposé que ν_t est un processus stationnaire ergodique^[1] et d'autocovariance finie.

Réduction de tendance GLS

Le test ERS utilise une réduction de tendance par moindres carrés généralisés (GLS) ([17]) pour améliorer la puissance du test. Cette méthode transforme les données originales en soustrayant une tendance ajustée localement autour de l'unité.

1) Transformation des données : Les données sont transformées en utilisant la formule suivante :

$$Y_t^* = Y_t - \left(\frac{\bar{c}}{T}\right) Y_{t-1}$$

où Y_t^* sont les données transformées, et c/T ajuste localement autour de l'unité pour améliorer la puissance du test ($c = T(\alpha - 1)$).

2) Application du test ADF modifié : Ensuite, on applique un test ADF modifié sur ces données transformées pour tester si le coefficient autorégressif reste proche de 1 ou non.

Statistique de test ERS

La statistique du test ERS se base sur celle du Dickey-Fuller mais appliquée aux données transformées par GLS. Le test DF-GLS a une puissance supérieure au test ADF traditionnel, surtout lorsque la racine autoregressive est proche de 1 mais inférieure à 1.

Avantages

- Puissance accrue : Le test ERS offre une puissance supérieure au test ADF traditionnel dans certaines conditions.
- Flexibilité : Il peut être utilisé pour détecter des racines unitaires même lorsque les données présentent des tendances ou des constantes.

[1]. Nous utilisons ici le terme "ergodique" dans un sens général. En termes simples, un processus est dit ergodique si la majorité de ses points initiaux visitent chaque région d'un espace des phases avec une probabilité à peu près égale.

1.4.3 Test de Kwiatkowski, Phillips, Schmidt et Shin(KPSS)

Contrairement aux tests précédents, l'hypothèse nulle du test KPSS ([38]), est la stationnarité de la série temporelle étudiée contre l'alternative d'une racine unitaire. Les auteurs expriment la série temporelle sous étude comme la somme d'une tendance déterministe, d'une marche aléatoire et d'une erreur stationnaire. L'hypothèse nulle de stationnarité autour d'une tendance est équivalente à l'hypothèse selon laquelle la variance de la marche aléatoire est nulle. Plus formellement la série temporelle Y_t , $t=1,2,\dots,T$ sous étude est générée comme suit :

$$\begin{cases} Y_t = w_t + \lambda t + u_t, & u_t \sim i.i.d.(0; \sigma_u^2) \\ w_t = w_{t-1} + v_t, & v_t \sim i.i.d.(0; \sigma_v^2) \end{cases}$$

Les hypothèses de ce test :

- a) La série temporelle Y_t étant stationnaire avec tendance. Le test KPSS vérifie alors s'il existe une racine unitaire dans w_t lorsque λ est non nul. Les hypothèses pour le test sont alors :

$$\begin{cases} H_0 : \text{La série temporelle est stationnaire avec tendance} \\ H_1 : \text{La série temporelle contient une racine unitaire} \end{cases}$$

- b) Le test KPSS peut également être utilisé pour tester la stationnarité au niveau. Dans ce cas, $\lambda = 0$ les hypothèses deviennent :

$$\begin{cases} H_0 : \text{La série temporelle est stationnaire au niveau} \\ H_1 : \text{La série temporelle contient une racine unitaire} \end{cases}$$

La statistique de test :

La statistique de test utilisée par ce test est un multiplicateur de Lagrange unilatéral et est donnée par :

$$LM_{KPSS} = \frac{\sum_{t=1}^T S_t^2}{\hat{\sigma}_\varepsilon^2} \quad (1.4.3)$$

Où S_t est la somme partielle des résidus $S_t = \sum_{i=1}^t e_i$, $t=1,\dots,T$
Et $\hat{\sigma}_\varepsilon^2$ est l'estimateur de la variance des résidus .

Chapitre 2

cointégration linéaire et non linéaire

2.1 Introduction

La cointégration concerne les séries non stationnaires, et au sein de celles-ci, les séries intégrées. Le concept a été présenté par Granger en 1981 ([25]) et re-précisé en 1987 par Engle et Granger ([18]). Il a intéressé les économistes en raison de la connexion existant entre les équations d'équilibre de leurs modèles et cette propriété de cointégration. Par ailleurs, le théorème de représentation de Granger, qui s'applique aux séries cointégrées, a rationalisé la structure des modèles à correction d'erreur dans lesquels la dynamique d'une variable intégrée va dépendre à la fois de l'état de la conjoncture et des ajustements vers un équilibre de long terme.

Cependant, il faut noter que les relations entre variables économiques ne sont pas toujours linéaires. Dans ces cas, les tests classiques de cointégration et de stationnarité, qui reposent sur l'hypothèse de linéarité, peuvent donner des résultats erronés ou biaisés. En effet, ces tests conventionnels ignorent les dynamiques non linéaires présentes dans les données, ce qui peut conduire à ne pas détecter une cointégration réelle ou à conclure à tort à son existence. Pour pallier ces limites, des approches plus récentes ont étendu la notion de cointégration au cadre non linéaire, en développant des modèles à correction d'erreur non linéaires et des tests adaptés à ces situations ([16]).

Dans la première partie de ce chapitre, nous allons nous intéresser à la cointégration linéaire, qui est la forme la plus couramment utilisée pour identifier les relations de long terme entre des variables non stationnaires. Ensuite, dans la deuxième partie, nous aborderons la cointégration non linéaire, qui permet de capturer des relations plus complexes entre ces variables.

Définition 2.1.1 *On considère deux séries X_t et Y_t dont les processus sont intégrés d'ordre un ($I(1)$). Elles sont dites cointégrées, s'il existe une combinaison linéaire unique des deux variables qui se révèle intégrée d'ordre 0 ($I(0)$). Une telle combinaison linéaire peut s'écrire sous la forme suivante :*

$$\varepsilon_t = Y_t - aX_t - b \quad (2.1.1)$$

Autrement dit, deux séries non stationnaires ($X_t \sim I(1)$ et $Y_t \sim I(1)$) sont dites cointégrées si on a :

$$Y_t - aX_t - b = \varepsilon_t \sim I(0)$$

Les séries X_t et Y_t sont alors notées :

$$X_t, Y_t \sim Cl(1, 1)$$

De manière générale, si X_t et Y_t sont deux séries intégrées d'ordre d ($I(d)$), alors il est possible que la combinaison linéaire (1) ne soit pas $I(d)$ mais $I(d-b)$ où b est un entier positif non nul ($0 < b \leq d$). Le vecteur $[1, -a, -b]$ est appelé vecteur de cointégration, les séries sont alors cointégrées.

Exemple

$$\begin{aligned} X_t &= X_{t-1} + u_t, \\ \text{où } u_t &\sim \text{i.i.d. } N(0, \sigma_u^2), \text{ avec } \sigma_u^2 \text{ une constante positive,} \\ Y_t &= X_t + v_t, \\ \text{où } v_t &\sim \text{i.i.d. } N(0, \sigma_v^2), \text{ avec } \sigma_v^2 \text{ une constante positive,} \\ Z_t &= Y_t - X_t \end{aligned}$$

Ici, (X_t) qui est une marche aléatoire est intégrée d'ordre 1, Y_t qui est une marche aléatoire augmentée d'une composante stationnaire est également intégrée d'ordre 1, mais $Z_t = Y_t - X_t = v_t$ est intégrée d'ordre 0. Ainsi, (X_t, Y_t) est constitué de variables cointégrées de vecteur cointégrant $(-1, 1)$. La Figure 2.1 reproduit une simulation sur 200 observations du système d'équations précédent.

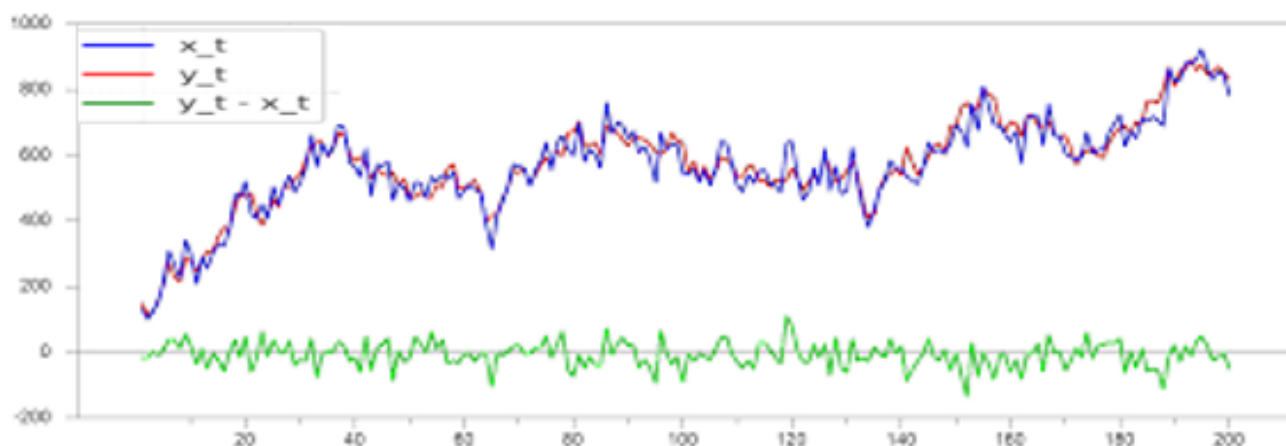


FIGURE 2.1 – Exemple de deux séries $I(1)$ cointégrées

Remarque 2.1.1 .

- Toute transformation linéaire d'un vecteur cointégrant est également cointégrante.

En effet, si (α, β) avec $\alpha \in \mathbb{R}$ et $(\alpha, \beta) \neq (0, 0)$ est un vecteur cointégrant, alors $\alpha X_t + \beta Y_t \sim I(0)$, mais alors pour toute constante c , la combinaison linéaire réalisée avec la transformée $(c\alpha, c\beta)$ vérifie $c\alpha X_t + c\beta Y_t = c(\alpha X_t + \beta Y_t)$ et est donc bien $I(0)$: $(c\alpha, c\beta)$ est cointégrant.

- Dans un cadre bivarié, et à ses transformées linéaires près, le vecteur cointégrant est unique.

Preuve

Montrons Cette propriété par l'absurde

Soient deux variables X et Y intégrées d'ordre 1 telles que :

$$Z_1 = \alpha_1 X_t + \beta_1 Y_t \sim I(0)$$

Et un autre vecteur cointégrant, non proportionnel au premier tel que :

$$Z_2 = \alpha_2 X_t + \beta_2 Y_t \sim I(0)$$

avec :

$$\alpha_1, \alpha_2 \in \mathbb{R} \setminus \{0\}, \quad (\alpha_1, \beta_1) \neq (0, 0), \quad (\alpha_2, \beta_2) \neq (0, 0)$$

Cette dernière propriété, si elle est vraie, assure que :

$$-(\alpha_1/\alpha_2)[\alpha_2 X_t + \beta_2 Y_t] \sim I(0)$$

L'addition de Z_1 et de $-(\alpha_1/\alpha_2)Z_2$, qui sont toutes deux $I(0)$, donnerait une variable également $I(0)$ définie par :

$$(\beta_1 - (\alpha_1/\alpha_2)\beta_2)Y_t$$

Ce qui impliquerait que Y_t est $I(0)$ or elle est $I(1)$.

- Une condition nécessaire de la cointégration est que les séries doivent être intégrées de même ordre, sinon elles ne peuvent pas être cointégrées.
- La combinaison linéaire de ces deux séries permet de ramener à une série d'ordre d'intégration inférieur.

2.2 Les tests de cointégration

Il existe plusieurs grandes catégories de tests de cointégration ; nous n'en présenterons que trois.

Le premier repose sur des tests de stationnarité des résidus d'une régression entre les composantes de X_t (dite régression cointégrante). Le second, plus efficace mais aussi plus coûteux, utilise l'estimation par le maximum de vraisemblance d'un autorégressif vectoriel sur X_t . Le dernier repose sur des statistiques de type Phillips-Ouliaris, qui généralisent le premier test avec des propriétés asymptotiques améliorées.

2.2.1 Test d'Engle et Granger (1987)

Selon Engle et Granger (1987) ([18]), tester la cointégration revient à opérer un test de racine unitaire sur les résidus de l'équation de cointégration.

Concrètement, il s'agit dans un premier temps d'estimer par la méthode des MCO la relation de long terme entre X_t et Y_t . On note par $\hat{\epsilon}_t$ le résidu d'estimation.

Dans un deuxième temps, d'appliquer la méthodologie des tests de Dickey-Fuller et de Phillips et Perron à ce résidu (Si le résidu est stationnaire, c'est-à-dire qu'il n'y a pas de racine unitaire, cela indique que les variables sont cointégrées.) :

$$\hat{\epsilon}_t = \alpha \hat{\epsilon}_{t-1} + \epsilon_t \tag{2.2.1}$$

Où ϵ_t représente le terme d'erreur.

Étape 1 : Estimer la relation de long terme par MCO pour obtenir les résidus

Si on a

$$X_t \sim I(1) \quad \text{et} \quad Y_t \sim I(1)$$

On estime par les MCO la relation de long terme :

$$Y_t = aX_t + b + \varepsilon_t$$

Pour qu'il y ait cointégration, il faut que le résidu e_t issu de la régression soit stationnaire :

$$e_t = Y_t - \hat{a}X_t - \hat{b} \sim I(0)$$

Étape 2 : Test de racine unitaire sur les résidus.

Après avoir estimé la relation de long terme entre x_t et y_t par MCO, on obtient les résidus \hat{e}_t . Ensuite, on applique un test de racine unitaire sur ces résidus pour vérifier s'ils sont stationnaires :

$$\hat{e}_t = \alpha \hat{e}_{t-1} + \epsilon_t$$

Si les résidus sont stationnaires ($\hat{e}_t \sim I(0)$), cela indique que les variables sont cointégrées

Remarques

1. La stationnarité du résidu est testée à l'aide du test de Dickey-Fuller (simple ou augmenté).
2. Ici la relation porte sur les résidus estimés et non pas sur les **vrais** résidus de l'équation de cointégration. Par conséquent, nous ne pouvons pas nous référer aux tables de Dickey-Fuller pour mener le test de stationnarité.
3. Si le résidu est stationnaire nous pouvons alors estimer un modèle appelé modèle à correction d'erreur (MCE) qui intègre les variables en variation et en niveau (le théorème de la représentation de Granger met en évidence le lien entre cointégration et modèle à correction d'erreur).
4. L'emploi d'un modèle à correction d'erreur dans le cas de la cointégration permet d'obtenir des prévisions plus fiables que si on avait utilisé la relation de long terme car les résultats de l'estimation de cette relation sont faussés par la non stationnarité des séries.

2.2.2 Exemple R pour la méthode d'Engle et Granger (1987)

```
# 1. Charger le package
library(urca)

# 2. Créer des données simulées
set.seed(123)
Xt <- cumsum(rnorm(100)) # Série I(1)
Yt <- 0.7 * Xt + rnorm(100) # Relation de cointégration + bruit

# 3. Méthode Engle & Granger
```

```
# Étape 1 : Régression MCO
model <- lm(Yt ~ Xt)
residus <- residuals(model)

# Étape 2 : Test ADF sur les résidus
adf_test <- ur.df(residus, type = "none", lags = 1)
summary(adf_test)

#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.0157 -0.6784 -0.1029  0.6504  3.1015

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.23719    0.15359  -8.055 2.17e-12 ***
z.diff.lag   0.09603     0.10191   0.942  0.348
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.965 on 96 degrees of freedom
Multiple R-squared:  0.5661,    Adjusted R-squared:  0.557
F-statistic: 62.61 on 2 and 96 DF,  p-value: < 2.2e-16

Value of test-statistic is: -8.0553

Critical values for test statistics:
      1pct  5pct 10pct
taul -2.6 -1.95 -1.61
```

Interprétation

La p-valeur du test ADF sur les résidus étant très faible ($p = 2,17 \times 10^{-12}$), donc inférieure à 5%, on rejette l'hypothèse nulle de non-stationnarité. Les résidus sont donc stationnaires, ce qui confirme que les séries X_t et Y_t sont cointégrées.

Conclusion de l'analyse :

- Les résidus du modèle de cointégration apparaissent stationnaires ($I(0)$)
- On en déduit que les séries X_t et Y_t sont cointégrées
- Cela valide l'existence d'une relation d'équilibre de long terme entre les deux variables

2.2.3 Test de Johansen (1988)

L'approche en deux étapes d'Engle et Granger présente des limitations importantes. Non seulement elle ne permet de détecter qu'une seule relation de cointégration, mais elle souffre également du problème de normalisation arbitraire de la variable dépendante dans la régression statique.

Søren Johansen a développé sa méthode révolutionnaire à la fin des années 1980, offrant une approche multivariée complète basée sur le maximum de vraisemblance. Cette méthode permet :

- De déterminer le nombre de relations de cointégration (rang cointégrant).
- D'estimer simultanément l'ensemble des vecteurs cointégrants.
- De tester des restrictions économiques sur les paramètres de cointégration.

Le cadre méthodologique s'appuie sur un modèle vectoriel à correction d'erreur (VECM) :

$$\Delta \mathbf{X}t = \mathbf{\Pi} \mathbf{X}t - 1 + \sum_{i=1}^{p-1} \mathbf{\Gamma}i \Delta \mathbf{X}t - i + \boldsymbol{\varepsilon}_t \quad (2.2.2)$$

Où :

- $\mathbf{\Pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$ est la matrice de long terme.
- $\boldsymbol{\alpha}$ représente la vitesse d'ajustement vers l'équilibre.
- $\boldsymbol{\beta}$ contient les vecteurs de cointégration
- Le rang de $\mathbf{\Pi}$ (noté r) détermine le nombre de relations de cointégration.

Johansen teste le rang de cointégration (nombre de relations de long terme) parmi plusieurs séries non stationnaires. Il utilise le test de la trace et du maximum de valeur propre.

Test de la trace

$$\lambda_{\text{trace}}(r) = -T \sum_{i=r+1}^k \ln(1 - \hat{\lambda}_i) \quad (2.2.3)$$

$$\left\{ \begin{array}{l} H_0 : \text{rang}(\mathbf{\Pi}) \leq r \\ H_1 : \text{rang}(\mathbf{\Pi}) > r \end{array} \right.$$

Test du maximum de valeur propre

$$\lambda_{\text{max}}(r) = -T \ln(1 - \hat{\lambda}_{r+1}) \quad (2.2.4)$$

$$\left\{ \begin{array}{l} H_0 : \text{rang}(\mathbf{\Pi}) = r \\ H_1 : \text{rang}(\mathbf{\Pi}) = r + 1 \end{array} \right.$$

Avantages clés de l'approche Johansen :

- Non-sensibilité au choix de la variable dépendante
- Possibilité de tester des hypothèses économiques sur $\boldsymbol{\beta}$
- Meilleure puissance en échantillon fini que les approches alternatives

Exemple : Pour un système à $k = 3$ variables où on trouve $r = 2$ relations de cointégration, cela implique l'existence de deux combinaisons linéaires stationnaires des variables, révélant ainsi deux équilibres de long terme stables([30]).

Remarque importante : La mise en œuvre pratique nécessite des choix appropriés de :

- La longueur des retards (p) dans le VECM
- La forme de la tendance déterministe
- La normalisation des vecteurs de cointégration

2.2.4 Test de Phillips-Ouliaris

Le test de Phillips-Ouliaris, développé par Peter C.B. Phillips et Sam Ouliaris en 1990 ([50]), représente une avancée méthodologique majeure en offrant une procédure robuste pour détecter la présence de relations cointégrées.

Contrairement à l'approche initiale de Engle et Granger (1987) ([18]) qui repose sur une application standard des tests de racine unitaire aux résidus d'une régression statique, le test de Phillips-Ouliaris incorpore des corrections non paramétriques pour tenir compte de l'autocorrélation et de l'hétéroscédasticité présentes dans les résidus. Cette innovation permet d'obtenir des statistiques de test dont les distributions asymptotiques sont libres de paramètres de nuisance, garantissant ainsi une plus grande fiabilité des inférences.

Cadre théorique et fondements mathématiques

Considérons un vecteur $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{nt})'$ de n séries temporelles intégrées d'ordre 1, notées $I(1)$.

Le test de Phillips-Ouliaris se concentre sur l'analyse des résidus $\hat{\epsilon}_t$ obtenus par l'estimation de la relation de long terme via les moindres carrés ordinaires (MCO) :

$$y_{1t} = \alpha + \beta_2 y_{2t} + \dots + \beta_n y_{nt} + \epsilon_t. \quad (2.2.5)$$

Sous l'hypothèse nulle d'absence de cointégration, les résidus $\hat{\epsilon}_t$ sont non stationnaires ($I(1)$), tandis que sous l'hypothèse alternative de présence de cointégration, ils sont stationnaires ($I(0)$). La contribution essentielle de Phillips et Ouliaris consiste à dériver les distributions asymptotiques des tests de racine unitaire appliqués aux résidus $\hat{\epsilon}_t$ dans ce contexte particulier, en corrigeant les biais induits par la dépendance sérielle et l'hétéroscédasticité.

Statistiques du test

Le test de Phillips-Ouliaris propose deux statistiques principales, notées Z_α et Z_t , qui sont des versions modifiées des statistiques de Phillips-Perron adaptées au contexte de la cointégration.

1. La statistique Z_α est définie par :

$$Z_\alpha = T(\hat{\rho} - 1) - \frac{1}{2} \frac{\hat{\sigma}_T^2 - \hat{s}_T^2}{T^{-2} \sum_{t=1}^T \hat{\epsilon}_{t-1}^2}, \quad (2.2.6)$$

où :

- T est le nombre d'observations,
- $\hat{\rho}$ est l'estimateur MCO de l'autorégression d'ordre 1 des résidus : $\hat{\epsilon}_t = \rho \hat{\epsilon}_{t-1} + u_t$,
- $\hat{\sigma}_T^2$ est un estimateur de la variance de long terme de u_t ,
- \hat{s}_T^2 est un estimateur de la variance instantanée de u_t .

2. La statistique Z_t est définie par :

$$Z_t = \sqrt{\frac{\hat{s}_T^2}{\hat{\sigma}_T^2}} \cdot t_{\hat{\rho}} - \frac{1}{2} \frac{\hat{\sigma}_T^2 - \hat{s}_T^2}{\hat{\sigma}_T \sqrt{T^{-2} \sum_{t=1}^T \hat{\epsilon}_{t-1}^2}}, \quad (2.2.7)$$

où $t_{\hat{\rho}}$ est la statistique t usuelle associée à $\hat{\rho}$.

Procédure d'application

L'application du test de Phillips-Ouliaris suit une séquence logique en quatre étapes :

Étape 1 : Vérification de l'intégration : Tester que toutes les variables candidates à la cointégration sont bien intégrées d'ordre 1 à l'aide de tests de racine unitaire appropriés (ADF, Phillips-Perron, etc.). Si une ou plusieurs variables sont stationnaires ($I(0)$), le concept de cointégration ne s'applique pas directement.

Étape 2 : Estimation de la relation de long terme : Estimer par MCO la relation de cointégration présumée comme dans l'équation (2.2.5). Notons que l'estimateur MCO est convergent sous l'hypothèse alternative de cointégration, mais non normalement distribué (distribution de Dickey-Fuller).

Étape 3 : Extraction et analyse des résidus : Extraire les résidus $\hat{\epsilon}_t$ de la régression précédente. Ces résidus doivent être soumis à un test de racine unitaire, mais avec les corrections spécifiques de Phillips-Ouliaris.

Étape 4 : Calcul des statistiques et décision : Calculer les statistiques Z_α et Z_t selon les équations (2.2.6) et (2.2.7), en estimant les variances nécessaires avec une fenêtre de troncature appropriée. Comparer les valeurs obtenues aux valeurs critiques tabulées.

Le rejet de l'hypothèse nulle (H_0 : absence de cointégration) se fait si les statistiques sont inférieures aux valeurs critiques au seuil choisi.

2.3 Modèle à correction d'erreur

Pour tester la relation de cointégration selon Engle et Granger, on doit d'abord tester la relation de long terme et en récupérer le résidu de cette relation. Ensuite, on effectue le test ADF sur ce résidu pour tester sa stationnarité. L'étape qui suit est l'estimation d'un modèle à correction d'erreur.

Qu'est-ce qu'un modèle à correction d'erreur ?

Le modèle à correction d'erreur décrit l'ajustement de court terme des variables vers la relation d'équilibre de long terme.

Autrement dit, à $t = 0$, les variables forment une relation d'équilibre qu'on appelle la cointégration, mais elles ne sont pas à l'abri des chocs. Ainsi, lorsqu'un choc survient et affecte l'une ou plusieurs des variables, il peut casser la relation d'équilibre qui les relie. En conséquence, les variables vont s'éloigner les unes des autres, c'est-à-dire qu'elles vont s'écarter de la relation initiale. Cela produit une déviation temporaire par rapport à cette relation d'équilibre.

Et donc, cette erreur, cette déviation qu'on appelle l'erreur d'équilibre, c'est cette dérive à court terme. Car à long terme, il y a toujours une relation d'équilibre, il y a une force de rappel qui ramène les variables les unes vers les autres.

Donc, même s'il y a toujours un écart à court terme, il y a toujours un retour vers le long terme, exactement : il y a toujours une correction d'erreur.

Définition 2.3.1 *Si deux séries sont cointégrées, c'est-à-dire que leur combinaison linéaire est stationnaire ($Y_t - \hat{a}X_t - \hat{b} \sim I(0)$) On peut estimer un modèle à correction d'erreur (MCE) pour modéliser conjointement les dynamiques de court et de long terme.*

Équation du Modèle à Correction d'Erreur

On peut estimer le modèle à correction d'erreur (MCE) suivant :

$$\Delta Y_t = \gamma \Delta X_t + \underbrace{\delta (Y_{t-1} - aX_{t-1} - b)}_{\text{Terme de correction}} + \varepsilon_t \quad \text{avec} \quad \delta < 0 \quad (2.3.1)$$

On peut remarquer que le paramètre δ doit être négatif pour qu'il y ait un retour de Y_t à sa valeur d'équilibre de long terme qui est $(aX_{t-1} + b)$. En effet, lorsque Y_{t-1} est supérieur à $(aX_{t-1} + b)$, il n'y a une force de rappel vers l'équilibre de long terme que si $\delta < 0$ ([10]).

Le MCE permet de modéliser conjointement les dynamiques de court terme (représentées par les variables en différence première) et de long terme (représentées par les variables en niveau).

— La dynamique de court terme s'écrit :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + \nu_t \quad (2.3.2)$$

— La dynamique de long terme s'exprime de la manière suivante :

$$Y_t = aX_t + b + \varepsilon_t \quad (2.3.3)$$

2.4 Autres approches multivariées

2.4.1 Cointégration linéaire basée sur l'approche ARDL

Dans ce paragraphe, nous présentons un test de cointégration linéaire basé sur les modèles autorégressifs à retards échelonnés appelés ARDL (AutoRegressive Distributed Lag).

Pour tester l'existence ou non de la cointégration entre séries, la littérature économétrique fournit plusieurs tests ou approches dont le test de Engel et Granger (1987) ([18]), ceux de Johansen (1988, 1991) ([30],[31]) et Johansen et Juselius (1990) ([33]), et celui de Pesaran et al. (1997) ([47]), Pesaran et Shin (1998) ([48]) et Pesaran et al. (2001) ([49]). Le test de cointégration de Engle et Granger n'aide à vérifier la cointégration qu'entre deux séries intégrées de même ordre (soit ordre d'intégration = 1), il est donc adapté au cas bivarié et s'avère ainsi moins efficace pour des cas multivariés. Les tests de cointégration de Johansen permettent de vérifier plutôt la cointégration sur plus de deux séries, il ont été conçus pour des cas multivariés. Toutefois, bien que le test de Johansen fondé sur une modélisation vectorielle autorégressive à correction d'erreur (VECM) constitue un remède aux limites du test de Engle et Granger pour le cas multivarié, il exige aussi que toutes les séries ou variables soient intégrées de même ordre, ce qui n'est pas toujours le cas en

pratique. Alors, lorsqu'on dispose de plusieurs variables intégrées d'ordres différents ($I(0)$, $I(1)$), l'on peut recourir au test de cointégration de Pesaran .

Au début des années 2000, Pesaran et Shin (1995)([47]) et Pesaran et al. (2001)([49]) ont introduit une approche de test de cointégration appelée le test de limites de l'autorégression à retards échelonnés (ARDL bounds test). Cette approche est devenue populaire car elle brise la restriction traditionnelle des tests de cointégration selon laquelle les variables testées doivent être non stationnaires et que toutes les variables soient intégrées du même ordre. Certains chercheurs favorisent cette approche, car de nombreuses applications impliquent des variables économiques d'ordre inconnu ou d'interprétation difficile.

De nombreux chercheurs utilisent le test ARDL bounds comme test de cointégration. Par exemple, Muscatelli et Spinelli (2000) ([42]) ont adopté ce test pour examiner la stabilité de la demande de monnaie en Italie; Rushdi et al. (2012) ([56]) ont étudié la relation à long terme entre les recettes fiscales et la production en Australie; (Jalil et al. 2013)([29]) ont étudié la relation entre la consommation de pétrole et la croissance économique au Pakistan; Kim et Baek (2013) ([36]) ont réexaminé la relation entre les volumes d'échanges et les prix du pétrole en Corée; et Asteriou et al. (2016)([3]) ont examiné l'influence de la croissance économique, du taux de change et de la volatilité des marchés internationaux sur les prix du pétrole au Mexique, au Nigeria et en Turquie. Beaucoup de ces études comportent des implications importantes pour la politique, il est donc essentiel que les chercheurs utilisent la version correcte du test ARDL amélioré afin d'éviter des résultats erronés ou des conclusions incorrectes. L'exactitude et la robustesse de ces résultats sont cruciales pour fournir de bonnes bases aux décideurs politiques et pour des recommandations efficaces sur les politiques économiques ou fiscales.

À côté de la stationnarité, le problème de *causalité*^[1] a aussi attiré l'attention de plusieurs chercheurs, entre autres Granger (1980)([26]) qui va en proposer une définition simple : selon Granger, l'on dira d'une variable X qu'elle cause Y si la seule connaissance du passé de X améliore les prévisions sur Y que ne le ferait le passé de Y . Ainsi, on sait que la corrélation (existence d'un lien positif ou négatif entre deux séries) de la causalité, et avons compris que causalité implique corrélation, mais non l'inverse.

Granger a proposé un test de causalité qui va porter sur des séries stationnaires intégrées de même ordre, et mais pas aux séries intégrées à des ordres différents. Pour y remédier, Toda et Yamamoto (1995)([62]) proposent une procédure non séquentielle de test de causalité adaptée aux séries intégrées à des ordres différents. Pour ces deux auteurs, les tests préliminaires de stationnarité et cointégration (procédures séquentielles de Granger) importent peu pour l'économiste qui doit se soucier à tester plutôt les restrictions théoriques (elles sécurisent l'information en niveau). Toda et Yamamoto ont proposé d'estimer un VAR en niveau corrigé (sur-paramétré), devant servir de base au test de causalité, sous l'hypothèse d'une potentielle cointégration probable entre séries qu'ils intègrent dans le modèle sans l'étudier explicitement.

2.4.2 Modèles ARDL

Les modèles *AutoRegressive Distributed Lag* (ARDL), ou « modèles autorégressifs à retards échelonnés ou distribués » (ARRE), sont des modèles dynamiques ([24]). Ces

[1]. Dans ce contexte, le terme « causalité » doit être compris dans le sens spécifique de Granger : une relation prédictive entre séries temporelles où l'information passée d'une variable améliore la prévision d'une autre. Il ne s'agit pas d'une relation causale au sens philosophique ou physique, mais d'une causalité purement statistique et temporelle.

derniers ont la particularité de prendre en compte la dynamique temporelle (délai d'ajustement, anticipations, etc.) dans l'explication d'une variable (série chronologique), améliorant ainsi les prévisions et l'efficacité des politiques (décisions, actions, etc.), contrairement au modèle simple (non dynamique) dont l'explication instantanée (effet immédiat ou non étalé dans le temps) ne restitue qu'une partie de la variation de la variable à expliquer. Dans la famille des modèles dynamiques, l'on distingue trois types de modèles.

Si on considère une variable endogène (Y_t) et une variable exogène (X_t), qu'on définit comme suit :

Définition 2.4.1 Variables endogènes : ce sont des variables dont la valeur est déterminée par le modèle.

Définition 2.4.2 Variables exogènes : ce sont les variables du modèle qu'on considère déterminées à l'extérieur du modèle.

- Les modèles autorégressifs (AR) : c'est des modèles dynamiques où l'on trouve, parmi les variables explicatives (Y_t), la variable dépendante décalée (expliquée par ses propres valeurs passées (Y_{t-i})), il s'exprime de la manière suivante :

$$Y_t = a_0 + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + \varepsilon_t$$

ou bien sous la forme :

$$Y_t = a_0 + \sum_{i=1}^p a_i Y_{t-i} + \varepsilon_t \quad (2.4.1)$$

Avec

$\varepsilon_t \sim \text{iid}(0, \sigma)$: terme d'erreur.

- Les modèles à retards échelonnés ou distributed lag (DL) : ce sont des modèles dynamiques qui ont pour variables explicatives (X_t), et ses valeurs passées ou décalées. Ils prennent la forme suivante :

$$Y_t = \beta + b_0 X_t + \dots + b_q X_{t-q} + z_t$$

ou bien sous forme sommée :

$$Y_t = \beta + \sum_{i=0}^q b_i X_{t-i} + z_t \quad (2.4.2)$$

- Les modèles autorégressifs à retards échelonnés (ARDL) : ces modèles combinent les caractéristiques de deux précédents ; on y trouve, parmi les variables explicatives (X_t), la variable dépendante décalée (Y_{t-i}) et les valeurs passées de la variable indépendante (X_{t-i}), ainsi sa forme s'écrit :

$$Y_t = \varphi + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + b_0 X_t + \dots + b_q X_{t-q} + e_t$$

ou bien sous la forme :

$$Y_t = \varphi + \sum_{i=1}^p a_i Y_{t-i} + \sum_{i=0}^q b_i X_{t-i} + e_t \quad (2.4.3)$$

Le terme retards échelonnés montre que les effets à court terme de X_t sur Y_t sont différents de ceux à long terme. D'un point de temps à un autre, les échelles de réaction de Y_t au changement de X_t différent.

On note que (b_0) traduit l'effet à court terme de (X_t) sur (Y_t) , et pour expliquer l'effet à long terme de (X_t) sur (Y_t) , on peut calculer λ à partir de la relation de long terme notée (2.3.3)

Avec :

$$\lambda = \frac{\sum_{i=0}^q b_i}{1 - \sum_{i=1}^p a_i}$$

Ces modèles dynamiques souffrent généralement de problèmes d'autocorrélation d'erreurs, avec la présence de la variable endogène décalée comme explicative (modèles AR et ARDL), et de multi-colinéarité (modèles DL et ARDL), ce qui complique l'estimation des paramètres par les Moindres Carrés Ordinaires/MCO. Ici, il tient de recourir aux techniques d'estimation robuste (méthode SUR, etc.) pour pallier à ces problèmes. Aussi, l'on retiendra que les variables considérées dans ces modèles se doivent d'être stationnaires pour éviter des régressions fallacieuses.

L'on retiendra d'un modèle ARDL que, faisant partie de la famille des modèles dynamiques, il permet d'estimer les dynamiques de court terme et les effets de long terme pour des séries cointégrées ou même intégrées à des ordres différents, mais on ne sait pas appliquer un modèle ARDL pour des séries intégrées à des ordres qui dépassent 1 ([49]).

2.4.3 Test de Cointégration aux Bornes (Bounds-test)

Le test de cointégration aux bornes est basé sur l'estimation des modèles vectoriels autorégressifs à retards échelonnés. Il est utilisé lorsque l'on suspecte la présence d'au moins deux séries non stationnaires, ce qui conduit à rechercher l'existence d'une relation de cointégration (relation d'équilibre à long terme) entre les variables du modèle.

Ce test est applicable lorsque les variables sont intégrées d'ordre 0 ($I(0)$) ou d'ordre 1 ($I(1)$). Cette caractéristique réduit le problème lié à l'incertitude des résultats du test de racine unitaire.

Équation de Base

L'équation de base pour le test de cointégration aux bornes est une forme modifiée du modèle ARDL, incluant les différences des variables pour capturer les effets à court terme et les termes en niveau pour les effets à long terme :

$$\Delta Y_t = a_0 + \sum_{i=1}^p a_i \Delta Y_{t-i} + \sum_{i=0}^q b_i \Delta X_{t-i} + \delta_1 Y_{t-1} + \delta_2 X_{t-1} + \mu_t \quad (2.4.4)$$

Avec :

- δ_1 : Coefficient qui représente l'effet à long terme de Y_{t-1} sur ΔY_t . Un coefficient négatif indique une correction d'erreur vers l'équilibre.
- δ_2 : Coefficient qui représente l'effet à long terme de X_{t-1} sur ΔY_t . Ce coefficient peut être utilisé pour évaluer l'influence de X sur la dynamique de Y .
- μ_t : Terme d'erreur .

La procédure du test est telle que l'on devra comparer les valeurs de Fisher obtenues aux valeurs critiques (bornes) simulées pour plusieurs cas et différents seuils par Pesaran et al. L'on notera des valeurs critiques que la borne supérieure reprend les valeurs pour

lesquelles les variables sont intégrées d'ordre 1 $I(1)$ et la borne inférieure concernent les variables $I(0)$. Ainsi :

Bornes Critiques

- Borne inférieure : Supposée que toutes les variables explicatives sont $I(0)$.
- Borne supérieure : Supposée que toutes les variables explicatives sont $I(1)$.

Après avoir calculer la F-stat (Fisher calculé) on teste comme suit :

- Si F-stat $>$ Borne supérieure : Existence de cointégration.
- Si F-stat $<$ Borne inférieure : Non-existence de cointégration.
- Si Borne inférieure $<$ F-stat $<$ Borne supérieure : On ne peut rien conclure.

Remarque 2.4.1 (*Mise en œuvre avec R*)

Le test de cointégration aux bornes peut être réalisé en utilisant la commande `bounds_f_test` du package *ARDL* dédié à cet effet.

Les objets principaux nécessaires pour ce test sont :

- un objet de classe `ardl` ou `uecm` correspondant au modèle ARDL ou au modèle à correction d'erreur non restreint estimé sur vos données,
- un paramètre `case` qui spécifie la forme du modèle (par exemple, présence ou absence d'intercept et/ou de tendance),
- des paramètres optionnels tels que le niveau de signification `alpha`, le choix entre calcul asymptotique ou exact, etc.

Dans cet exemple, nous utilisons le jeu de données `denmark` inclus dans le package *ARDL*.

Il s'agit de données macroéconomiques danoises (Johansen & Juselius, 1990) ([33]) qui contiennent plusieurs variables économiques mesurées sur une période donnée.

Le but du modèle est d'étudier la relation de long terme entre la demande réelle de monnaie (LRM) et ses déterminants économiques clés : le revenu réel (LRY) et les coûts d'opportunité liés aux taux d'intérêt (IBO et IDE).

- LRM : logarithme de la demande réelle de monnaie,
- LRY : logarithme du revenu réel,
- IBO : taux d'intérêt obligataire,
- IDE : taux d'intérêt sur dépôts bancaires.

Ces variables sont utilisées pour illustrer la modélisation ARDL et le test de cointégration aux bornes.

Avant de procéder au test de cointégration aux bornes, il est nécessaire de vérifier la stationnarité des quatre séries. Les résultats de test KPSS sont présentés dans la table 2.1.

Passons maintenant au programme R qui illustre l'application du test de cointégration aux bornes.

Chapitre 2 : La cointégration linéaire

Série	KPSS p-valeur (niveau)	Conclusion
LRM	0.01 (non stationnaire)	Non stationnaire
LRY	0.01 (non stationnaire)	Non stationnaire
IBO	0.04 (non stationnaire)	Non stationnaire
IDE	0.10 (stationnaire)	Stationnaire

TABLE 2.1 – Résultats des tests KPSS en niveau pour les variables du jeu de données `denmark`

```

> library(ARDL)
> data(denmark)
> head(denmark)
      LRM      LRY      LPY      IBO      IDE
1974 Q1 11.63255 5.903658 -0.6187359 0.1547356 0.0940
1974 Q2 11.60415 5.873820 -0.5807479 0.1779912 0.0955
1974 Q3 11.58152 5.837818 -0.5428478 0.1705647 0.0955
1974 Q4 11.60185 5.812255 -0.5046041 0.1522273 0.0955
1975 Q1 11.58630 5.803945 -0.4864585 0.1342276 0.0885
1975 Q2 11.60450 5.786761 -0.4544386 0.1334805 0.0790
> ardl_model <- ardl(
+   formula = LRM ~ LRY + IBO + IDE,
+   data = denmark,
+   order = c(3, 1, 3, 2)
+ )
> summary(ardl_model)

Time series regression with "zooreg" data:
Start = 1974 Q4, End = 1987 Q3

Call:
dynlm::dynlm(formula = full_formula, data = data, start = start,
  end = end)

Residuals:
      Min       1Q   Median       3Q      Max
-0.029939 -0.008856 -0.002562  0.008190  0.072577

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.6202     0.5678   4.615 4.19e-05 ***
L(LRM, 1)      0.3192     0.1367   2.336 0.024735 *
L(LRM, 2)      0.5326     0.1324   4.024 0.000255 ***
L(LRM, 3)     -0.2687     0.1021  -2.631 0.012143 *
LRY            0.6728     0.1312   5.129 8.32e-06 ***
L(LRY, 1)     -0.2574     0.1472  -1.749 0.088146 .
IBO           -1.0785     0.3217  -3.353 0.001790 **
L(IBO, 1)     -0.1062     0.5858  -0.181 0.857081

```

```

L(IBO, 2)      0.2877      0.5691      0.505 0.616067
L(IBO, 3)     -0.9947      0.3925     -2.534 0.015401 *
IDE           0.1255      0.5545      0.226 0.822161
L(IDE, 1)    -0.3280      0.7213     -0.455 0.651847
L(IDE, 2)     1.4079      0.5520      2.550 0.014803 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0191 on 39 degrees of freedom
Multiple R-squared:  0.988,    Adjusted R-squared:  0.9843
F-statistic: 266.8 on 12 and 39 DF,  p-value: < 2.2e-16

> bounds_test_result <- bounds_f_test(ardl_model, case = 3, alpha = 0.05)
> print(bounds_test_result)

      Bounds F-test (Wald) for no cointegration

data:  d(LRM) ~ L(LRM, 1) + L(LRY, 1) + L(IBO, 1) + L(IDE, 1)
+ d(L(LRM, 1)) + d(L(LRM, 2)) + d(LRY) + d(IBO) + d(L(IBO, 1))
+ d(L(IBO, 2)) + d(IDE) + d(L(IDE, 1))
F = 6.2059, Lower-bound I(0) = 3.2244, Upper-bound I(1) = 4.3542,
p-value = 0.004146
alternative hypothesis: Possible cointegration
null values:
      k      T
      3 1000

> print(bounds_test_result$tab)
      statistic Lower-bound I(0) Upper-bound I(1) alpha      p.value
F 6.205875          3.224392          4.354236  0.05 0.004146084

```

- Comme la p -value = 0.004, donc nous rejetons H_0 au seuil 5%. Ce résultat indique une relation de cointégration forte entre les variables du modèle, validant ainsi l'existence d'un équilibre de long terme.

Pour conclure, nous pouvons donc dire que Pesaran et al (2001) ([49]) ont introduit l'approche ARDL bounds pour tester la cointégration. L'un des avantages de cette approche est qu'elle permet aux régressions d'avoir des variables explicatives d'ordres d'intégration mixtes ou inconnus, $I(0)$ ou $I(1)$, ce que les tests de cointégration conventionnels n'autorisent pas. Cependant, il existe la possibilité de tomber dans des cas dégénérés de non-cointégration à partir des tests de limites ARDL (C'est à dire que les tests classiques de cointégration ne permettent pas de conclure correctement à l'existence d'une relation d'équilibre à long terme entre les variables, en raison de la non-significativité des coefficients sur les niveaux retardés de la variable dépendante ou des variables indépendantes.). Cet aspect a souvent été ignoré dans les applications empiriques.

2.5 Cointégration non linéaire

La cointégration non linéaire est une extension du concept classique de cointégration, qui permet de modéliser des relations d'équilibre à long terme entre séries temporelles tout en tenant compte de dynamiques non linéaires. Contrairement à la cointégration linéaire, où une combinaison linéaire des séries est stationnaire, la cointégration non linéaire introduit des fonctions non linéaires pour capturer des comportements plus complexes.

Une première question importante concerne l'intégration des deux hypothèses de non-linéarité et de non-stationnarité. Les tests stationnaires conventionnels, qui ignorent la dynamique non linéaire, peuvent être biaisés. Leur configuration peut ne pas être pertinente dans un monde non linéaire. Voici un exemple. Considérons les deux modèles suivants

$$X_t = f(X_{t-1}) + \sum_{i=1}^t \phi^{t-i} \varepsilon_t, \quad t = 1, 2, \dots \quad (2.5.1)$$

et

$$X_t = \phi X_{t-1} + f(X_{t-1}) + \varepsilon_t, \quad t = 1, 2, \dots \quad (2.5.2)$$

Où $\varepsilon_t \sim iid(0, \sigma^2)$ et f est une fonction non linéaire qui satisfait la condition $|f(X_{t-1})| < \infty \quad \forall t$. Supposons que $\phi = 1$. Alors, l'équation (2.5.1) a une composante linéaire non stationnaire donnée par son deuxième terme.

Cependant, que peut-on dire de X_t ?

De toute évidence, la réponse dépend des propriétés de la fonction non linéaire. Si la fonction est bornée, il n'est pas certain que X_t aura une composante permanente. Et nous ne pouvons pas utiliser l'approche de la racine unitaire pour étudier les propriétés de stationnarité de f , car c'est une fonction non linéaire. Sous l'hypothèse que $\phi = 1$, l'équation (2.5.2) s'écrit

$$\Delta X_t = f(X_{t-1}) + \varepsilon_t, \quad t = 1, 2, \dots \quad (2.5.3)$$

Demander si ΔX_t est $I(0)$ peut n'avoir aucun sens ici en raison de la non-linéarité de la fonction f .

Donc, probablement, nous avons besoin d'une analyse plus fine des propriétés de persistance dans un cadre non linéaire afin de distinguer entre la mémoire à court terme et la mémoire à long terme dans les séries chronologiques ([16]).

Définition 2.5.1 (Relation de cointégration non linéaire) Une série temporelle y_t est dite cointégrée de manière non linéaire avec x_t s'il existe une fonction non linéaire $f(\cdot)$ telle que :

$$u_t = y_t - f(x_t)$$

est un processus stationnaire, même si y_t et x_t sont non stationnaires ($I(1)$).

Remarque 2.5.1 Cas particulier : Si $f(\cdot)$ est linéaire, on retrouve la cointégration standard ([18]).

2.6 Modèles à correction d'erreur non linéaires (NECM)

Les modèles à correction d'erreur non linéaire représentent une généralisation fondamentale du cadre linéaire traditionnel, permettant de modéliser des mécanismes d'ajustement asymétriques ou dépendants de régimes.

La formulation générale des modèles NECM pour un vecteur $X_t = (y_t, x_t)'$ s'exprime sous la forme dynamique suivante :

$$\Delta X_t = \psi \Delta X_{t-1} + F(X_{t-1}) + v_t \quad (2.6.1)$$

où la fonction $F(X_{t-1}) = J(z_{t-1})$ intègre la non-linéarité via l'écart à l'équilibre $z_t = y_t - \beta x_t$. Pour garantir la stabilité et la cointégration du système, les auteurs énoncent des conditions essentielles, nous présenterons ses conditions sous forme d'un théorème.

Théorème 2.6.1 (Théorème de représentation de Granger ([19]))

Considérons le modèle de correction non linéaire (NECM) pour le vecteur $X_t = (y_t, x_t)'$ de dimension (2×1) , donné par l'équation (2.6.1). Supposons que :

- (a) v_t est un processus α -mélangeant (défini plus loin) d'ordre $-s/(s-2)$ pour $s > 2$.
- (b) $\sum_t v_t$ n'est pas NED^[2] (défini plus loin) sur une suite α -mélangeante .
- (c) $E\|v_t\|^2 < \Delta_v$.
- (d) $F(X_{t-1}) = J(Z_{t-1})$, où $Z_t = y_t - \beta x_t$, et $J(\cdot)$ est une fonction continûment différentiable qui satisfait une condition de Lipschitz généralisée .
- (e) Soit $SR(M) < 1$, où $SR(M)$ est le rayon spectral de la matrice M , et
- (f) pour un certain $\delta > 0$,

$$SR \left(\begin{pmatrix} \Psi_1 & \nabla J(Z) \\ \beta' \Psi_1 & 1 + \beta' \nabla J(Z) \end{pmatrix} \right) \leq 1 - \delta.$$

Ces conditions assurent que

1. ΔX_t et Z_t sont simultanément NED sur la suite α -mélangeante (v_t) , où $u_t = v_{y,t} - \beta' v_{x,t}$, et
2. X_t est I(1).

Ce théorème donne des conditions suffisantes pour que des variables cointégrées soient générées par un modèle de correction d'erreur non linéaire.

Définition 2.6.1 (Processus α -mélangeant)

Soit $\{v_t\}$ une suite de variables aléatoires. Soit $\mathcal{S}_s^t = \sigma(v_s, \dots, v_t)$ la tribu engendrée. On définit les coefficients de α -mélange par :

$$\alpha_m = \sup_t \sup_{\substack{F \in \mathcal{S}_{-\infty}^t \\ G \in \mathcal{S}_{t+m}^+}} |P(G \cap F) - P(G)P(F)|$$

Le processus $\{v_t\}$ est dit fortement mélangeant (ou α -mélangeant) si $\alpha_m \rightarrow 0$ lorsque $m \rightarrow \infty$. Si $\alpha_m \leq m^{-a}$, on dit que $\{v_t\}$ est fortement mélangeant d'ordre $-a$.

[2]. NED signifie Near-Epoch Dependent (ou dépendant à proximité d'époque en français).

Définition 2.6.2 (Processus NED)

Soit $\{w_t\}$ une suite de variables aléatoires avec $E\{w_t^2\} < \infty$ pour tout t . On dit que $\{w_t\}$ est NED (Near-Epoch Dependent) sur la suite sous-jacente $\{v_t\}$ de taille $-a$ si $\varphi(n)$ est de taille $-a$, où $\varphi(n)$ est donnée par :

$$\sup \|w_t - E_t^{t+n} w_t\|_2 = \varphi(n)$$

où $E_t^{t+n}(w_t) = E(w_t | v_t, \dots, v_{t+n})$ et $\|\cdot\|_2$ est la norme L_2 d'un vecteur aléatoire, définie comme $E^{1/2}|z|^2$ où $|\cdot|$ désigne la norme euclidienne.

2.6.1 Tests de cointégration non linéaire

Face aux limitations des tests de cointégration linéaire traditionnels, plusieurs approches spécifiques ont été développées pour détecter des relations non linéaires.

Aparicio, Escribano et Sipols (2006) ([2]) ont proposé le test RUR (*Range Unit Root*), un test de racine unitaire basé sur les premières différences des étendues d'une série. L'étendue représente l'écart entre les valeurs maximale et minimale observées à un instant donné. La fréquence de nouveaux records diminue plus rapidement pour les séries stationnaires que pour les séries non stationnaires, en raison de la variance croissante de ces dernières. Cette propriété permet au test RUR d'offrir une alternative robuste aux tests traditionnels comme ADF.

Escribano, Santos et Sipols (2008) ([20]) ont développé un test de cointégration basé sur les statistiques d'ordre induit (IOC). Ce test exploite la synchronisation des records entre séries cointégrées : leur statistique compare les distributions empiriques des séries ordonnées à l'aide d'une distance de Kolmogorov-Smirnov. Il est robuste aux non-linéarités monotones et aux ruptures structurelles, et ne nécessite pas d'estimation préalable de la relation de cointégration.

ATIL, FELLAG et al. (2024) ([4]) ont affirmé que, généralement, les tests traditionnels ne parviennent pas à détecter la présence de non linéarités dans les séries cointégrées. Dans leur article, les auteurs ont alors construit une nouvelle procédure pour les tests de cointégration basée sur une modification du test de Engle et Granger. Ces tests ainsi construits en utilisant des statistiques d'ordre ont la capacité de détecter des preuves de cointégration là où les tests de cointégration classiques ne la détectent pas.

2.6.2 Test basé sur les rangs (Range Unit Root (RUR))

Le test RUR, développé par Aparicio, Escribano et Sipols ([2]), exploite les propriétés des étendues mobiles définies comme $R_t(k) = \max_{t-k \leq i \leq t} y_i - \min_{t-k \leq i \leq t} y_i$. La statistique du test s'exprime formellement :

$$RUR = \frac{1}{T} \sum_{t=2}^T \frac{\Delta R_t(k)}{\hat{\sigma}_{\Delta R}}$$

où $\hat{\sigma}_{\Delta R}$ représente un estimateur robuste de la variance. Ce test présente l'avantage majeur de demeurer valide en présence de valeurs aberrantes et de transformations monotones des données, conservant une puissance élevée contre l'hypothèse alternative de cointégration non linéaire.

2.6.3 Test RCC (Record Counting Cointegration)

Le test RCC, proposé par Escribano et al (2006) ([21]), repose sur la synchronisation des valeurs extrêmes (*records*) entre séries cointégrées. La statistique du test compare la proportion de *records* simultanés selon l'expression suivante :

$$\text{RCC} = \frac{\sum_{t=2}^T \mathbf{1}(\text{record à } t \text{ pour } y_t \text{ et } x_t)}{\sqrt{\widehat{\text{Var}}(\mathbf{1}_{y,t})\widehat{\text{Var}}(\mathbf{1}_{x,t})}}$$

où :

- $\mathbf{1}(\cdot)$ est la fonction indicatrice qui vaut 1 si y_t et x_t battent simultanément un record (maximum ou minimum historique) à la date t

Remarque

- Un RCC significativement élevé suggère une cointégration
- Un RCC proche de 0 indique l'absence de relation à long terme

2.6.4 Test basé sur les statistiques d'ordre induit (IOC)

Un autre test de cointégration robuste aux non-linéarités et aux ruptures structurelles est basé sur les statistiques d'ordre induit. En particulier, Escribano et al (2008) ([20]) a considéré que deux séries y_t et x_t sont cointégrées (qu'elles soient linéaires ou non linéaires) si les séries d'ordre induit correspondantes sont tracées autour de la droite à 45°. Leur statistique de test compare ces deux séries d'ordre induit en comparant leurs distributions empiriques respectives, la distribution empirique de y_t et la distribution empirique de y_t induite par x_t , en utilisant la statistique de type Kolmogorov-Smirnov suivante

$$KS = \sup_{j=1,\dots,n} \left| \hat{F}_Y^{(n)}(\hat{y}_{l_s}^{(j,n)}) - \hat{F}_Y^{(n)}(y_{j,n}) \right| \quad (2.6.2)$$

où $\hat{F}_Y^{(n)}(y)$ est la fonction de distribution empirique obtenue à partir d'un échantillon de taille n de la série y_t , c'est-à-dire

$$\hat{F}_Y^{(n)}(y) = \frac{1}{n} \sum_{t=1}^n \mathbf{1}(y_t \leq y),$$

2.7 Approche Paramétrique

L'approche paramétrique postule une forme fonctionnelle spécifique pour la relation de cointégration et le mécanisme d'ajustement, permettant une interprétation économique directe des paramètres estimés.

La relation de long terme s'exprime généralement sous la forme

$$y_t = f(x_t, \beta) + v_t \quad (2.7.1)$$

Où f désigne une fonction connue dépendant de paramètres β . Parmi les spécifications courantes, la fonction logistique

$$y_t = \beta_1 + \beta_2 / (1 + \exp(-\beta_3(x_t - \beta_4))) \quad (2.7.2)$$

$\nu \setminus n$	100	250	500	1,000
0.01	0.3564	0.3665	0.3812	0.3726
0.025	0.4059	0.4382	0.4591	0.4346
0.05	0.4653	0.4980	0.4830	0.4915
0.10	0.5248	0.5697	0.5649	0.5794
0.90	0.9505	0.9681	0.9721	0.9780
0.95	0.9703	0.9801	0.9860	0.9870

TABLE 2.2 – Les valeurs critiques de la statistique de KS

Modélise des transitions progressives entre régimes économiques, tandis que la fonction exponentielle

$$y_t = \beta_1 x_t + \beta_2 \exp(-\beta_3 x_t) \quad (2.7.3)$$

Convient particulièrement aux processus de convergence avec saturation. Pour les mécanismes de correction d'erreur, les modèles NECM définis plus haut adoptent généralement la structure

$$\Delta y_t = \alpha(y_{t-1} - g(x_{t-1}, \beta))h(z_{t-1}; \gamma) + \dots \quad (2.7.4)$$

Où $h(\cdot)$ représente une fonction paramétrique définie a priori.

2.7.1 Méthode d'estimation

L'estimation des paramètres s'effectue principalement par maximum de vraisemblance ou moindres carrés non linéaires, ce dernier minimisant la fonction objective

$$\min_{\theta} \sum_{t=1}^T [\Delta y_t - f(\mathbf{x}_t; \theta)]^2 \quad (2.7.5)$$

où θ regroupe l'ensemble des paramètres structurels.

Cependant, les auteurs soulignent que cette approche souffre de risques substantiels de biais en cas de mauvaise spécification fonctionnelle, particulièrement lorsque la non-linéarité affecte simultanément la relation de long terme et le mécanisme d'ajustement. De plus, l'estimation des modèles à seuils présente des difficultés computationnelles importantes liées à la non-convexité de la fonction objective.

2.8 Approche Non Paramétrique

Des estimations non paramétriques des relations de cointégration non linéaire ont déjà été effectuées par Granger et Hallman (1991) ([28]), mais ce n'est qu'avec les travaux récents de Karlsen HA, Tjøstheim D (2001) ([35]), Karlsen HA et al (2007) ([34]) et de Phillips PCB, Park JY (1998) ([54]) qu'une théorie d'estimation non paramétrique pour les processus non stationnaires a été développée. Phillips PCB, Park JY (1998) ([54]) ont considéré une théorie basée sur des arguments de temps local, tandis que Karlsen HA,

Tjøstheim D (2001) ([35]), Karlsen HA et al (2007) ([34]) ont utilisé la théorie des processus de Markov null-récurrents. Une comparaison des deux méthodologies est discutée dans l'article de Bandi FM (2004) [5], où des asymptotiques non paramétriques proches de l'intégration sont étudiées.

2.8.1 Fonction de transfert

Les auteurs estiment la fonction de transfert $f(x_t)$ dans le modèle de régression non linéaire

$$y_t = f(x_t) + v_t,$$

où les séries y_t et x_t sont des processus non stationnaires univariés observés, et v_t est un processus stationnaire non observé. Karlsen et al (2007) ([34]) ont étudié l'estimation non paramétrique par noyau de $f(x_t)$ définie par

$$\hat{f}(x) = \frac{\sum_{t=0}^n y_t K_{x,h}(x_t)}{\sum_{t=0}^n K_{x,h}(x_t)},$$

où

$$K_{x,h}(x_t) = h^{-1} K\left(\frac{y - x}{h}\right),$$

K est une fonction noyau et h un paramètre de lissage (largeur de bande). Ces auteurs étudient la théorie asymptotique pour $\hat{f}(x)$ sous certaines conditions de régularité et différentes hypothèses sur la dépendance entre x_t et v_t . Les hypothèses particulièrement adaptées au cadre de la cointégration non linéaire sont celles qui permettent une dépendance entre x_t et v_t .

La famille de processus non stationnaires considérée est la classe dite des processus de Markov β -null-récurrents ([41]) satisfaisant une restriction sur la distribution de queue du temps de récurrence. Cette classe est suffisamment large pour contenir la marche aléatoire, les processus à racine unitaire ainsi que d'autres processus non stationnaires non linéaires. Il est montré que la théorie d'estimation non paramétrique diffère de celle du cas stationnaire, avec des vitesses de convergence plus lentes, comme attendu. Cette nouvelle théorie asymptotique non paramétrique ouvre la voie à de futurs développements dans l'inférence des modèles de cointégration non linéaire.

Remarque 2.8.1 *Karlsen et al (2007) ([34]) ont établi que dans un contexte non stationnaire, la vitesse de convergence des estimateurs non paramétriques chute significativement à $T^{1/4}$ contre $T^{1/2}$ en environnement stationnaire, avec une distribution limite de type mixte gaussienne. Les auteurs identifient plusieurs défis majeurs associés à cette approche, notamment la malédiction de la dimension qui limite sévèrement son applicabilité aux systèmes multivariés, la sélection délicate de la fenêtre h via des méthodes de validation croisée souvent instables, et une interprétation économique moins directe que dans le cadre paramétrique traditionnel.*

2.9 Review sur la cointégration non linéaire

Il est connu, à travers une large littérature, que lorsque deux séries sont non stationnaires et intégrées du même ordre, une combinaison linéaire de ces deux dernières peut

être stationnaire, ce qui les rend cointégrées. Cependant, il est utile de se poser la question de savoir si une combinaison non linéaire peut aussi aboutir à une série stationnaire.

Cette question est abordée par divers auteurs même si, à ce jour, il n'existe pas de théorie générale de la cointégration non linéaire. Par exemple, Dufrénot et Lignon (2012) ([16]) consacrent un ouvrage entier aux domaines de l'économétrie des séries chronologiques qui traitent simultanément de la non linéarité et la non stationnarité. En effet, selon ces auteurs, l'étude simultanée de ces deux aspects ne s'est développée que lorsque des arguments d'ordre économique l'exigent.

Dufrénot et Mignon (2012) ([16]) affirment que l'idée prédominante aujourd'hui est qu'une série non stationnaire présente des écarts persistants par rapport à ses composantes de long terme (tendances déterministes ou stochastiques). Ces écarts sont modélisés de diverses manières : modèles à racine unitaire, processus fractionnellement intégrés, modèles avec décalages de tendance temporelle, etc. Cependant, de nombreux autres comportements inhérents aux processus non stationnaires ne sont pas reflétés dans les modèles linéaires. Par exemple, les variables économiques à distribution mixte, ou les processus dépendants de l'état, subissent des épisodes de dynamique changeante. Étant donné que les modèles non linéaires présentent fréquemment des équilibres multiples, une dynamique non réversible et des changements de régime, ils sont susceptibles de présenter des comportements non stationnaires qui diffèrent de ce qui est habituellement observé dans les modèles linéaires.

Finalement, force est de constater que les modèles de cointégration linéaire conventionnels ne sont pas capables de reproduire le type de non-stationnarité inhérent aux modèles non linéaires. De nouveaux outils sont nécessaires. La présentation de ces nouveaux outils constitue le cœur de leur ouvrage dans lequel ils rendent compte des sujets récents apparus dans la littérature.

Les auteurs présentent dans leur ouvrage plusieurs interprétations de la cointégration non linéaire ainsi que les modèles non linéaires à correction d'erreur sont introduits, en particulier, lorsque qu'on observe le mécanisme d'ajustement à l'équilibre rendant la relation à long terme entre les variables non linéaire.

À titre d'exemple, Dufrénot et Mignon (2012) ([16]) proposent l'exemple basique suivant mettant en relation deux variables X_t et Y_t . Dans le contexte classique, (X_t) et (Y_t) sont $I(1)$ et $Z_t = Y_t - aX_t$ est $I(0)$. Mais dans le contexte non linéaire, Z_t s'écrit sous la forme $Z_t = f(X_t) - g(Y_t)$ avec $(f(X_t))$ et $(g(Y_t))$ non stationnaires mais (Z_t) stationnaire.

Lin and Granger (2004) ([39]) posent trois importantes questions reprises dans la thèse de Haichen Shi (2018) ([57]) :

Première question : *Quelles propriétés peuvent changer ou ne pas changer après avoir effectué une transformation non linéaire sur une série intégrée ?*

Granger et Hallman (1991) ([27]) étudient l'effet de transformations non linéaires de séries intégrées, en particulier $I(1)$ sur les niveaux d'interrogation suivants.

- (a) Si X_t est intégrée, est-ce qu'on peut avoir $Z_t = f(X_t)$ intégrée pour une certaine fonction f ?
- (b) Si X_t et Y_t sont $I(1)$ cointégrées, est-ce qu'on peut avoir $g(X_t)$ et $g(Y_t)$ le sont aussi ? Ils montrent alors que les séries intégrées ne sont généralement pas cointégrées avec une transformation non linéaire. Cependant, si la même transformation est appliquée à une paire de séries cointégrées, la série transformée peut toujours être cointégrée (voir plus bas).

Dittmann et Granger (2002) ([15]) s'intéressent, quant à eux, aux séries fractionnaires

intégrées. L'une des conclusions intéressantes de leur étude est que toute transformation de puissance entière d'une série $I(d)$ non stationnaire produit une tendance déterministe de la moyenne et de la variance.

Deuxième question : *Comment estimer la fonction de transformation non linéaire ?*

Il existe deux approches sur cette question : paramétrique et non paramétrique. Comme l'estimation des paramètres dépend fortement de la forme des transformations non linéaires, divers articles publiés se concentrent sur l'établissement de la distribution asymptotique de ces estimateurs non linéaires. Haichen Shi (2018) ([58]) cite, entre autres, Park and Phillips (2001) ([45]), Chang et al. (2001) ([8]), Karlsen et al. (2007) ([34]) et Wang et Phillips (2009) ([64]).

Troisième question : *Comment tester l'existence d'une relation de cointégration non linéaire ?*

Haichen Shi (2018) ([58]) cite, par exemple, Breitung (2001) ([6]) qui a proposé un test de rang pour la cointégration non linéaire, basé sur la différence entre les séquences de rangs. Il donne aussi le cas de Choi et Saikkonen (2010) ([9]) qui développent une méthode de test pour la cointégration non linéaire des variables $I(1)$ consistant à diviser le résidu du modèle en sous-échantillons et à appliquer un test de racine unitaire sur ces sous-échantillons. Ils démontrent qu'à mesure que la taille de l'échantillon s'approche de l'infini, la distribution limite des statistiques de test est indépendante de la forme des transformations non linéaires.

Haichen Shi (2018) ([58]) affirme que de nombreux systèmes d'ingénierie présentent des relations non linéaires entre les caractéristiques mesurées d'intérêt et que dans le contexte de la surveillance de la santé structurelle notée SHM (Structural Health Monitoring), la nécessité d'une théorie non linéaire s'est toutefois rapidement imposée. Il donne des exemples pratiques intéressants (voir chapitre 3).

Haichen Shi (2018) ([58]) explique qu'une approche potentiellement simple de la cointégration non linéaire, comme discutée par Dufrénot et Mignon (2012) ([16]) et partiellement adoptée par Cross et Worden (2011) ([12]), consiste simplement à étendre la deuxième étape de la démarche de Engle-Granger à une régression non linéaire. Par exemple du type

$$Y_t = f(X_t) + \epsilon_t$$

où $f(\cdot)$ est une fonction de régression estimée.

Haichen Shi (2018) ([58]) ajoute que, par exemple, dans le cas pratique des réseaux de neurones suggérés par Dufrénot et Mignon (2012) ([16]), tout algorithme d'apprentissage automatique approprié pourrait en principe être utilisé, et c'est sur cette idée que s'est concentrée une partie de la littérature récente sur la cointégration non linéaire. Gestel et al (2006) ([63]) intègrent pour cela une approche Bayésienne. C'est l'intégration de l'approche Bayésienne dans la régression non linéaire qui motive Haichen Shi (2018) ([57]) pour adopter la régression par processus gaussien notée GP afin d'apprendre à la machine la relation requise dans l'équation $Y_t = f(X_t) + \epsilon_t$ à partir des données. La régression par processus gaussien qui est une approche d'apprentissage automatique Bayésienne est présentée dans le chapitre 3.

Revenons maintenant à Granger et Hallman qui se sont intéressés aux transformations non linéaires de séries intégrées. Ces auteurs affirment que le test de Dickey-Fuller classique (DF) apparaît souvent comme trop sensible aux transformations non linéaires que la fonction d'autocorrélation empirique. Ils répondent, par leur étude, aux questions posées citées plus haut que si X_t est intégrée, alors $Z_t = f(X_t)$ peut être intégrée pour une fonction f affine et que si X_t et Y_t sont $I(1)$ cointégrées, alors $g(X_t)$ et $g(Y_t)$ le sont

Chapitre 2 : La cointégration non linéaire

aussi si $g(\cdot)$ est une fonction homogène, i.e $g(\alpha x) = \alpha^\lambda g(x)$ auquel cas le paramètre de cointégration est α^λ .

En fait, ils se sont posé la question de savoir si les tests DF et ADF fonctionnent convenablement avec des séries transformées par des fonctions non linéaires. Ils donnent deux tables de fonctions de répartitions empiriques des statistiques DF et ADF obtenues après diverses transformations non linéaires d'une marche aléatoire gaussienne du type

$$X_t = X_{t-1} + \epsilon_t$$

avec $X_0 = 0$ et $\epsilon_t; t = 1, 2, \dots, n$ des v.a *iid* telles que $Prob(\epsilon = 1) = Prob(\epsilon = -1) = 1/2$ sous des conditions données par Phillips (1987). Dans ces tests, l'hypothèse nulle H_0 est que les séries sont $I(1)$ et que sous H_1 , les séries sont $I(0)$. Les différentes fonctions non linéaires de X_t sont $X_t, X_t^2, X_t^3, |X_t|, sgn(X_t), \sin(X_t), \exp(X_t), \ln(X_t + 75)$ et $1/(X_t + 75)$. Les auteurs aboutissent à la conclusion que, non seulement H_0 est toujours rejetée pour la fonction $\sin(X_t)$, mais aussi pour les fonctions $sgn(X_t)$ et $\exp(X_t)$ qui sont des processus à mémoire longue. Les autres fonctions aboutissent aussi à des rejets fréquents sauf pour les deux dernières. Rappelons que dans la stratégie de Box et Jenkins, le corrélogramme permet de décider si la série est $I(0)$ ou $I(1)$. Cela veut dire que si l'ACF décroît lentement, on choisit souvent une série $I(1)$. Dans leur étude, DF et ADF semblent plus sensibles que l'usage de l'ACF. Ils ajoutent que l'on ne peut pas construire un test invariant à toute transformation de la série à tester. Mais ils précisent que Granger et Hallman (1991) ([27]) montrent que la transformation $\sin(X_t)$ aboutit à une stationnarité de l'AR(1). Cela suggère que toute transformation périodique aboutit à une stationnarité de la série.

Granger et Hallman (1991) ([27]) proposent alors de construire des tests basés sur les rangs comme cela se fait habituellement dans le cas non paramétrique. Ils s'appuient sur le fait que les tests basés sur les rangs ont des lois non affectées par des transformations monotones des données. Ils définissent les rangs R_t de la série X_t par

$$R_t = \text{rang de } X_t \text{ parmi } X_1, X_2, \dots, X_n$$

Ils choisissent H_0 comme étant l'hypothèse qu'il existe une transformation strictement monotone de la série à tester qui a présenté une racine unitaire. Ils construisent le test DF et ADF sur les rangs et ils les appellent RDF et RADF. Ils concluent alors que ces derniers sont bien plus puissants et même plus robustes que les tests usuels DF et ADF. Toutefois, les auteurs recommandent d'utiliser aussi bien les tests usuels DF et ADF que leurs versions RDF et RADF à cause des incertitudes qui concernent la linéarité des processus étudiés.

En termes de cointégration, pour répondre à la question "si X_t est $I(1)$, est-ce que X_t et $g(X_t)$ peuvent être cointégrées pour une certaine fonction $g(\cdot)$?", Granger et Hallman (1991) ([27]) proposent le modèle $X_t = m + X_{t-1} + e_t$ avec (e_t) *iid* $N(0, \sigma^2)$ et ils montrent qu'on peut répondre par "oui", si $g(x) = \alpha x^k$ avec $k = 1$. Ce qui n'est pas le cas pour une fonction du type $g(x) = \exp(\lambda x)$ par exemple.

Concernant l'autre question "Si X_t et Y_t sont $I(1)$ cointégrées, est-ce qu'on peut avoir $g(X_t)$ et $g(Y_t)$ le sont aussi?", ils proposent d'y répondre sur le modèle $Y_t = \alpha X_t + \epsilon_t$ avec (ϵ_t) *iid* de moyenne zéro et indépendantes de e_t . Ils aboutissent à la conclusion que les $g(X_t)$ et $g(Y_t)$ sont cointégrées dans le cas où $\alpha = 1$ ou si $g(\cdot)$ vérifie la propriété d'homogénéité $g(\alpha x) = \alpha^\lambda g(x)$ est homogène. Finalement, on ne peut pas répondre exclusivement, ni par oui, ni par non, aux deux questions posées plus haut. On peut même avoir des situations où X_t et Y_t non cointégrées mais X_t^2 et Y_t^2 le sont.

Pour terminer ce review, on peut citer Phillips et Ouliaris (1990) ([51]) qui développent une théorie asymptotique pour les tests de cointégration basés sur les résidus. Ces tests utilisent des procédures conçues pour détecter la présence d'une racine unitaire dans les résidus de régressions (cointégrantes) entre les niveaux de séries chronologiques économiques. À cet effet, deux tests sont proposés par ces auteurs, dont l'un est invariant à la normalisation de la régression cointégrante. Ils montrent alors que ces tests qui présentent des distributions limites données en termes de mouvement brownien standard. Park et Philips (1999) ([44]) proposent une théorie asymptotique de processus stochastiques générés par des transformations non linéaires de séries intégrées. Diverses fonctions non linéaires de séries intégrées, telles que les séries temporelles ARIMA, sont étudiées, et les distributions asymptotiques des moments d'échantillon de ces fonctions sont obtenues et analysées. Les transformations considérées dans leur article incluent diverses fonctions utilisées en analyse statistique non linéaire pratique. Il est démontré que leur théorie asymptotique est très différente de celle des processus intégrés et des séries temporelles stationnaires. Chang et al (1999) ([7]) développent une théorie asymptotique pour une classe générale de modèles de régression non linéaires en généralisant les résultats de Phillips et Hansen (1990) ([50]) portant sur la cointégration des modèles de régression linéaires. Enfin, dans Park et Philips (2001) ([45]), une théorie asymptotique est développée pour la régression non linéaire avec des processus intégrés. Les modèles tiennent compte des effets non linéaires des séries temporelles à racine unitaire et traitent donc du cas de la cointégration non linéaire paramétrique. La théorie couvre les fonctions intégrables et asymptotiquement homogènes. Des conditions suffisantes pour une cohérence faible sont données et une théorie de distribution limite est fournie.

Chapitre 3

Simulations et applications

2.1 Introduction

Dans ce chapitre, nous proposons d'illustrer les aspects pratiques de la cointégration linéaire et non linéaire à travers quelques études de simulations et applications basées sur des données réelles. L'objectif est de montrer comment les outils théoriques développés dans le chapitre précédent peuvent être mobilisés pour étudier des phénomènes économiques, technologiques ou financiers.

Nous débutons par une application de la cointégration linéaire fondée sur l'étude de Cowpertwait et Metcalfe (2009) ([11]), qui s'intéressaient à la relation entre la livre sterling (UK) et l'euro (EU) avant le Brexit.

Nous abordons ensuite deux applications centrées sur la cointégration non linéaire, conceptuellement plus complexe mais indispensable pour modéliser des relations durables dans des systèmes dont le comportement évolue de manière non linéaire. La première est inspirée de travaux de Cross et Worden (2011) ([12]), qui appliquent ces méthodes à la surveillance de l'intégrité des structures (SHM – Structural Health Monitoring). Dans ce contexte, la cointégration permet de séparer les effets dus à l'environnement (comme la température) des signaux indiquant des dommages structurels. Cette démarche sera complétée par une étude spécifique du pont Z24, un pont en béton situé en suisse et surveillé expérimentalement avant sa démolition, dont les données présentent des relations non linéaires particulièrement bien adaptées à ce type d'analyse.

Enfin, nous terminons ce chapitre par une application financière plus exploratoire, à travers une simulation de stratégie d'arbitrage statistique par retour à la moyenne. Nous modélisons un portefeuille dont le comportement suit un processus AR(1), en supposant qu'il a été construit à l'aide de techniques de cointégration. L'objectif est ici d'évaluer la performance d'une stratégie simple de liquidation conditionnelle, via des simulations Monte Carlo.

Ces applications variées mettent en lumière la puissance de la cointégration – qu'elle soit linéaire ou non linéaire – pour détecter des relations persistantes entre séries temporelles, et pour construire des modèles à forte valeur explicative ou prédictive dans des contextes réels très différents.

3.2 Une application financière de la cointégration linéaire

Dans ce paragraphe, nous proposons de reproduire une application de la cointégration linéaire réalisée par Cowpertwait et Metcalfe où ils se sont intéressés à la relation qui pouvait exister entre la livre Sterling du Royaume uni et l'euro pendant une période située avant le Brexit. Dans cet exemple, les données portent sur les séries de taux de change quotidiens de la livre sterling notée UK, de l'euro EU, pour la période de janvier 2004 à décembre 2007, toutes exprimées avec le dollar américain. Le Royaume-Uni faisant partie, à cette époque, de la Communauté économique européenne (CEE), on pouvait penser que toute variation du taux de change de l'euro était susceptible de se répercuter sur le taux de change de la livre sterling. On peut donc conclure à des fluctuations communes aux deux séries ; en particulier, on pourrait penser, par exemple, que les deux séries pourraient partager une tendance stochastique commune. Voyons de près ces deux séries liées à l'euro et à la livre sterling que nous appelons EU et UK respectivement. La représentation de l'évolution des deux séries est donnée par la figure 3.1 suivante :

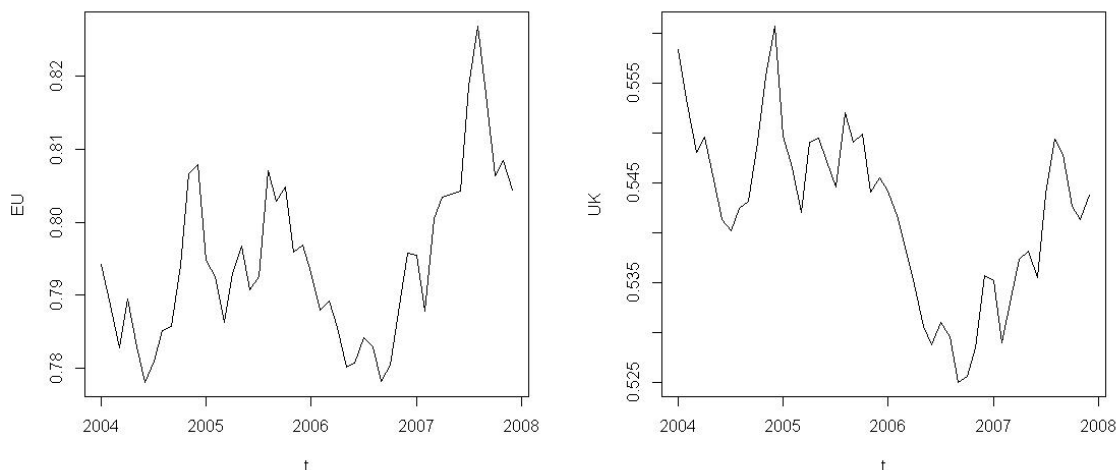


FIGURE 3.1 – Variation des monnaies EU (à gauche) et UK (à droite) de Janvier 2004 à décembre 2007

L'étude, par la méthode de Box et Jenkins, en particulier, des ACF et PACF des deux séries a conduit les auteurs à modéliser chacune par un AR(1).

Maintenant, la question qui nous intéresse est de savoir si ces deux séries sont cointégrées ou non. D'abord, nous allons vérifier leur stationarité. Le test de Dickey-Fuller donne, après application des p-valeur égales à 0.7399 pour EU et 0.4463 pour UK. Ceci veut dire qu'une racine unitaire est présente dans les deux séries. Les p-valeurs du test ADF sur les séries différenciées $\text{Diff}(\text{EU})$ et $\text{Diff}(\text{UK})$, quant à elles, sont égales à 0.01 pour les deux. Ce qui veut dire, finalement, les séries EU et UK sont toutes les deux intégrées $I(1)$.

Pour voir si elles ont une tendance commune, on peut commencer par voir comment elles sont liées. La figure 3.2 suivante montre qu'il y a une relation qui semble linéaire entre elles.

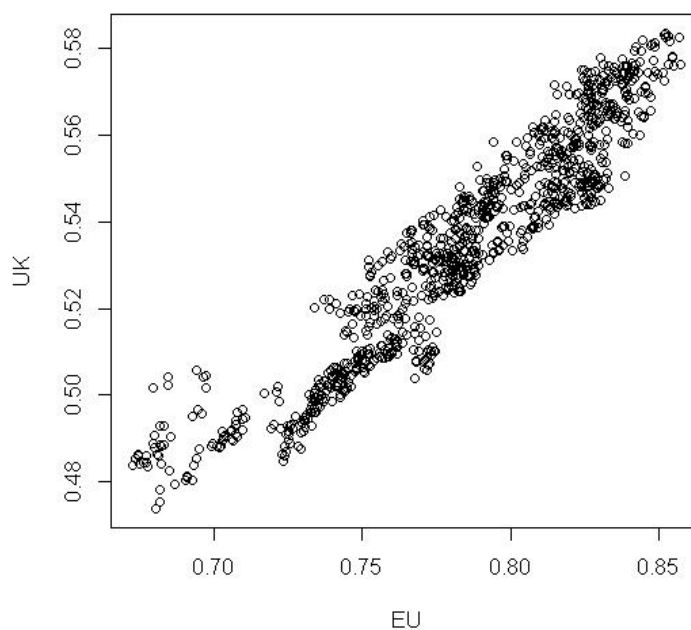


FIGURE 3.2 – Relation linéaire entre EU et UK

D'ailleurs le coefficient de corrélation entre EU et UK est égal à 0.946. Au nal, pour tester si une cointégration linéaire est possible entre ces deux séries, nous réalisons le test de Phillips-Ouliaris suivant qui donne

Phillips-Ouliaris Cointegration Test

data : cbind(fileUK, fileEU) Phillips-Ouliaris demeaned=-21.6618, Truncation lag parameter = 10

p-value = 0.04118

la p-valeur suggère donc que les deux séries sont cointégrées. Cowpertwait et Metcalfe (2009) ont également utilisé l'AIC pour affirmer que la série résiduelle obtenue est aussi stationnaire.

3.3 Une application de la cointégration non linéaire : Tendances linéaire versus une tendance quadratique

Dans ce qui suit, nous proposons d'étudier un modèle de cointégration proposé par Cross and Worden (2011) ([12]). Dans cet article il est question d'étendre la technique de cointégration, habituellement appliqué en économétrie vers le domaine technologique. En effet, un des problèmes largement abordé par différents auteurs est ce qu'on appelle "la fiabilité des structures" (Security Health Monitoring, en anglais) qui est utilisé sous l'abréviation SHM. Ces auteurs affirment que l'un des principaux problèmes rencontrés par l'application de la SHM aux structures comme les grands ouvrages (exemple : ponts) est de dissocier l'effet des changements environnementaux de celui des dommages physiques.

Cependant, dans le nouveau contexte de la SHM, les dommages peuvent souvent rendre une structure donnée non linéaire. Par conséquent, les séries chronologiques linéaires ne sont pas très adaptées à ces systèmes. A cet effet, Cross et Worden proposent deux approches de la cointégration non linéaire que nous ne développerons pas dans ce travail : la première est une méthode basée sur l'optimisation proposée par Storn et Price (1997) ([61]) et la seconde est une variante de ce qu'on appelle la procédure de Johansen basée sur la méthode du maximum de vraisemblance ([32]). Enfin, ils appliquent les concepts de la cointégration non linéaire avec des données réelles de SHM issues d'un projet de référence sur le pont de l'autoroute Z24, situé en Suisse, que nous décrirons plus loin.

Par ailleurs, des problèmes similaires peuvent survenir dans le contexte technologique où le machine learning joue un rôle prépondérant. Considérons, à titre d'exemple, un problème de régression appliqué à la robotique, où l'on souhaite que le contrôle robotique apprenne la dynamique inverse d'un bras robotisé. Il s'agit ici de relier l'état du bras (donné par les positions, vitesses et accélérations des articulations) aux couples correspondants. Un tel modèle peut ensuite être utilisé pour calculer les couples nécessaires au déplacement du bras le long d'une trajectoire donnée. Un autre exemple serait celui d'une usine chimique, où l'on pourrait souhaiter prédire le rendement en fonction de paramètres de procédé tels que la température, la pression, la quantité de catalyseur, etc. En général, nous désignons l'entrée par x et la sortie (ou cible) par y . L'entrée du jeu de données est généralement représentée par un vecteur x , car il existe généralement de nombreuses variables d'entrée. La cible y peut être continue (comme dans le cas de la régression) ou discrète (comme dans le cas de la classification). On dispose donc d'un ensemble de données D de n observations

$$D = \{(x_i, y_i); i = 1, 2, \dots, n\}$$

À partir de ces données d'entraînement, nous souhaitons établir des prédictions pour de nouvelles entrées x , l'entraînement étant inductif, que nous n'avons pas observées dans l'ensemble d'entraînement. Il est donc clair que le problème traité est inductif; nous devons passer des données d'entraînement finies D à une fonction f qui effectue des prédictions pour toutes les valeurs d'entrée possibles. Pour ce faire, nous devons formuler des hypothèses sur les caractéristiques de la fonction sous-jacente, car sinon toute fonction cohérente avec les données d'apprentissage serait également valide.

De nombreuses méthodes ont été proposées pour traiter le problème de l'apprentissage supervisé; nous en décrivons ici deux approches courantes. La première consiste à restreindre la classe de fonctions considérées, par exemple en ne considérant que les fonc-

tions linéaires de l'entrée. La seconde approche consiste (de manière assez générale) à attribuer une probabilité a priori à chaque fonction possible, en attribuant des probabilités plus élevées aux fonctions que nous considérons comme plus probables, par exemple parce qu'elles sont plus lisses que d'autres fonctions.

La première approche présente un problème évident : il faut déterminer la richesse de la classe de fonctions considérée ; Si nous utilisons un modèle basé sur une classe de fonctions (par exemple, des fonctions linéaires) et que la fonction cible n'est pas bien modélisée par cette classe, les prédictions seront mauvaises. On pourrait être tenté d'augmenter la flexibilité de la classe de fonctions, mais cela présente un risque de surapprentissage : Ceci peut conduire à un bon ajustement aux données d'apprentissage, mais obtenir de mauvais résultats lors des prédictions de test.

La deuxième approche semble poser un sérieux problème : il existe certainement un ensemble infini et dénombrable de fonctions possibles, et comment allons-nous calculer avec cet ensemble en temps fini ? C'est là que le processus gaussien entre en jeu. Un processus gaussien est une généralisation de la distribution de probabilité gaussienne. Alors qu'une distribution de probabilité décrit des variables aléatoires scalaires ou vectorielles (pour les distributions multivariées), un processus stochastique régit les propriétés des fonctions. On peut imaginer une fonction comme un très long vecteur, chaque entrée du vecteur spécifiant la valeur de la fonction $f(x)$ à une entrée x particulière.

De nombreux chercheurs se sont rendu compte que les réseaux neuronaux n'étaient pas si faciles à appliquer dans la pratique, en raison des nombreuses décisions à prendre : quelle architecture, quelles fonctions d'activation, quel taux d'apprentissage, etc., et de l'absence d'un cadre de principe pour répondre à ces questions.

Une grande partie de la théorie fondamentale et de nombreux algorithmes sont communs aux communautés des statistiques et de l'apprentissage automatique. Les principales différences résident peut-être dans les types de problèmes abordés et dans l'objectif de l'apprentissage. Au risque de simplifier, on pourrait dire qu'en statistique, l'accent est souvent mis sur la compréhension des données et des relations à l'aide de modèles fournissant des résumés approximatifs tels que des relations linéaires ou des indépendances. En revanche, l'apprentissage automatique vise principalement à réaliser des prédictions aussi précises que possible et à comprendre le comportement des algorithmes d'apprentissage.

Les processus gaussiens sont mathématiquement équivalents à de nombreux modèles connus, notamment les modèles linéaires bayésiens et les modèles splines (Ait Ramdane, 2024) ([1]). Du point de vue des processus gaussiens, les modèles peuvent être plus faciles à manipuler et à interpréter que les réseaux de neurones. Au sein de la communauté statistique, les processus gaussiens ont également été maintes fois évoqués, même s'il serait probablement excessif d'affirmer que leur utilisation est répandue, sauf pour certaines applications spécifiques telles que les modèles spatiaux en météorologie et en géologie, et l'analyse d'expériences informatiques. Il existe également une théorie riche pour les modèles de processus gaussiens dans la littérature sur l'analyse des séries chronologiques ; quelques pistes vers cette littérature sont données dans Rasmussen et Williams (2006, Annexe B) ([55]).

3.3.1 Processus gaussien

Définition 3.3.1 *Un processus gaussien est un ensemble de variables aléatoires, dont un nombre fini a une distribution gaussienne conjointe*

Soit un processus réel noté $f(x)$ de fonction moyenne $m(x)$ et de fonction variance

$k(x, x')$ définis par

$$m(x) = \mathbb{E}(x) \quad ; \quad k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x')))]$$

Un processus Gaussien est donc complètement déterminé par sa moyenne et sa matrice variance-covariance. On le note $f(x) \sim \mathcal{GP}(m(x), k(x, x'))$. Dans le cas présent, les variables aléatoires représentent la valeur de la fonction $f(x)$ au point x . Les processus gaussiens sont souvent définis dans le temps, c'est-à-dire lorsque l'ensemble d'indices des variables aléatoires est le temps. Ce n'est généralement pas le cas dans l'utilisation des GP ici où l'ensemble d'indices est \mathcal{X} et l'ensemble des entrées possibles, dont le domaine d'entrée pourrait être plus général, par exemple, \mathbb{R}^D .

Dans son travail, Shi (2018) ([57]) considère que la régression par processus gaussien peut être vue comme une approche d'apprentissage automatique bayésienne qui peut traiter les problèmes de régression non linéaire de telle manière que les intervalles de confiance pour les prédictions soient produits de manière naturelle. Dans ce cas par exemple, Shi (2018) ([57]) commence pour ce faire, par assigner au GP une loi a priori de la forme $f \sim \mathcal{GP}(0, K)$. La matrice de variance-covariance K a pour terme général $[K]_{ij} = k(x_i, x_j)$ avec

$$k_{x_i, x_j} = \sigma_f^2 \exp\left(-\frac{1}{2}(x_i - x_j)'l^{-2}(x_i - x_j)\right) + \sigma_n^2 \mathbb{I}$$

où σ_f^2 , l et σ_n sont considérés comme des hyperparamètres. Le modèle ainsi construit fait que $f(X)$ suit une loi $\mathcal{N}(0, K(X, X'))$. De plus, la fonction de vraisemblance de y sachant X et f est une loi $\mathcal{N}(f(X), \sigma_n^2 \mathbb{I})$. Enfin, l'auteur obtient la densité a posteriori donnée par

$$\log p(y|X) = -\frac{1}{2} \log |K + \sigma_n^2 \mathbb{I}| - \frac{1}{2} y'(K + \sigma_n^2 \mathbb{I})^{-1} y - \frac{N}{2} \log 2\pi$$

Haichen Shi ([58]) effectue l'étude de la cointégration non linéaire, via des méthodes d'apprentissage automatique des méthodes de séries chronologiques. L'approche de la cointégration non linéaire adoptée est celle discutée dans Dufrenot et Mignon (2012) ([16]) étendue à la deuxième étape du cadre Engle-Granger en vue d'une régression non linéaire du type

$$y_t = f(x_t) + \epsilon_t$$

pour une fonction de régression f . Alors, la régression par processus gaussien est utilisée dans Shi (2018) ([57]) pour estimer la fonction de cointégration non linéaire $f(\cdot)$ du modèle ci-dessus. Les étapes suivies sont comme suit :

1. Sélectionner les variables appropriées pour le modèle et calculer l'ordre des variables à intégrer. Le test ADF est appliqué aux variables pour déterminer l'ordre d'intégration.
2. Séparer les ensembles de données en deux parties : un ensemble de données d'entraînement et un ensemble de données de test. Les données d'entraînement servent à entraîner le modèle de régression GP ; l'ensemble de données de test sert à surveiller les variations potentielles du système.
3. Utiliser les données d'apprentissage pour entraîner le modèle de régression GP, puis appliquer à nouveau le test ADF pour vérifier si la série résiduelle du modèle est intégrée à

un ordre inférieur à celui des variables d'origine. Une fois cet objectif atteint, on peut dire que la relation de cointégration non linéaire est établie avec succès et que les tendances communes sont éliminées; la série résiduelle du modèle peut donc constituer un bon indicateur des variations induites par les dommages.

4. Comme la régression cointégrante non linéaire peut suivre plusieurs formes de modèles, la dernière étape consiste à étudier toutes les formes de modèles possibles, à les évaluer et à sélectionner le meilleur comme résultat final. Afin de mesurer la « qualité » d'un modèle, on peut utiliser des indicateurs d'aptitude, des critères d'information ou des statistiques de test pour quantifier les performances de différents modèles.

Dans ce qui suit, nous proposons de résumer un cas pratique étudié par Shi (2018) [57]) qu'il a repris lui-même de Cross et Worden ([12]) pour réaliser une cointégration non linéaire. Ce cas est intéressant car il porte sur des séries linéaires à tendance stationnaire confrontées aux séries quadratiques à tendance stationnaire. De quoi s'agit-il exactement ?

Il n'est pas rare de voir des situations où des caractéristiques sensibles aux dommages sont en réalité liées de manière non linéaire à des perturbations externes. La cointégration linéaire conventionnelle ne permet donc pas de modéliser précisément cette relation.

C'est pour reproduire ce comportement que Cross et Worden ([12]) ont simulé deux séries temporelles à tendance stationnaire, l'une avec une tendance linéaire et l'autre avec une tendance quadratique. Cross et Worden proposent alors le problème suivant :

Il est connu que les résidus d'une régression linéaire entre les variables d'une série chronologique non cointégrée sont stationnaires. Supposons que l'on ait deux séries, dont l'une de manière linéaire ou en cassure, la maintienne sous la stationnarité tandis que l'autre série de manière non linéaire (quadratique dans leur simulation) la maintienne également sous la stationnarité. Pour illustrer ce cas, ils simulent deux séries chronologiques, l'une étant une combinaison linéaire d'un bruit blanc et d'une perturbation externe, l'autre étant une combinaison quadratique d'un bruit blanc et d'une perturbation externe. L'une des séries est donc définie comme une perturbation externe, T_i , et l'autre (également stationnaire) est une fonction non linéaire de façon quadratique, en ajoutant un bruit blanc. En notant respectivement, pour le modèle proposé, les séquences de ces variables comme suit :

$$X_i = T_i + \epsilon_i \quad ; \quad Y_i = T_i^2 + \mu_i \quad ; \quad i = 1, \dots, 1000 \quad (3.3.1)$$

avec (ϵ_i) et (μ_i) des bruits blancs, T_i , T_i^2 , des valeurs déterministes.
Soit la série chronologique définie par

$$Z_i = a_1 X_i^2 + a_2 Y_i \quad ; \quad i = 1, \dots, 1000 \quad (3.3.2)$$

La question que se posent les auteurs est : puisqu'il est clair qu'une combinaison linéaire de (X_i) et de (Y_i) ne peut pas être stationnaire, est-il possible de trouver des valeurs de a_1 et de a_2 telles que la série (Z_i) soit stationnaire? Ce qui reprend le test des séries originales (X_i) et (Y_i) cointégrées. Cross et Worden (2011) ([12]) abordent ce problème en utilisant les deux approches citées plus haut.

Le but de notre travail est d'aborder le problème posé par ces auteurs, moyennant des méthodes de simulation intensives avec les données du modèle étudié par Shi (2018)

([57]). C'est à dire, comme suit : on se donne deux séries de 1000 données simulées de (X_i) et (Y_i) et on suppose que la variable déterministe T représentant la température est une fonction strictement monotone croissante de i ou de t . Ce contexte correspond à la procédure d'Engle-Granger qui se récapitule comme suit : si la stationnarité de X_t et Y_t soit inconnue, trois scénarios sont possibles :

- i) les deux sont stationnaires, $I(0)$
- ii) l'un des deux est $I(1)$
- iii) les deux sont $I(1)$.

Pour le premier scénario, les méthodes de modélisation classiques et largement développés dans la littérature, peuvent être appliquées. Pour ce qui est du second scénario, le modèle d'une série $Z_t = X_t + Y_t$ peut aboutir à un résultat non stationnaire et donc aucune cointégration n'existe dans ce cas. Par contre, pour le dernier scénario où on s'intéresse ici, X_t et Y_t contiennent des racines unitaires ou des tendances stochastiques. On peut alors les exprimer sous la forme suivante :

$$X_t = (\text{tendance stochastique de } x) + \text{processus stationnaire}$$

$$Y_t = (\text{tendance stochastique de } y) + \text{processus stationnaire}$$

Dans le cas le plus probable, si les termes de tendance stochastique des X_t et Y_t sont totalement indépendants, $X_t + Y_t$ produira toujours un résidu non stationnaire et si seulement si les tendances stochastiques sont proportionnelles, alors la combinaison Z_t c'est-à-dire si X_t ou Y_t ont une tendance commune,

$$(\text{tendance stochastique de } x) = a \cdot (\text{tendance stochastique de } y)$$

peut alors s'annuler, de sorte que le résidu stationnaire $\epsilon_t = X_t - aY_t$ peut être obtenu. X_t et Y_t sont alors dits cointégrés au sens de la cointégration ordinaire. On peut consulter Shi (2018, section 4.1) ([57]) pour plus de détails sur cette procédure, en particulier sur les inconvénients éventuels.

3.3.2 Étude de Monte Carlo

Nous nous proposons de simuler deux séries de 1000 données simulées de (X_t) et (Y_t) et supposons que la variable déterministe T représente la température sous forme fonction strictement monotone comprise entre 0 et 10 (Shi, 2018) ([57]).

Les données de départ sont donc deux vecteurs x et y à 1000 composantes correspondant aux modèles (3.3.1).

Ces deux séries, comme nous pouvons facilement le constater, ne sont pas stationnaires. Nous obtenons les séries représentées par la figure 3.3 qui représente ces deux séries ainsi que leur résidu. Cette figure illustre bien la non stationarité aussi bien des deux séries que du résidu.

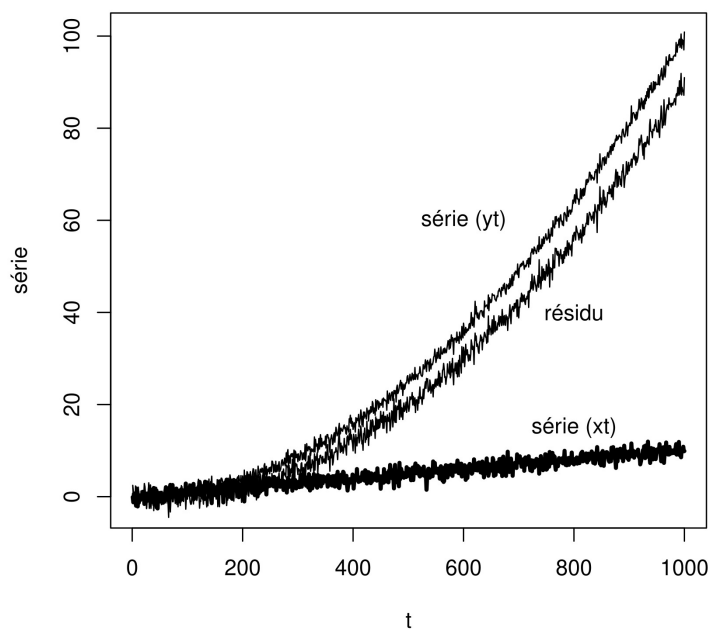


FIGURE 3.3 – Valeurs simulées des deux séries (X_t) et (Y_t) et de leur résidu

Pour chaque valeur choisie du couple (a_1, a_2) ,

1. nous obtenons le vecteur $z = a_1x^2 + a_2y$.
2. Nous calculons la valeur de la statistique de Dickey-Fuller DF correspondante
3. Nous déduisons la p-valeur p du test.

Si $p \geq 0,05$, nous décidons que z n'est pas stationnaire, donc, x et y non cointégrées. Par contre, si $p < 0,05$, nous décidons que z est stationnaire, donc, x et y cointégrées.

Le table 3.1 donne les p-valeurs des premiers résultats de notre simulation où nous avons choisi $a_1 = 0, 0,5, 1,0, 1,5, \dots, 10$ et $a_2 = 0,5, 1,0, 1,5, 2,0, \dots, 5,0$.

TABLE 3.1 – p-valeurs du test **ADF** pour $0 \leq a_1 \leq 10$ et $0.5 \leq a_2 \leq 5.0$

a_2	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0.5	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985
1.0	0.478	0.775	0.909	0.965	0.985	0.985	0.985	0.985	0.985	0.985
1.5	0.191	0.478	0.686	0.905	0.976	0.985	0.985	0.985	0.985	0.985
2.0	0.081	0.304	0.478	0.705	0.860	0.976	0.985	0.985	0.985	0.985
2.5	0.049	0.191	0.353	0.478	0.576	0.721	0.877	0.985	0.985	0.985
3.0	0.035	0.098	0.191	0.304	0.385	0.566	0.721	0.877	0.985	0.985
3.5	0.028	0.081	0.142	0.242	0.338	0.477	0.566	0.721	0.877	0.985
4.0	0.023	0.060	0.098	0.191	0.304	0.385	0.477	0.566	0.721	0.877
4.5	0.019	0.048	0.081	0.142	0.242	0.338	0.385	0.477	0.566	0.721
5.0	0.017	0.036	0.060	0.098	0.191	0.304	0.338	0.385	0.477	0.566
5.5	0.015	0.029	0.048	0.081	0.142	0.242	0.304	0.338	0.385	0.477
6.0	0.013	0.024	0.036	0.060	0.098	0.191	0.242	0.304	0.338	0.385
6.5	0.012	0.021	0.029	0.048	0.081	0.142	0.191	0.242	0.304	0.338
7.0	0.011	0.019	0.024	0.036	0.060	0.098	0.142	0.191	0.242	0.304
7.5	0.011	0.017	0.021	0.029	0.048	0.081	0.098	0.142	0.191	0.242
8.0	0.010	0.015	0.019	0.024	0.036	0.060	0.081	0.098	0.142	0.191
8.5	0.010	0.014	0.017	0.021	0.029	0.048	0.060	0.081	0.098	0.142
9.0	0.010	0.013	0.015	0.019	0.024	0.036	0.048	0.060	0.081	0.098
9.5	0.010	0.012	0.014	0.017	0.021	0.029	0.036	0.048	0.060	0.081
10.0	0.010	0.018	0.026	0.053	0.093	0.129	0.151	0.171	0.191	0.211

Nous remarquons qu'au fur et à mesure que le paramètre a_1 augmente et si les valeurs de a_2 sont petites, il est possible de rendre stationnaire la série z car la p-valeur est inférieure à 5%. Par exemple, si $a_1 = 10$ et $0 < a_2 < 2$, les deux séries sont non linéairement cointégrées.

Nous nous sommes posé la question de savoir si, en augmentant encore plus les valeurs de a_2 , est ce que la p-valeur peut descendre au-dessous de 5% pour des valeurs plus grandes de a_2 .

Le table 3.2 ci-dessous nous donne les éléments de réponse.

Chapitre 3 : Simulations et applications

TABLE 3.2 – p-valeurs du test **ADF** pour $10 \leq a_1 \leq 100$ et $0.5 \leq a_2 \leq 5.0$

a_2	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
10	0.010	0.017	0.024	0.037	0.049	0.068	0.087	0.111	0.152	0.191
15	0.010	0.012	0.017	0.022	0.028	0.037	0.045	0.055	0.068	0.081
20	0.010	0.013	0.017	0.021	0.024	0.030	0.037	0.043	0.049	0.059
25	0.010	0.011	0.014	0.017	0.020	0.022	0.027	0.030	0.032	0.037
30	0.010	0.010	0.012	0.015	0.017	0.020	0.022	0.024	0.028	0.035
35	0.010	0.010	0.011	0.013	0.015	0.017	0.019	0.022	0.021	0.023
40	0.010	0.010	0.010	0.011	0.013	0.015	0.017	0.019	0.019	0.021
45	0.010	0.010	0.010	0.010	0.011	0.013	0.015	0.017	0.019	0.021
50	0.010	0.010	0.010	0.010	0.010	0.013	0.014	0.014	0.017	0.019
55	0.010	0.010	0.010	0.010	0.010	0.011	0.012	0.013	0.014	0.017
60	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.012	0.013	0.015
65	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.012	0.014
70	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.012	0.013
75	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.012
80	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.011
85	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011
90	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
95	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
100	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010

Ainsi, nous voyons bien que si a_1 est suffisamment grand, nous pouvons obtenir la cointégration des deux séries pour des valeurs même plus grandes de a_2 . Donc, plus a_1 est fixé à une valeur assez grande, plus des valeurs de a_2 sont disponibles pour rendre cointégrées les deux séries. Cela est parfaitement illustré dans la figure 3.4 qui montre la variation de la p-valeur en fonction de a_2 pour trois valeurs de $a_1 = 5, 10, 20$.

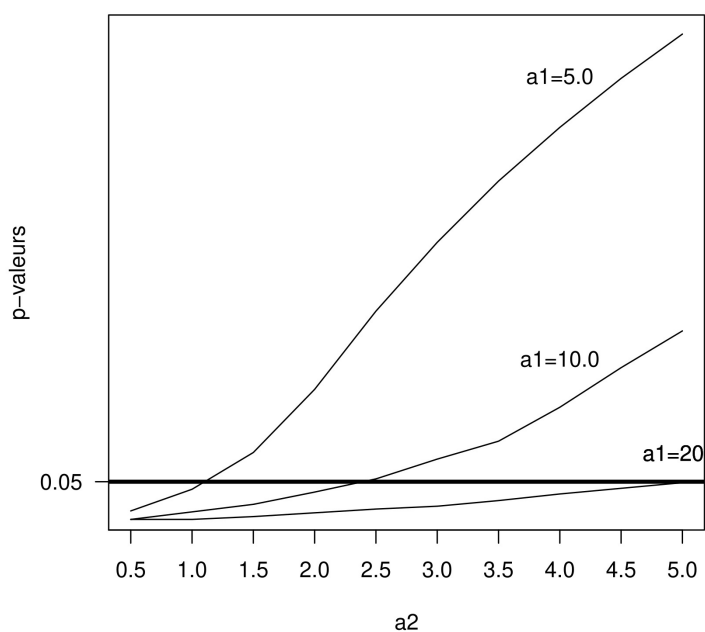


FIGURE 3.4 – Variation de la p-valeur en fonction de a_2 pour $a_1 = 5, 10, 20$

De même, si on fixe la valeur de a_2 à une valeur donnée, la cointégration des deux séries n'est obtenue que si a_2 est petit et a_1 dépassant un certain seuil minimal. Plus a_2 augmente, plus la cointégration des deux séries ne sera possible que de grandes valeurs de a_1 . La figure 3.5 suivante illustre bien cette situation pour $a_2 = 1.0, 2.0, 3.0$.

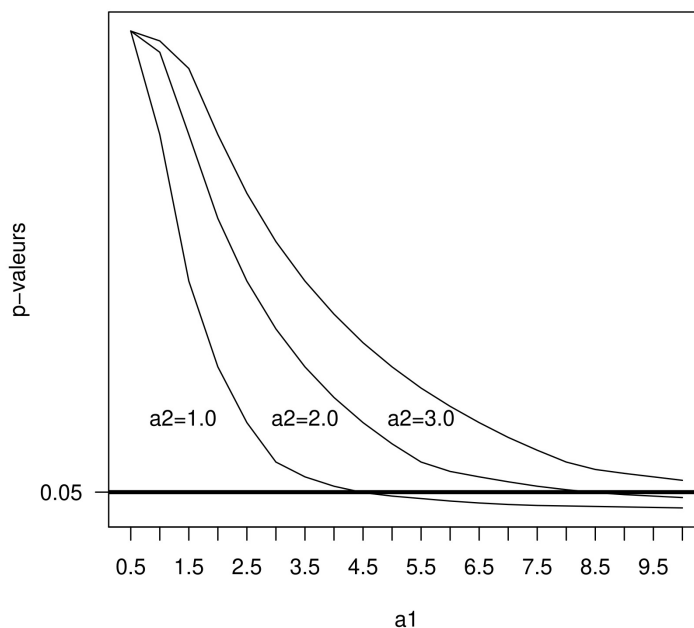


FIGURE 3.5 – Variation de la p-valeur en fonction de a_1 pour $a_2 = 1, 2, 3$

3.3.3 La démarche par processus gaussien

Shi (2018) ([58]) a étudié le problème que nous avons étudié ci-dessus en introduisant ce qu'il appelle la procédure dite " par processus Gaussien" qu'il justifie comme suit : une des applications pertinentes de la cointégration non linéaire porte sur son utilisation dans le Machine Learning.

Shi propose d'étudier deux cas pratiques de cointégration non linéaire utilisant ce qu'on appelle régression par processus gaussien (GP) qui est une approche d'apprentissage automatique bayésienne pouvant traiter les problèmes de régression non linéaire. En clair, on peut dire qu'un GP peut être considéré comme une loi de probabilité sur des fonctions, dont tous les échantillons finis sont distribués conjointement selon une distribution gaussienne. Rasmussen et Williams (2006) ([55]) développent largement l'usage de cet outil dans le machine learning. En effet, comme nous l'avons souligné dans l'introduction, il existe une théorie riche pour les modèles de processus gaussiens dans la littérature sur l'analyse des séries chronologiques. Dans sa thèse, Shi a utilisé la procédure par processus GP pour capter la non linéarité comme suit.

Pour ce faire, l'auteur propose un autre modèle que celui de Cross et Worden (2011) ([12]) de la forme

$$X_t = \mathcal{GP}(Y_t)$$

Avec \mathcal{GP} qui correspond au processus gaussien décrit plus haut.

L'objectif de cette étude est de voir s'il est possible de produire un résidu stationnaire des deux séries même si elles ne sont pas stationnaires comme nous l'avons déjà montré. L'idée est donc de pouvoir capter la non linéarité tout en supprimant les nuisances liées à l'hétéroscédasticité. Pour résumer sa démarche, il donne dans la figure 3.6 la prédiction obtenue avec la démarche GP où on voit bien que la tendance de X_t est bien capturée par le modèle proposé.

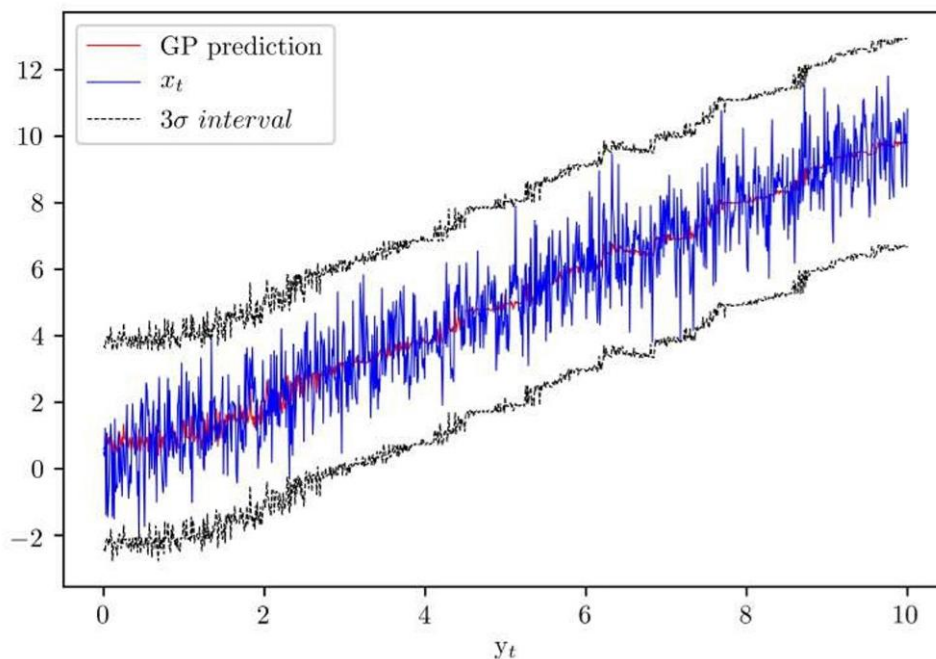


FIGURE 3.6 – Prédiction GP du modèle de cointégration non linéaire (source Shi, 2018)

Pour affirmer que la cointégration non linéaire entre X_t et Y_t est obtenue, il réalise un test ADF sur le résidu obtenu qui le confirme. La figure 3.7 montre bien les points en bleu des données initiales et les points en rouge correspondent aux valeurs prédites par cette procédure. Cette figure montre comment la non linéarité a été prise en compte par la démarche GP mettant en évidence la cointégration non linéaire entre les deux séries.

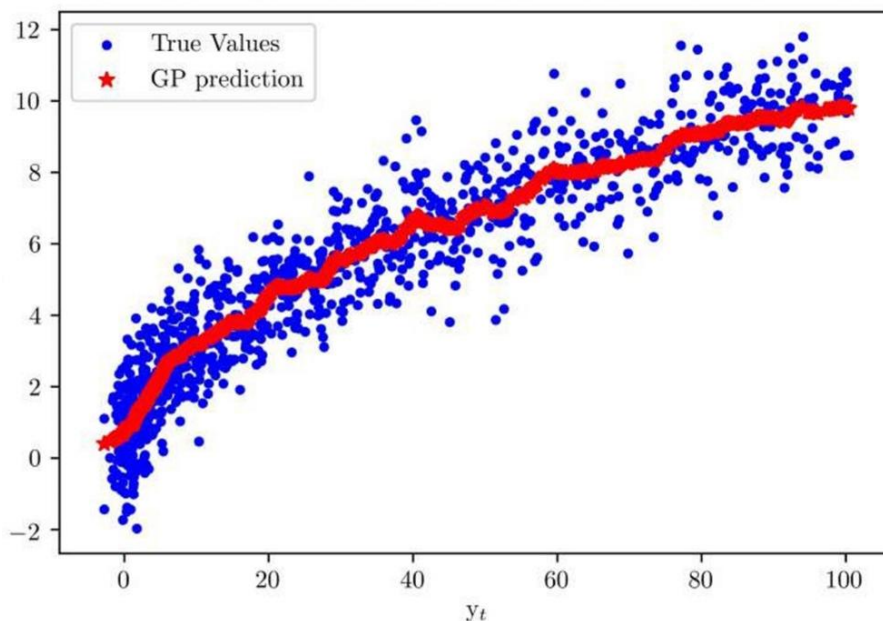


FIGURE 3.7 – Prédiction GP du modèle de cointégration non linéaire (source Shi, 2018)

3.3.4 Application au pont de référence Z24

Le but de cette partie est de souligner l'importance de l'application de la cointégration non linéaire dans d'autres domaines, notamment à travers les techniques d'évaluation des propriétés dynamiques des ponts à partir de données expérimentales. Le pont Z24, un pont en béton armé à trois travées situé en Suisse, a été choisi comme étude de cas dans plusieurs articles. Peeters et Ventura (2003) ([46]) ont comparé, sur ce cas, les techniques d'analyse modale habituellement employées dans les laboratoires et les résultats des techniques d'excitation classiques pour les grandes structures de génie civil. Plus d'information sur ce pont peut être trouvée dans KrÄamer et al (1999) ([37]). Steenackers et Guillaume (2004) ([60]) ont utilisé la régression linéaire et non linéaire pour mettre en relation les paramètres du pont comme les températures du sol, du pont et de l'air.

Quel est le problème posé exactement avec ce pont? Le pont Z24 a été construit en 1963 pour relier Koppigen et Utzenstorf, en Suisse, en surmontant l'autoroute A1 Berne-Zurich. Sa structure classique, un pont à poutres-caissons en béton post-tendu, était composée de trois travées. Ce pont devait être démoli à la fin de 1998 pour permettre la construction d'une nouvelle voie ferrée adjacente à l'autoroute, bien que dans un état relativement bon. Afin d'identifier ses caractéristiques dynamiques, le pont Z24 a été lourdement équipé d'accéléromètres et de capteurs environnementaux (température, humidité, vitesse et direction du vent).

Il a été remarqué par les expérimentateurs que les séries chronologiques liées à l'évolution des paramètres naturels (environnement, etc.) du pont ne sont pas stationnaires et la relation entre la température et les autres caractéristiques n'est pas linéaire.

C'est pourquoi la plupart des méthodes d'apprentissage supervisé appliquées à ce pont

visaient souvent à trouver la relation de régression entre les propriétés dynamiques et la température. Par exemple, Cross et Worden (2011) [12]), cités plus haut ont appliqué des méthodes de régression sur les données de ce pont. C'est le problème de cette relation avec les températures qu'a abordé Shi (2018) ([57]) dans sa thèse par le biais de la cointégration non linéaire. Il a aussi été mis en évidence que les propriétés dynamiques du système peuvent être fortement influencées par la température, et que cette influence peut être trompeuse pour la détection des dommages. Cet auteur a pu montrer que la méthode de cointégration non linéaire peut permettre de construire un résidu stationnaire, insensible aux variations de température, mais néanmoins sensible aux dommages.

3.4 Arbitrage statistique et trading financier

Les stratégies d'investissement utilisant l'arbitrage statistique sont historiquement nées aux États-Unis dans la banque Morgan Stanley durant les années 80. Ces méthodes de trading se sont alors grandement popularisées avec l'avènement de l'informatisation des marchés financiers ayant entraîné la création d'une masse de données toujours plus importante à analyser. Avec l'essor des big data dans le secteur financier, l'arbitrage statistique est devenu une stratégie d'investissement incontournable. C'est pour cela que, de nos jours, la plupart des sociétés de gestion élaborent des méthodes d'arbitrage statistique.

La définition consensuelle est que l'arbitrage statistique est un ensemble de méthodes statistiques visant à analyser et à exploiter les mouvements de milliers d'actifs traités sur les marchés. Cela se traduit par une analyse des mouvements historiques de ces actifs aussi bien que leurs corrélations. Cependant, il y a lieu de préciser qu'il existe différents types d'arbitrage et que celui-ci ne doit pas être confondu avec le trading à haute fréquence. La raison est que l'arbitrage cherche seulement à exploiter des bâtir un raisonnement sur une analyse statistique de moyen long terme. C'est là que la cointégration en est un outil précieux qui justifie pourquoi nous l'abordons dans dans ce chapitre en plus du fait que tous les acteurs de la finance de marché utilisent des modèles d'arbitrage pour construire leurs positions ou étudier empiriquement les marchés. Et la tendance de l'utilisation de ses modèles est loin de s'inverser.

Pour mieux comprendre l'importance de la cointégration dans ce domaine, il suffit de définir l'arbitrage à la base comme l'achat et la vente simultanés d'actifs similaires sur différents marchés afin de tirer parti des différentiels de prix. Par exemple, en pratique, Lorsqu'un opérateur utilise l'arbitrage, il achète essentiellement un actif moins cher et le vend à un prix plus élevé sur un marché différent, réalisant ainsi un profit sans aucun flux de trésorerie net. Techniquement, l'arbitrage statistique est dérivé d'un ensemble de stratégies d'investissement algorithmiques quantitatives dont le but ultime est souvent de générer des profits de trading supérieurs à la normale pour les gros investisseurs. Dans l'arbitrage statistique, un trader ouvrira simultanément une position longue et une position courte afin de tirer parti de la tarification inefficace des actifs corrélés. Par exemple, si un opérateur estime qu'Amazon est surévalué et que Facebook est sous-évalué, il ouvrira une position longue sur Amazon et, en même temps, une position courte sur Facebook. Ceci est souvent appelé le trading par paires qui justifie l'usage de la cointégration en pratique. En effet, le test le plus courant pour le trading par paires est le test de cointégration. Dans ce qui suit, nous présentons un exemple de simulation d'une situation financière faisant appel à cette technique.

3.4.1 Simulation d'un fonds de couverture avec Modèle AR(1)

Supposons qu'un fonds de couverture détienne une position de 1 000 000 \$ dans un portefeuille, en utilisant 50 000 \$ de son propre capital et 950 000 \$ d'argent emprunté pour l'achat. Si la valeur du portefeuille descend en dessous de 950 000 \$ à la fin d'une journée de trading, alors le fonds de couverture doit liquider et rembourser le prêt.

Le portefeuille a été sélectionné par analyse de cointégration et son prix suit un processus AR(1) :

$$(P_t - \mu) = \phi(P_{t-1} - \mu) + \varepsilon_t \quad (3.4.1)$$

où

- P_t est le prix du portefeuille à la fin du jour t
- $\mu = \$1\,030\,000$ (niveau moyen)
- $\phi = 0.99$ (paramètre de persistance)
- ε_t a un écart-type de \$5 000

Le fonds de couverture sait que le prix reviendra éventuellement à \$1 030 000 (en supposant le modèle correct, ce qui est une forte hypothèse).

Il décide de liquider sa position le jour t si $P_t \geq 1\,020\,000$ \$, réalisant ainsi un profit d'au moins 20 000 \$. Cependant, si le prix tombe sous 950 000 \$, il doit liquider avec une perte de 50 000 \$ plus la différence entre 950 000 \$ et le prix de liquidation.

En résumé, le fonds de couverture liquidera à la fin du premier jour où :

- $P_t \geq 1\,020\,000$ \$ (profit $\geq 20\,000$ \$), **ou**
- $P_t < 950\,000$ \$ (perte $\geq 50\,000$ \$)

Cette configuration soulève quatre interrogations fondamentales :

1. Quel est le gain moyen anticipé par cette stratégie ?
2. Quelle est la probabilité d'atteindre le seuil de liquidation à perte ?
3. Combien de jours en moyenne la position est-elle maintenue ?
4. Quel rendement peut-on espérer sur le capital investi ?

Pour répondre à ces questions, nous avons implémenté une simulation Monte Carlo avec un processus AR(1).

Cette simulation illustre bien le concept d'arbitrage statistique, plus précisément une stratégie de mean-reversion trading (retour à la moyenne).

Le portefeuille est construit via une analyse de cointégration (comme mentionné), ce qui est la base classique des stratégies d'arbitrage statistique. La cointégration identifie des actifs dont les prix divergent temporairement mais reviennent à une relation d'équilibre à long terme.

Chapitre 3 : Simulations et applications

Le code R suivant présente l'implémentation :

```
> set.seed(2009)
> niter <- 10000
> mu <- 1030000
> phi <- 0.99
> sigma <- 5000
> P0 <- 1000000
> seuil_haut <- 1020000
> seuil_bas <- 950000
>
> results <- replicate(niter, {
+   P <- P0
+   day_exited <- 1000 # valeur par défaut si pas de sortie anticipée
+   for(day in 1:1000) {
+     P <- mu + phi * (P - mu) + rnorm(1, 0, sigma)
+     if(P >= seuil_haut || P < seuil_bas) {
+       day_exited <- day
+       break # sortir de la boucle for
+     }
+   }
+   c(P - 1e6, day_exited) # retourner le profit et le jour de sortie
+ })
>
> # Résultats
> cat("Profit moyen:", mean(profits), "$\n")
Le profit moyen est : 15651.19 $

> cat("Probabilité de perte:", mean(profits < -50000) * 100, "%\n")
La probabilité de perte est: 9.66 %

> cat("Temps moyen avant liquidation:", mean(jours), "jours\n")
Le temps moyen avant liquidation est: 44.1267 jours

> cat("Rendement annuel:", (mean(profits) / 50000) * (253 / mean(jours))
* 100, "%\n")
Le rendement annuel est : 179.4719 %
```

Les résultats de la simulation montrent que, sur 10,000 itérations :

- Le profit moyen par position est de 15,651 \$, ce qui indique une performance globale positive de la stratégie malgré les risques de perte.
- La probabilité de perte importante (i.e., forcer la liquidation sous le seuil de 950,000 \$) est estimée à 9.66%, ce qui reste relativement faible mais non négligeable.
- Le temps moyen de détention du portefeuille est d'environ 44 jours, ce qui correspond à une stratégie de moyen terme, adaptée à un horizon d'investissement tactique.

- Enfin, le rendement annualisé sur le capital investi est particulièrement élevé, atteignant 179.47%, ce qui s'explique par l'effet de levier important (ratio 20 :1) utilisé par le fonds.

Conclusion

Cette simulation met en évidence le potentiel d'une stratégie de retour à la moyenne basée sur un portefeuille construit par cointégration, couplée à un effet de levier élevé. Le processus AR(1) permet de modéliser de manière simple, mais efficace, la dynamique des prix autour de leur moyenne historique.

Malgré une probabilité non nulle de perte significative (près de 10%), le rendement espéré est substantiel, démontrant l'intérêt économique de cette approche dans un contexte de gestion alternative. Cependant, plusieurs hypothèses restent fortes, notamment la stationnarité du processus et la stabilité des paramètres (μ , ϕ , σ).

De plus, la simulation ne tient pas compte de facteurs de marché externes (liquidité, volatilité structurelle, frais de transaction, etc.) ni de contraintes réglementaires, qui pourraient affecter la robustesse de la stratégie en conditions réelles.

Enfin, cette analyse constitue une base quantitative utile pour l'évaluation de stratégies d'arbitrage statistique, et peut être enrichie par des extensions comme l'optimisation de seuils ou l'introduction de stop-loss dynamiques.

Conclusion générale

Ce mémoire a parcouru méthodiquement les trois dimensions fondamentales de l'analyse des séries temporelles modernes. La première partie a établi les bases théoriques indispensables, clarifiant les concepts de stationnarité et les subtilités des différents tests de racine unitaire. La seconde partie a approfondi l'évolution conceptuelle de la cointégration, depuis ses formulations linéaires classiques jusqu'aux développements non linéaires contemporains, offrant ainsi un panorama complet des outils disponibles pour analyser les relations entre variables non stationnaires.

Les applications pratiques présentées dans la troisième partie ont démontré la pertinence opérationnelle de ces méthodes. L'étude de la relation livre sterling/euro a illustré la puissance de la cointégration linéaire en macroéconomie financière. L'analyse du pont Z24 a révélé l'apport décisif des approches non linéaires en ingénierie structurelle, tandis que la simulation financière a montré leur utilité pour les stratégies de marché. Ces cas concrets soulignent combien le choix approprié entre méthodes linéaires et non linéaires dépend crucialement des caractéristiques propres à chaque problème étudié.

Les principaux enseignements de cette recherche mettent en lumière plusieurs points clés. D'abord, la nécessité d'un diagnostic précis des propriétés des séries avant toute modélisation. Ensuite, l'importance croissante des approches non linéaires pour capturer la complexité des phénomènes économiques réels. Enfin, la complémentarité entre rigueur théorique et validation empirique dans le développement de modèles pertinents.

Les perspectives de recherche ouvertes par ce travail sont multiples. L'intégration des méthodes d'apprentissage automatique pourrait permettre de détecter des formes plus complexes de cointégration. L'adaptation de ces techniques aux données haute fréquence représente un autre chantier important. Enfin, le développement de tests plus robustes aux différentes formes de non-linéarité constituerait une avancée méthodologique significative.

Au terme de ce parcours, il apparaît clairement que l'analyse des séries temporelles, en particulier à travers le prisme de la cointégration linéaire et non linéaire, reste un champ extrêmement vivant de l'économétrie. Ses applications pratiques ne cessent de s'étendre, tandis que ses fondements théoriques continuent à se raffiner. Ce mémoire aura atteint son objectif s'il contribue à montrer comment ces outils sophistiqués peuvent éclairer des problèmes concrets tout en ouvrant de nouvelles voies pour la recherche future.

Bibliographie

- [1] AIT RAMDANE, T. « Modèles de Régression Appliqués au Machine Learning ». Master's thesis. Université Mouloud Mammeri de Tizi-Ouzou, 2024.
- [2] APARICIO, F., ESCRIBANO, A. et SIPOLS, A. E. « Range Unit-Root (RUR) Tests : Robust against Nonlinearities, Error Distributions, Structural Breaks and Outliers ». In : *Journal of Time Series Analysis* 27.4 (2006), p. 477-496.
- [3] ASTERIOU, D., MASATCI, K. et PILBEAM, K. *Exchange Rate Volatility and International Trade : International Evidence from the MINT Countries*. Vol. 58, pp. 133–140. Economic Modelling, 2016.
- [4] ATIL, L., FELLAG, H. et SIPOLS, A. « Non-linear Cointegration Test, Based on Record Counting Statistic ». In : *Computational Economics* 64 (2024), p. 2205-2230. DOI : 10.1007/s10614-024-10608-2.
- [5] BANDI, F. M. *On persistence and nonparametric estimation (with an application to stock return predictability)*. Manuscrit non publié. 2004.
- [6] BREITUNG, J. « Rank Tests for Nonlinear Cointegration ». In : *Journal of Business & Economic Statistics* 19.3 (2001), p. 331-340.
- [7] CHANG, Y., PARK, J. et PHILLIPS, P. *Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors*. Rapp. tech. 1495. Cowles Foundation, 1999.
- [8] CHANG, Y., PARK, J. Y. et PHILLIPS, P. C. B. « Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors ». In : *Econometrics Journal* 4.1 (2001), p. 1-36.
- [9] CHOI, I. et SAIKKONEN, P. « Tests for Nonlinear Cointegration ». In : *Econometric Theory* 26.3 (2010), p. 682-709. DOI : 10.1017/S0266466609100043.
- [10] COLLETAZ. *Introduction à la cointégration*. 2020.
- [11] COWPERTWAIT, P. S. P. et METCALFE, A. V. *Introductory Time Series with R*. Springer, 2009.
- [12] CROSS, E. J. et WORDEN, K. « Approaches to Nonlinear Cointegration with a View Towards Applications in SHM ». In : *Journal of Physics : Conference Series*. T. 305. 2011, p. 012069.
- [13] DICKEY, D. et FULLER, W. *Distribution of the Estimators for Autoregressive Time Series With a Unit Root*. Vol. 74, pp. 427–431. Journal of the American Statistical Association, 1979.
- [14] DICKEY, D. A. et FULLER, W. A. « Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root ». In : *Econometrica* 49.4 (1981), p. 1057-1072.
- [15] DITTMANN, I. et GRANGER, C. W. J. « Properties of Nonlinear Transformations of Fractionally Integrated Processes ». In : *Journal of Econometrics* 110.2 (2002), p. 113-133.
- [16] DUFRÉNOT, G. et MIGNON, V. *Recent Developments in Nonlinear Cointegration with Applications to Macroeconomics and Finance*. Springer Science & Business Media, 2012.
- [17] ELLIOTT, G., T, R. et STOCK, J. *Efficient Tests For an Autoregressive Unit Root*. Vol. 64, No. 4, pp. 813–836. Econometrica, 1996.
- [18] ENGLE, R. et GRANGER, C. *Co-Integration and Error Correction : Representation, Estimation, and Testing*. Vol. 55, No. 2, pp. 251–276. Econometrica, 1987.
- [19] ESCRIBANO, A. et MIRA, S. « Nonlinear error correction models ». In : *Journal of Time Series Analysis* 23.5 (2002), p. 509-522. DOI : 10.1111/1467-9892.00271.

Chapitre 3 : Simulations et applications

- [20] ESCRIBANO, A., SANTOS, M. T. et SIPOLS, A. E. « Testing for cointegration using induced order statistics ». In : *Computational Statistics* 23 (2008), p. 131-151.
- [21] ESCRIBANO, A., SIPOLS, A. E. et APARICIO, F. « Nonlinear cointegration and nonlinear error correction : Applications of tests based on first differences of ranges ». In : *Communications in Statistics - Simulation and Computation* 35.5 (2006), p. 939-956. DOI : 10.1080/03610910600880351. URL : <https://doi.org/10.1080/03610910600880351>.
- [22] FORTIER, S. *Les modeles AR, MA et ARMA multidimensionnels : estimation et causalité*. Département de mathématique. Université de SHERBROOK, 2020.
- [23] FULLER, W. *Introduction to Statistical Time Series*. New York : Wiley, 1976.
- [24] GOURIEROUX, C. et MONFORT, A. *Time Series and Dynamic Models*. Cambridge : Cambridge University Press, 1997.
- [25] GRANGER, C. *Some Properties of Time Series Data and Their Use in Econometric Model Specification*. Vol. 16, pp. 121–130. *Journal of Econometrics*, 1981.
- [26] GRANGER, C. W. J. « Testing for Causality : A Personal Viewpoint ». In : *Journal of Economic Dynamics and Control* 2 (1980), p. 329-352.
- [27] GRANGER, C. W. J. et HALLMAN, J. « Nonlinear Transformations of Integrated Time Series ». In : *Journal of Time Series Analysis* 12.3 (1991), p. 207-224.
- [28] GRANGER, C. W. et HALLMAN, J. « Long memory series with attractors ». In : *Oxford Bulletin of Economics and Statistics* 53.1 (1991), p. 11-26.
- [29] JALIL, A., MAHMOOD, T. et IDREES, M. *Tourism-Growth Nexus in Pakistan : Evidence from ARDL Bounds Test*. Vol. 35, pp. 185–191. DOI : 10.1016/j.econmod.2013.06.034. *Economic Modelling*, 2013.
- [30] JOHANSEN, S. *Statistical Analysis of Cointegration Vectors*. Received September 1987, final version January 1988. *Journal of Economic Dynamics et Control*, 1988.
- [31] JOHANSEN, S. *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models*. T. 59. *Econometrica*, 1991, p. 1551-1580.
- [32] JOHANSEN, S. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, 1995.
- [33] JOHANSEN, S. et JUSELIUS, K. « Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money ». In : *Oxford Bulletin of Economics and Statistics* 52.2 (1990), p. 169-210.
- [34] KARLSEN, H. A., MYKLEBUST, T. et TJØSTHEIM, D. « Nonparametric Estimation in a Nonlinear Cointegration Type Model ». In : *Annals of Statistics* 35.1 (2007), p. 252-299.
- [35] KARLSEN, H. A. et TJØSTHEIM, D. « Nonparametric estimation in null recurrent time series ». In : *The Annals of Statistics* 29.2 (2001), p. 372-416.
- [36] KIM, H. et BAEK, J. « Assessing Dynamics of Crude Oil Import Demand in Korea ». In : *Economic Modelling* 35 (2013), p. 260-263. DOI : 10.1016/j.econmod.2013.07.010.
- [37] KRÄMER, C., DE SMET, C. et DE ROECK, G. « Z24 Bridge Damage Detection Tests ». In : *Proceedings of the 17th International Modal Analysis Conference*. Orlando, FL, 1999, p. 1023-1029.
- [38] KWIATKOWSKI, D. et al. *Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. How Sure Are We that Economic Time Series Have a Unit Root ?* Vol. 54, pp. 159–178. *Journal of Econometrics*, 1992.
- [39] LIN, J. et GRANGER, C. « Testing Nonlinear Cointegration ». In : *COMPSTAT 2004 Proceedings*. 2004, p. 1413-1419.
- [40] MADDALA, G. S. et KIM, I.-M. *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press, 1998.
- [41] MEYN, S. P. et TWEEDIE, R. L. « Stability of Markovian processes III : Foster-Lyapunov criteria for continuous-time processes ». In : *Advances in Applied Probability* 25.3 (1993), p. 518-548. DOI : 10.2307/1427494.

- [42] MUSCATELLI, V. et SPINELLI, F. « The Long-Run Stability of the Demand for Money : Italy 1861–1996 ». In : *Economic Notes* 29.1 (2000), p. 72-76.
- [43] NICOLAU, F. *Séries Chronologiques*. Cours, Département STID, Université Nice d'Azur. Université Nice d'Azur, 2005.
- [44] PARK, J. et PHILLIPS, P. « Asymptotics for Nonlinear Transformations of Integrated Time Series ». In : *Econometric Theory* 15.3 (1999), p. 269-298.
- [45] PARK, J. et PHILLIPS, P. « Nonlinear Regressions with Integrated Time Series ». In : *Econometrica* 69.1 (2001), p. 117-161.
- [46] PEETERS, B. et VENTURA, C. « Comparative Study of Modal Analysis Techniques for Bridge Structures ». In : *Journal of Structural Engineering* 129.8 (2003). Z-24 Bridge case study, p. 965-972.
- [47] PESARAN, M. et PESARAN, B. *Working with Microfit 4.0 : Interactive Econometric Analysis*. Oxford University Press, 1997.
- [48] PESARAN, M. et SHIN, Y. « An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis ». In : *The Ragnar Frisch Centennial Symposium*. 1998, p. 371-413.
- [49] PESARAN, M., SHIN, Y. et SMITH, R. « Bounds Testing Approaches to the Analysis of Level Relationships ». In : *Journal of Applied Econometrics* 16.3 (2001), p. 289-326.
- [50] PHILLIPS, P. et HANSEN, B. « Statistical Inference in Instrumental Variables Regressions with I(1) Processes ». In : *Review of Economic Studies* 57 (1990), p. 99-125.
- [51] PHILLIPS, P. et OULIARIS, S. « Asymptotic Properties of Residual Based Tests for Cointegration ». In : *Econometrica* 58.1 (1990), p. 165-193.
- [52] PHILLIPS, P. et PERRON, P. *Testing for a Unit Root in a Time Series Regression*. *Biometrika* 75, 335-346., 1988.
- [53] PHILLIPS, P. (*Time Series Regression With a Unit Root*. *Econometrica* 55, 277-301., 1987.
- [54] PHILLIPS, P. C. et PARK, J. Y. *Nonstationary density estimation and kernel autoregression*. Manuscrit non publié. 1998.
- [55] RASMUSSEN, C. E. et WILLIAMS, C. K. I. *Gaussian Processes for Machine Learning*. Cambridge, MA : MIT Press, 2006.
- [56] RUSHDI, M., KIM, J. et SILVAPULLE, P. « ARDL Bounds Tests and Robust Inference for the Long Run Relationship between Real Stock Returns and Inflation in Australia ». In : *Economic Modelling* 29 (2012), p. 535-543.
- [57] SHI, H. « On Nonlinear Cointegration Methods for Structural Health Monitoring ». Thèse de doct. Sheffield, UK : University of Sheffield, mars 2018.
- [58] SHI, H., WORDEN, K. et CROSS, E. J. « A Regime-Switching Cointegration Approach for Removing Environmental and Operational Variations in Structural Health Monitoring ». In : *Mechanical Systems and Signal Processing* 103 (2018), p. 381-397.
- [59] SHUMWAY, R. et STOFFER, D. « Time series Analysis and its Applications with R ». In : *Springer* ().
- [60] STEENACKERS, G. et GUILLAUME, P. « Structural Health Monitoring of the Z-24 Bridge in Presence of Environmental Changes Using Modal Analysis ». In : *Applied Mechanics, Vrije Universiteit Brussel* (2004). Conference paper.
- [61] STORN, R. et PRICE, K. « Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces ». In : *Journal of Global Optimization* 11.4 (1997), p. 341-359.
- [62] TODA, H. Y. et YAMAMOTO, T. « Statistical Inference in Vector Autoregressions with Possibly Integrated Processes ». In : *Journal of Econometrics* 66.1-2 (1995), p. 225-250.
- [63] VAN GESTEL, T. et al. « A Bayesian Nonlinear Support Vector Machine Error Correction Model ». In : *Journal of Forecasting* 25.2 (2006), p. 77-100.
- [64] WANG, Q. et PHILLIPS, P. C. B. « Structural Nonparametric Cointegrating Regression ». In : *Econometrica* 77.6 (2009), p. 1901-1948.

Abstract

Ce document présente la cointégration en séries chronologiques, une méthode permettant de détecter des relations d'équilibre à long terme entre plusieurs séries non stationnaires. Dans le chapitre sur les notions générales, il est expliqué que l'objectif principal d'une série chronologique est de modéliser des phénomènes temporels pour réaliser des prévisions, ce qui nécessite souvent de rendre la série stationnaire à l'aide de différenciations successives. Il est donc essentiel de tester la stationnarité, à l'aide de tests comme Dickey-Fuller, Phillips-Perron ou KPSS, car la non-stationnarité peut fausser les modèles. La distinction entre stationnarité stricte et faible, ainsi qu'entre non-stationnarité déterministe et stochastique, permet de choisir les méthodes d'analyse adaptées. L'auteur rappelle que la différenciation peut entraîner une perte d'information, et que la cointégration permet de préserver cette information en associant plusieurs séries non stationnaires. Cette approche n'est pas limitée à l'économie : elle s'applique aussi en ingénierie, en environnement ou en sciences sociales, pour l'analyse de données temporelles variées.

This document presents cointegration in time series, a method to detect long-term equilibrium relationships between several non-stationary series. In the chapter on general concepts, it is explained that the main goal of a time series is to model temporal phenomena for forecasting, which often requires making the series stationary through successive differencing. It is therefore essential to test stationarity using tests such as Dickey-Fuller, Phillips-Perron, or KPSS, since non-stationarity can distort models. The distinction between strict and weak stationarity, as well as between deterministic and stochastic non-stationarity, allows for choosing appropriate analysis methods. The author notes that differencing can lead to information loss, and cointegration helps preserve this information by associating multiple non-stationary series. This approach is not limited to economics : it also applies in engineering, environmental studies, and social sciences for analyzing various temporal data.

Mots-clés :Cointégration, Séries chronologiques, Stationnarité, Test de racine unitaire, Différenciation,Test de cointégration, Modèle à correction d'erreur, Relations d'équilibre , Prévision.