

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 :

Inférence dans les modèles ARDL et application

Réalisé par :

LKOUCHA Hanane ép.AMENOUCHE

Encadré par :

M^{me} ATIL Lynda

Soutenu devant le jury d'examen composé de :

M^{me} Belkacem Cherifa,

MCB, UMMTO

Présidente

M^r Fellag Hocine,

Professeur, UMMTO

Examineur

M^{me} Atil Lynda,

MCA, UMMTO

Rapporteur

Soutenu le : 03/07/2025

Dédicaces

Je dédie ce modeste travail :

À mes chers parents, piliers inébranlables de ma vie, dont l'amour inconditionnel, les sacrifices silencieux et le soutien indéfectible ont guidé chacun de mes pas. Votre sagesse et votre dévouement ont fait de moi la personne que je suis aujourd'hui. Recevez ici l'expression de ma gratitude éternelle et de mon affection sans limites.

À mon frère bien-aimé Mouhand, compagnon de route précieux, merci pour ta présence réconfortante, tes conseils avisés et ta confiance inaltérable. Tu es bien plus qu'un frère - un ami, un soutien, une source d'inspiration quotidienne.

À ma sœur chérie Soumeya, rayon de soleil dans ma vie, dont la douceur, l'énergie positive et le soutien constant ont illuminé mon parcours. Je te souhaite un avenir à la hauteur de ta générosité et de ta bienveillance.

À mon cher époux, mon roc et mon havre de paix, dont la présence apaisante, les encouragements judicieux et l'amour sans faille ont été le moteur de ma persévérance. Ce travail est autant le tien que le mien, fruit de nos sacrifices partagés.

À ma belle-famille, pour votre accueil chaleureux, votre soutien discret mais précieux, et les liens de cœur que nous avons su tisser. Votre bienveillance a été une source de réconfort tout au long de ce parcours.

À mes chers camarades de promotion Master 2 Probabilités et Statistique (Ryma, soumeya, Amel.L, Lisa, Lila, Rachida, Thanina, Amel.K, Lydia, Karima), compagnons d'étude et de labeur, merci pour l'entraide précieuse, les échanges enrichissants et les moments de complicité qui ont rendu ce parcours plus léger et plus joyeux. Vous resterez à jamais ma "deuxième famille universitaire".

À tous ceux qui, par leur affection, leurs conseils ou leur simple présence, ont contribué de près ou de loin à l'aboutissement de ce travail, recevez ici l'expression de ma profonde reconnaissance.

Remerciements

Avant toute chose, je rends grâce à Dieu Tout-Puissant de m'avoir donné la force et la persévérance nécessaires pour accomplir ce travail.

Ma profonde gratitude va à ma promotrice, Mme ATIL Lynda, qui a été bien plus qu'une simple guide. Sa confiance, sa patience et son soutien constant ont été des forces précieuses tout au long de ce parcours. Je lui suis sincèrement reconnaissante pour son encadrement précieux, ses conseils éclairés et sa disponibilité sans faille.

Je remercie sincèrement les membres du jury, Mme BELKACEM Cherifa, M FELLAG Hocine, pour avoir accepté d'évaluer ce mémoire et pour leurs remarques constructives.

Je remercie chaleureusement l'ensemble des enseignants du département de mathématiques pour leur transmission du savoir et leur accompagnement pédagogique durant tout mon cursus universitaire.

Que toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce mémoire trouvent ici l'expression de ma gratitude.

Table des matières

| | | |
|----------|---|-----------|
| 1 | Notions générales sur les séries chronologiques | 7 |
| 1.1 | Définitions et propriétés | 7 |
| 1.1.1 | Série chronologique | 7 |
| 1.1.2 | Composantes d'une série chronologique | 7 |
| 1.1.3 | Modélisation d'une série chronologique | 8 |
| 1.1.4 | Opérateurs sur les séries chronologiques | 9 |
| 1.1.5 | Processus Bruit Blanc (White Noise Process) | 10 |
| 1.1.6 | La stationnarité | 11 |
| 1.1.7 | La non stationnarité | 11 |
| 1.1.8 | Fonction d'autocovariance | 12 |
| 1.1.9 | Autocorrélations simples et partielles | 12 |
| 1.1.10 | Les modèles autorégressifs | 13 |
| 1.2 | Tests de stationnarité | 14 |
| 1.2.1 | Test de Dickey-Fuller Augmenté (ADF) | 15 |
| 1.2.2 | Test de Phillips-Perron (PP) | 17 |
| 1.2.3 | Test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) | 19 |
| 2 | La cointégration linéaire | 22 |
| 2.1 | Introduction | 22 |
| 2.2 | Modèle à correction d'erreur | 23 |
| 2.3 | Tests de cointégration | 24 |
| 2.3.1 | Test de cointégration de Engle et Granger (1987) | 24 |
| 2.3.2 | Test de cointégration de Johansen (1988) | 25 |
| 2.3.3 | Test de cointégration de Phillips-Ouliaris (1990) | 28 |
| 2.4 | Comparaison des tests de cointégration | 30 |
| 3 | Les modèles Auto Regressive Distributed Lag (ARDL) | 31 |
| 3.1 | Introduction | 31 |
| 3.1.1 | Propriétés du modèle ARDL | 33 |
| 3.1.2 | Critères de sélection de modèle | 33 |
| 3.2 | L'approche du test aux bornes ou test de cointégration de Pesaran et al. (2001) | 34 |
| 3.3 | Estimation dans les modèles ARDL | 38 |
| 3.4 | Méthodes d'estimation du modèle ARDL | 40 |

| | | |
|-------|--|-----------|
| 3.4.1 | Estimation du modèle ARDL par la méthode des moindres carrés ordinaires (MCO) | 40 |
| 3.4.2 | Estimation d'un modèle ARDL par la méthode du maximum de vraisemblance (MV) | 43 |
| 3.5 | Inférence Bayésienne dans les Modèles ARDL et Extension des Bounds Tests | 47 |
| 3.5.1 | Inférence Bayésienne dans les Modèles ARDL | 47 |
| 3.6 | Application sur données réelles | 51 |
| 3.6.1 | Présentation des données "Denmark" | 51 |
| 3.6.2 | Test de Johansen (avec constante dans la relation de cointégration) | 55 |
| 3.6.3 | Présentation de la fonction <code>ardl()</code> et de ses composantes | 59 |
| 3.6.4 | Présentation de la fonction <code>auto_ardl()</code> et de ses composantes | 61 |
| 3.6.5 | Test des bornes de cointégration avec <code>bounds_f_test</code> | 64 |
| 3.6.6 | Test des bornes basé sur la statistique t (<code>bounds_t_test</code>) | 67 |
| 3.7 | Exemple appliqué sur des données simulées d'un modèle ARDL avec détails sur l'estimation | 69 |
| | Conclusion générale | 77 |

Introduction générale

Les séries chronologiques constituent un outil fondamental en économétrie et en statistique appliquée, permettant d'analyser l'évolution temporelle de variables économiques, financières ou environnementales. Leur utilisation s'étend à des domaines variés tels que la prévision macroéconomique, l'évaluation de politiques publiques ou l'étude des marchés financiers. L'importance des séries temporelles réside dans leur capacité à capturer des dynamiques complexes et à révéler des relations structurelles entre variables.

Pour exploiter pleinement ces données, la notion de stationnarité est essentielle. Une série est dite stationnaire lorsque ses propriétés statistiques (moyenne, variance, autocorrélation) restent stables dans le temps. Cette hypothèse est cruciale pour éviter des résultats fallacieux dans les modélisations. Pour vérifier la stationnarité, des tests de racine unitaire tels que le test de Dickey-Fuller augmenté (ADF), Phillips-Perron (PP) ou KPSS sont employés. Ces outils permettent de détecter la présence de tendances stochastiques ou déterministes.

Lorsque plusieurs séries non stationnaires partagent une relation d'équilibre à long terme, on parle de cointégration. Cette notion, introduite par Engle et Granger (1987)[8], ouvre la voie à la modélisation des dynamiques conjointes via des modèles à correction d'erreur (ECM). Les tests de cointégration (Engle-Granger, Johansen, Phillips-Ouliaris) exigent toutefois que les séries soient intégrées du même ordre, une contrainte restrictive dans de nombreuses applications empiriques.

Pour pallier cette limite, les modèles ARDL (Autoregressive Distributed Lag) ont été développés. Leur flexibilité permet d'analyser des relations de court et long terme entre variables intégrées d'ordres différents, sans nécessiter de prétests de cointégration. Le *bounds test* de Pesaran et al. (2001)[16] et l'estimation par MCO/MV en font un cadre robuste, complété aujourd'hui par des approches bayésiennes pour mieux quantifier l'incertitude.

Dans ce mémoire, nous proposons une analyse approfondie de ces concepts et leur application à des données réelles. Le premier chapitre introduit les séries chronologiques, rappelle les concepts de stationnarité et présente les tests ADF, PP et KPSS. Le deuxième chapitre explore la cointégration, les modèles ECM et les tests associés (Engle-Granger, Johansen). Enfin, le troisième chapitre détaille le modèle ARDL, ses propriétés et son estimation (MCO, MV, bayésienne), avant une application empirique sur des données macroéconomiques du Danemark. Les résultats des tests de Johansen et du *bounds test* y sont comparés, mettant en lumière la cohérence des approches.

Chapitre 1

Notions générales sur les séries chronologiques

1.1 Définitions et propriétés

1.1.1 Série chronologique

Les séries chronologiques, ou séries temporelles, constituent un outil fondamental pour analyser l'évolution dynamique de phénomènes observés à intervalles réguliers. Leur utilisation traverse aujourd'hui une multitude de domaines : en économétrie pour modéliser la relation entre variables macroéconomiques, en médecine pour suivre l'évolution de maladies épidémiques, ou en démographie pour anticiper les tendances de population. Cette transversalité s'explique par leur capacité à décrire, expliquer et prédire des comportements temporels complexes, en combinant rigueur statistique et interprétation contextuelle.

Définition 1.1.1. *Une série chronologique (ou série temporelle) est une suite de variables aléatoires $\{X_t; t \in T\}$ où l'ensemble T est un intervalle de temps qui peut être discret ou continu.*

1.1.2 Composantes d'une série chronologique

Une série X_t se décompose généralement en trois composantes principales

a. La tendance (T_t)

La tendance (T_t) représente l'évolution à long terme de la série étudiée. Elle traduit le comportement « moyen » de la série dans le temps. La tendance peut prendre plusieurs formes, telles qu'une fonction linéaire, polynômiale, exponentielle ou logarithmique, selon la nature des données et des phénomènes observés.

— Tendance linéaire

La tendance linéaire est la forme la plus simple de tendance. Elle peut être estimée à l'aide de la méthode des moindres carrés, ce qui revient à effectuer une régression linéaire simple du type :

$$T_t = a + bt$$

— Tendance quadratique

Elle est généralement modélisée par l'équation suivante :

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

— **Tendance polynômiale d'ordre q**

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

— **Tendance logistique**

La fonction logistique permet de modéliser des processus dont la croissance est limitée par une valeur maximale c . Elle est souvent utilisée pour représenter des phénomènes de saturation ou de croissance asymptotique. La forme générale de la tendance logistique est la suivante :

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

où :

- c : valeur asymptotique maximale (capacité limite)
- a : paramètre lié à la vitesse de croissance
- b : paramètre lié aux conditions initiales
- e : base du logarithme népérien

b. La saisonnalité (S_t)

La composante saisonnière S_t représente les effets récurrents liés aux saisons météorologiques (comme la production agricole ou la consommation de gaz) ou à des activités économiques et sociales (fêtes, vacances, soldes, etc.). Mathématiquement, ce sont des fonctions périodiques, c'est-à-dire qu'il existe un entier p , appelé *période*, tel que :

$$s_t = s_{t+p} \quad \text{pour tout } t \geq 1$$

Au premier abord, cette composante est entièrement déterminée par ses p premières valeurs :

$$s_1, s_2, s_3, \dots, s_p$$

c. La composante résiduelle

La composante résiduelle, ou *résidu*, notée généralement ε_t , représentent la partie non structurée du phénomène. Elle est modélisée par une suite de variables aléatoires ε_t , centrées, non corrélées et de même variance.

1.1.3 Modélisation d'une série chronologique

Il existe trois types de schémas ou de modèles pour la décomposition des séries chronologiques :

- **Modèle additif** : La série temporelle Y_t est exprimée comme la somme de ses composantes :

$$Y_t = T_t + S_t + \varepsilon_t$$

où T_t est la tendance, S_t la composante saisonnière, et ε_t le terme d'erreur. Ce modèle suppose que l'amplitude des variations saisonnières est constante dans le temps.

— **Modèle multiplicatif** : La série est représentée par le produit de ses composantes :

$$Y_t = T_t \times S_t \times \varepsilon_t$$

Ici, les variations saisonnières et les erreurs varient proportionnellement à la tendance, ce qui signifie que l'amplitude des fluctuations augmente avec le niveau de la série.

— **Modèle mixte** : Ce modèle combine des aspects additifs et multiplicatifs, par exemple :

$$Y_t = (T_t \times S_t) + \varepsilon_t$$

où la composante saisonnière agit de manière multiplicative tandis que le terme d'erreur est additif.

1.1.4 Opérateurs sur les séries chronologiques

Pour modéliser les processus stochastiques (et donc les séries chronologiques), on utilise plusieurs opérateurs fondamentaux.

Opérateur retard (B) ou (L)

On note indifféremment B (Backward) ou L (lag), l'opérateur retard qui décale le processus d'une unité de temps vers le passé.

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

Si on applique n fois cet opérateur, on décale le processus de n unités de temps.

$$\underbrace{B(B(\dots B(X_t) \dots))}_{n \text{ fois}} = B^n(X_t) = X_{t-n}$$

Opérateur d'avance (F)

L'opérateur d'avance (Forward Shift Operator), généralement noté F ou parfois B^{-1} ou L^{-1} , est l'opérateur dual de l'opérateur retard. Il décale la série d'une unité vers le futur.

$$F(X_t) = X_{t+1}$$

Si on applique n fois cet opérateur, on décale le processus de n unités de temps.

$$\underbrace{F(F(\dots F(X_t) \dots))}_{n \text{ fois}} = F^n(X_t) = X_{t+n}$$

Opérateur différence (Δ)

L'opérateur différence Δ mesure l'écart entre un processus et sa valeur décalée d'une unité de temps. Il se définit à l'aide de l'opérateur retard (B) et de l'opérateur identité (I).

$$\begin{aligned} \Delta(X_t) &= X_t - X_{t-1} = X_t - B(X_t) \\ &\Leftrightarrow \Delta = I - B \end{aligned}$$

où

$$I(X_i) = X_i$$

Dans certains ouvrages Δ est noté ∇ .

Propriétés de l'opérateur différence Δ

— Élimination des tendances

L'opérateur Δ élimine les tendances linéaires. Pour un processus de la forme :

$$X_t = at + b + E_t \quad (E_t \text{ est stationnaire})$$

on a

$$\Delta(X_t) = X_t - X_{t-1} = (at + b + E_t) - (a(t-1) + b + E_{t-1}) = a + (E_t - E_{t-1})$$

Le résultat $a + (E_t - E_{t-1})$ est stationnaire (la tendance linéaire at est supprimée).

De façon générale, l'opérateur Δ^d élimine les tendances polynomiales de degré d . Par exemple, pour une tendance de degré 2 :

$$X_t = at^2 + bt + c + E_t,$$

on a :

$$\Delta^2 = (I - B)^2 = I - 2B + B^2$$

et donc

$$\Delta^2(X_t) = X_t - 2X_{t-1} + X_{t-2} = \dots = 2a + (E_t - 2E_{t-1} + E_{t-2})$$

— Élimination de la saisonnalité

L'opérateur de différence saisonnière $\Delta_\omega = (I - B^\omega)$ est utilisé pour éliminer une saisonnalité de période ω (ω est un entier, comme 12 pour des données mensuelles).

Si on a un processus de la forme

$$X_{k\omega+j} = s_j + E_{k\omega+j} \quad (E_t \text{ stationnaire}).$$

On a

$$\begin{aligned} \Delta_\omega(X_{k\omega+j}) &= (I - B^\omega)(X_{k\omega+j}) = X_{k\omega+j} - X_{(k-1)\omega+j} \\ &= (s_j + E_{k\omega+j}) - (s_j + E_{(k-1)\omega+j}) \\ &= E_{k\omega+j} - E_{(k-1)\omega+j} \end{aligned}$$

· **Résultat** : La saisonnalité s_j est éliminée, ne laissant que la composante aléatoire stationnaire.

1.1.5 Processus Bruit Blanc (White Noise Process)

Définition 1.1.2. "Bruit blanc faible"

On dit de la suite de variables aléatoires (ε_t) qu'elle constitue un bruit blanc faible si elle vérifie les propriétés suivantes :

— *Espérance nulle* :

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

— *variance constante* :

$$\text{var}(\varepsilon_t) = \sigma^2$$

— Absence d'autocorrélation :

$$\text{cov}(\varepsilon_t, \varepsilon_s) = 0 \quad \text{si } t \neq s$$

Définition 1.1.3. "bruit blanc fort"

On dit que (ε_t) est un bruit blanc fort s'il est un bruit blanc faible et que les variables aléatoires (v.a) (ε_t) sont indépendantes et identiquement distribuées (i.i.d).

Remarque 1.1.1. Tout bruit blanc faible gaussien est un bruit blanc fort.

1.1.6 La stationnarité

Une propriété importante des séries chronologiques est la stationnarité. La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série $(X_t)_{t \in \mathbb{Z}}$ est stable si elle ne présente ni tendance saisonnière, ni tendance à la hausse ou à la baisse. Plus formellement, on distingue deux types de stationnarité :

Définition 1.1.4. on dit qu'une série temporelle X_t est fortement stationnaire si la loi est invariante par translation dans le temps.

$$\forall k, \forall (t_1, t_2, \dots, t_k), \mathcal{L}(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = \mathcal{L}(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$$

La notion de stationnarité forte est généralement beaucoup trop forte pour les applications. Dans la pratique, il est habituellement suffisant de considérer des modèles faiblement stationnaires (ou stationnaires du second ordre), pour lesquels seules la moyenne, la variance et la covariance (à décalage fixe) sont constantes dans le temps.

Définition 1.1.5. Une suite $\{X_t : t \geq 0\}$ de variables aléatoires est dite stationnaire du second ordre si elle vérifie les propriétés suivantes :

1. Pour tout t , $\mathbb{E}(X_t^2) < +\infty$.
2. Pour tout t , $\mathbb{E}(X_t) = \mu$, indépendante de t .
3. Pour tout t et pour tout h , $\text{cov}(X_t, X_{t+h}) = \gamma(h)$ est indépendante de t .

1.1.7 La non stationnarité

La plupart des séries économiques sont non stationnaires, c'est-à-dire que le processus qui les décrit ne vérifie pas au moins l'une des conditions de la définition d'un processus stationnaire du second ordre. Cela conduit à distinguer deux types de processus non stationnaires :

- les **processus TS (Trend Stationary)** ;
- les **processus DS (Difference Stationary)**.

Processus TS

On dit que le processus Y_t est caractérisé par une non-stationnarité déterministe, ou encore que le processus Y_t est *TS* (Trend Stationary), s'il peut s'écrire :

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

où $f(t)$ est une fonction déterministe du temps (par exemple une tendance linéaire), et Z_t est un processus stationnaire du second ordre.

Processus DS

On dit que le processus Y_t est caractérisé par une non-stationnarité stochastique, ou encore que le processus Y_t est de type *DS* (Difference Stationary), si le processus différencié une fois, $(1 - L)Y_t$, est stationnaire.

On parle aussi de processus intégré d'ordre 1, ce que l'on note $Y_t \sim I(1)$:

$$(1 - L)Y_t = Z_t \quad \text{qui est Stationnaire} \quad \Rightarrow \quad Y_t = Y_{t-1} + Z_t$$

1.1.8 Fonction d'autocovariance

La fonction d'autocovariance mesure la dépendance linéaire entre deux valeurs d'une série temporelle à différents instants. Pour un processus stochastique stationnaire $\{X_t\}$ se moyenne constante μ , la fonction d'autocovariance au retard k est définie par :

$$\gamma(k) = \text{Cov}(X_t, X_{t-k}) = \mathbb{E}[(X_t - \mu)(X_{t-k} - \mu)]$$

Propriétés essentielles

Les propriétés importantes de la fonction d'autocovariance sont :

$$\gamma(0) = \text{Var}(X_t) \quad (\text{Variance du processus})$$

$$\gamma(-k) = \gamma(k) \quad (\text{fonction paire})$$

$$|\gamma(k)| \leq \gamma(0)$$

Cette fonction est essentielle pour identifier la structure de dépendance temporelle dans un modèle et pour la construction de modèles ARMA/ARIMA.

1.1.9 Autocorrélations simples et partielles

Autocorrélation simple

L'autocorrélation simple (ou fonction d'autocorrélation – ACF) mesure la corrélation entre les valeurs d'une série à différents retards (lags). La fonction d'autocorrélation pour un processus X_t au retard k est définie par :

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

Avec

$$\rho(0) = 1 \quad \text{et} \quad |\rho(h)| < 1$$

On définit la matrice de corrélation (de dimension m) de la manière suivante :

$$\mathbf{R}(m) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(m-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(m-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(m-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(m-1) & \rho(m-2) & \rho(m-3) & \cdots & 1 \end{pmatrix}$$

La fonction autocorrélation partielle

La fonction d'autocorrélation partielle mesure la liaison linéaire entre Y_t et Y_{t-h} , une fois retirés les effets transitant par les variables intermédiaires $Y_{t-1}, \dots, Y_{t-h+1}$.

Le coefficient d'autocorrélation partielle d'ordre h , noté $r(h)$, est défini comme le coefficient de corrélation entre :

$$Y_t - \mathbb{E}(Y_t \mid Y_{t-1}, \dots, Y_{t-h+1}) \text{ et} \\ Y_{t-h} - \mathbb{E}(Y_{t-h} \mid Y_{t-1}, \dots, Y_{t-h+1})$$

Autrement dit, $r(h)$ mesure la corrélation entre les résidus obtenus après avoir retiré l'effet des variables intermédiaires. C'est également le coefficient de Y_{t-h} dans la régression de Y_t sur $Y_{t-1}, \dots, Y_{t-h+1}, Y_{t-h}$.

Dans le cas d'un processus stationnaire, le coefficient d'autocorrélation partielle d'ordre h peut être calculé à l'aide des déterminants de matrices de corrélations. On a :

$$r(h) = \frac{|\mathbf{R}^*(h)|}{|\mathbf{R}(h)|}$$

où $\mathbf{R}(h)$ est la matrice de Toeplitz symétrique définie comme suit :

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

et $\mathbf{R}^*(h)$ est la même matrice où la dernière colonne est remplacée par :

$$\begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(h) \end{pmatrix}$$

Par exemple :

$$r(1) = \rho(1) \\ r(2) = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}$$

Ainsi, le coefficient d'autocorrélation partielle à l'ordre 1 est simplement égal à l'autocorrélation simple. Pour des ordres supérieurs, l'autocorrélation partielle corrige les effets de transitivité en supprimant l'influence des observations intermédiaires.

1.1.10 Les modèles autorégressifs

Un modèle autorégressif (AR) est un modèle statistique utilisé en analyse de séries temporelles dans lequel la valeur actuelle d'une variable est exprimée comme une combinaison linéaire de ses valeurs passées, plus un terme d'erreur aléatoire.

Définition 1.1.6. *Un processus autorégressif d'ordre p , noté $AR(p)$, est donné par :*

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t$$

où $\varphi_1, \dots, \varphi_p$ sont les paramètres du modèle, "c" est une constante, et ε_t un bruit blanc. En utilisant l'opérateur des retards L , on peut écrire le modèle :

$$\begin{aligned} (1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) X_t &= c + \varepsilon_t \\ \Leftrightarrow \varphi(L) X_t &= \varepsilon_t \end{aligned}$$

La stationnarité d'un processus AR(p)

Soit (X_t) un processus autorégressif d'ordre p , noté $AR(p)$. Ce processus est dit stationnaire si toutes les racines du polynôme caractéristique $\varphi(L)$ sont de module strictement supérieur à 1.

Cas d'un AR(1)

Considérons le processus autorégressif d'ordre 1 :

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

On peut réécrire le modèle avec l'opérateur retard L :

$$(1 - \varphi L) X_t = \varepsilon_t$$

Le polynôme caractéristique est alors :

$$\phi(L) = 1 - \varphi L$$

Condition de stationnarité

Le processus est stationnaire si toutes les racines du polynôme caractéristique sont de module strictement supérieur à 1. Résolvons :

$$\phi(L) = 0 \quad \Rightarrow \quad 1 - \varphi L = 0 \quad \Rightarrow \quad L = \frac{1}{\varphi}$$

on conclure donc :

$$\left| \frac{1}{\varphi} \right| > 1 \quad \Leftrightarrow \quad |\varphi| < 1$$

Le processus AR(1) est stationnaire si et seulement si :

$$\boxed{|\varphi| < 1}$$

Exemple 1.1.1. Soit $\varphi = 0,7$ et $\varepsilon_t \sim \mathcal{N}(0, 1)$.

Le processus :

$$X_t = 0,7 X_{t-1} + \varepsilon_t$$

est **stationnaire**, car $|\varphi| = 0,7 < 1$.

1.2 Tests de stationnarité

Les tests de stationnarité sont des méthodes statistiques utilisées pour déterminer si une série temporelle est stationnaire ou non, il existe plusieurs tests sur la stationnarité, les plus utilisés étant le test de Dickey-Fuller augmenté (ADF), le test de Phillips-Perron (PP), le test de Zivot-Andrews (ZA) et le test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS).

1.2.1 Test de Dickey-Fuller Augmenté (ADF)

Le test de Dickey-Fuller augmenté (Augmented Dickey-Fuller, ADF) est un test statistique utilisé pour déterminer si une série temporelle possède une racine unitaire, c'est-à-dire si elle est non stationnaire. Il a été développé par Dickey et Fuller en 1981 ([6]) ; il s'agit d'une version améliorée du test de Dickey-Fuller standard, conçue pour prendre en compte les corrélations sérielles d'ordre supérieur dans les données. Pour comprendre comment fonctionne ce test, il faut analyser la reparamétrisation du modèle.

Considérons une série temporelle Y_t modélisé par un processus AR(1)

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

on a

- Si $\rho = 1$, la série a une racine unitaire (non-stationnaire).
- Si $|\rho| < 1$, la série est stationnaire.

Le test Dickey-Fuller classique test l'hypothèse :

- $H_0 : \rho = 1$ (racine unitaire).
- $H_1 : |\rho| < 1$ (stationnaire).

Le test ADF réécrit le modèle en différence première :

$$\Delta Y_t = \alpha + (\rho - 1)Y_{t-1} + \varepsilon_t$$

En posant $\phi = \rho - 1$, on obtient :

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t$$

Les hypothèses de test ADF deviennent :

- $H_0 : \phi = 0$ (il existe une racine unitaire).
- $H_1 : |\phi| < 0$ (la série est stationnaire)

Généralisation au cas AR(p)

Pour un processus AR(p), l'équation de test ADF devient :

$$\Delta Y_t = \alpha + \phi Y_{t-1} - \sum_{j=1}^p \gamma_j \Delta Y_{t-j} + \varepsilon_t \quad (1.2.1)$$

Formulation du modèle

Il exist trois versions du modèle :

1. Modèle sans constante ni tendance

$$\Delta Y_t = \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

2. Modèle avec constante

$$\Delta Y_t = \alpha + \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

3. Modèle avec constante et tendance

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

$\Delta Y_t = Y_t - Y_{t-1}$ est la différence première de la série, ϕ est le coefficient clé (si $\phi = 0$, la série a une racine unitaire, γ_i sont les coefficients des différences (ajouter pour corriger l'autocorrélation), p est le nombre de retards.

Statistique de test

Pour tester l'hypothèse nulle $H_0 : \phi = 0$, Dickey et Fuller (1979)([5]) ont proposé la statistique studentisée suivante :

$$t_n = \frac{\hat{\phi}_n}{\text{Std}(\hat{\phi}_n)}$$

où $\text{Std}(\hat{\phi}_n)$ est un estimateur de l'écart-type de l'estimateur des moindres carrés ordinaires $\hat{\phi}_n$.

Distribution non standard

Sous l'hypothèse nulle, cette statistique ne suit pas une loi de Student standard, même asymptotiquement. La véritable loi asymptotique est tabulée dans Fuller (1976)([5]) et est la suivante :

$$t_n \sim \frac{W^2(1) - 1}{2 \left[\int_0^1 W^2(s) ds \right]^{1/2}}, \quad (1.2.2)$$

où $W(\cdot)$ désigne un mouvement brownien standard sur l'intervalle $[0, 1]$.

La décision de rejeter ou non l'hypothèse nulle dépend de la comparaison de la statistique observée avec ces valeurs critiques.

Pratique du test

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

Étape 1 : Estimation du modèle avec constante et tendance (modèle 3)

On estime le modèle suivant :

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

On teste si le coefficient de la tendance β est significativement différent de zéro (test de Student).

Si la statistique calculée $t_{\text{calc}} > t_{\text{tab}}$, alors β est significatif : on rejette $H_{0\beta} : \beta = 0$.

Sinon, on accepte $H_{0\beta}$.

On teste ensuite l'hypothèse $H_0 : \phi = 0$ (présence d'une racine unitaire).

— Si la statistique associée à $\hat{\phi}$ est supérieure à la valeur critique, on accepte H_0 : la série est non stationnaire, mais peut être stationnarisée en retirant la tendance.

— Sinon, on rejette H_0 : la série est stationnaire.

Si β n'est pas significatif, on passe à l'étape 2.

Étape 2 : Estimation du modèle avec constante seulement (modèle 2)

On estime le modèle suivant :

$$\Delta Y_t = \alpha + \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

On teste si la constante α est significative.

- Si $t_{\text{calc}} > t_{\text{tab}}$, alors α est significative, et on procède au test de racine unitaire sur ϕ .
- Si la statistique de test est supérieure à la valeur critique, on accepte H_0 : la série est non stationnaire. une différenciation est nécessaire.
- Sinon, on rejette H_0 : la série est stationnaire.

Si α n'est pas significative, on passe à l'étape 3.

Étape 3 : Estimation du modèle sans constante ni tendance (modèle 1)

On estime le modèle suivant :

$$\Delta Y_t = \phi Y_{t-1} - \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

On teste l'hypothèse $H_0 : \phi = 0$ (présence d'une racine unitaire).

- Si la statistique est supérieure à la valeur critique, on accepte H_0 : la série est non stationnaire (différenciation nécessaire).
- Sinon, la série est stationnaire.

1.2.2 Test de Phillips-Perron (PP)

Le test de Phillips-Perron (1988)([21]) est une méthode statistique utilisée en économétrie pour déterminer si une série temporelle possède une racine unitaire (non-stationnaire). Il offre une alternative au test de Dickey-Fuller pour traiter la corrélation sérielle dans les erreurs du modèle. Contrairement au test ADF, qui ajoute des retards des différences de la variable, le test de Phillips-Perron applique une correction non paramétrique aux statistiques du test. Cette correction permet de tenir compte de l'autocorrélation et de l'hétéroscédasticité sans modifier la structure de l'équation estimée. Les statistiques obtenues ont la même distribution asymptotique que celles du test de Dickey-Fuller.

Considérons une série temporelle Y_t modélisé par un processus AR(1)

$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t \tag{1.2.3}$$

Hypothèses du test

Les hypothèses du test sont les suivantes :

- **Hypothèse nulle** $H_0 : \rho = 1$ (la série possède une racine unitaire, c'est-à-dire qu'elle est non stationnaire)
- **Hypothèse alternative** $H_1 : |\rho| < 1$ (la série est stationnaire)

Les statistiques du test

Le test de Phillips-Perron propose trois statistiques du test selon la présence d'une constante ou d'une tendance, pour tester l'hypothèse nulle de racine unitaire ($\rho = 1$).

Pour le modèle (1)

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

La statistique t de Phillips-Perron pour ρ (l'estimateur de ρ) est donnée par :

$$Z(\hat{t}_\rho) = \left(\frac{\hat{s}}{\hat{\sigma}_{Tl}} \right) \hat{t}_\rho - \frac{1}{2} \frac{\hat{\sigma}_{Tl}^2 - \hat{s}^2}{\hat{\sigma}_{Tl} \left(T - 2 \sum_{t=2}^T Y_{t-1}^2 \right)^{1/2}} \quad (1.2.4)$$

où \hat{s}^2 est l'estimateur de la variance "court terme" de ε_t :

$$\hat{s}^2 = \frac{1}{T} \sum_{t=2}^T \hat{\varepsilon}_t^2. \quad (1.2.5)$$

et $\hat{\sigma}_{Tl}^2$ est l'estimateur de la variance "long terme" de ε_t (voir Newey et West (1987)[17]) :

$$\hat{\sigma}_{Tl}^2 = \hat{s}^2 + \frac{2}{T} \sum_{s=1}^l \left(1 - \frac{s}{l+1} \right) \sum_{t=s+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-s}. \quad (1.2.6)$$

Les tests de Dickey-Fuller apparaissent ainsi comme des cas particuliers des tests de Phillips-Perron, si $\hat{s}^2 = \hat{\sigma}_{Tl}^2$.

Pour le modèle (2)

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

les statistiques sont les suivantes :

$$Z(\tilde{t}_\rho) = \left(\frac{\tilde{s}}{\tilde{\sigma}_{Tl}} \right) \tilde{t}_\rho - \frac{1}{2} \frac{\tilde{\sigma}_{Tl}^2 - \tilde{s}^2}{\tilde{\sigma}_{Tl} \left(T - 2 \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{-1})^2 \right)^{1/2}}, \quad (1.2.7)$$

$$Z(\tilde{t}_\alpha) = \left(\frac{\tilde{s}}{\tilde{\sigma}_{Tl}} \right) \tilde{t}_\alpha + \frac{1}{2} \frac{\sqrt{T} (\tilde{\sigma}_{Tl}^2 - \tilde{s}^2)}{\tilde{\sigma}_{Tl} \left(T - 2 \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{-1})^2 \right)^{1/2}} \times \frac{\sum_{t=2}^T Y_{t-1}}{\left(\sum_{t=2}^T Y_{t-1}^2 \right)^{1/2}} \quad (1.2.8)$$

où $\tilde{\sigma}_{Tl}^2$ et \tilde{s}^2 sont donnés respectivement par (4) et (5), en substituant $\tilde{\varepsilon}_t$ à $\hat{\varepsilon}_t$ et en supposant que $\bar{Y}_{-1} = (T-1)^{-1} \sum_{t=2}^T Y_{t-1}$.

Pour le modèle (3)

$$Y_t = \alpha + \beta t + \rho Y_{t-1} + \varepsilon_t$$

les statistiques sont données par :

$$Z(t_\rho^*) = \left(\frac{s^*}{\sigma_{Tl}^*} \right) t_\rho^* - \frac{T^3}{4\sqrt{3}} \frac{\sigma_{Tl}^{*2} - s^{*2}}{D_{xx}^{1/2} \sigma_{Tl}^*}, \quad (1.2.9)$$

$$Z(t_\alpha^*) = \left(\frac{s^*}{\sigma_{Tl}^*} \right) t_\alpha^* - \frac{T^4 \sqrt{T} (\sigma_{Tl}^{*2} - s^{*2}) \sum_{t=2}^T Y_{t-1}}{4 \left[3\sigma_{Tl}^{*2} D_{xx} \left(12D_{xx} + T^3 \left(\sum_{t=2}^T Y_{t-1} \right)^2 \right) \right]^{1/2}}, \quad (1.2.10)$$

$$Z(t_\beta^*) = \left(\frac{s^*}{\sigma_{Tl}^*} \right) t_\beta^* - \frac{T\sqrt{T}(\sigma_{Tl}^{*2} - s^{*2}) \left(T \sum_{t=2}^T Y_{t-1} - 2 \sum_{t=2}^T tY_{t-1} \right)}{4 \left[\sigma_{Tl}^{*2} D_{xx} \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{-1})^2 \right]^{1/2}}, \quad (1.2.11)$$

avec D_{xx} étant le déterminant du produit intérieur de la matrice de données avec elle-même, c'est-à-dire :

$$D_{xx} = \frac{T^2(T^2 - 1)}{12} \sum_{t=2}^T Y_{t-1}^2 - T \left(\sum_{t=2}^T tY_{t-1} \right)^2 + T(T+1) \sum_{t=2}^T tY_{t-1} \sum_{t=2}^T Y_{t-1} - \frac{T(T+1)(2T+1)}{6} \left(\sum_{t=2}^T Y_{t-1} \right)^2 \quad (1.2.12)$$

où σ_{Tl}^{*2} et s^{*2} sont donnés respectivement par (1.2.5) et (1.2.6) en remplaçant $\hat{\varepsilon}_t$ par $\tilde{\varepsilon}_t$.

Généralement, on prend : $l = O(T^{1/4})$. Schwert (1989)[24] suggère de calculer les statistiques de Phillips-Perron pour deux valeurs de l :

$$l_4 = \text{int} \left[4 \left(\frac{T}{100} \right)^{1/4} \right] \quad \text{et} \quad l_{12} = \text{int} \left[12 \left(\frac{T}{100} \right)^{1/4} \right] \quad (1.2.13)$$

où la notation $\text{int}[x]$ désigne la fonction partie entière qui retourne le plus grand entier inférieur ou égal à x . Pour conclure, notons que si $l = 0$, les tests de Phillips-Perron se réduisent aux procédures simplifiées de Dickey-Fuller.

1.2.3 Test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

Contrairement aux tests classiques, le test KPSS, proposé par Kwiatkowski *et al.* (1992)([16]), formule comme hypothèse nulle la stationnarité de la série temporelle, contre l'alternative de l'existence d'une racine unitaire (non-stationnarité). La série temporelle est modélisée comme la somme d'une tendance déterministe, d'une marche aléatoire et d'une composante stationnaire. L'hypothèse nulle de stationnarité revient à supposer que la variance de la marche aléatoire est nulle. La série temporelle Y_t , pour $t = 1, \dots, T$, est générée selon :

$$Y_t = \alpha t + r_t + \varepsilon_t$$

où :

αt : Tendance déterministe linéaire

$r_t = r_{t-1} + u_t$ ($u_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_u^2)$) : Marche aléatoire

$\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$: Bruit blanc stationnaire

Hypothèses de test

- **Hypothèse nulle** (H_0) : $\sigma_u^2 = 0$, donc r_t est constant, ce qui implique que Y_t est stationnaire autour d'une tendance (ou d'un niveau si $\alpha = 0$).
- **Hypothèse alternative** (H_1) : $\sigma_u^2 > 0$, donc r_t suit une marche aléatoire. Y_t est non stationnaire.

La statistique de test

Kwiatkowski *et al.* (1992)([16]) proposent d'utiliser un test du multiplicateur de Lagrange (LM) pour tester l'hypothèse nulle de stationnarité, c'est-à-dire $\sigma_u^2 = 0$. La statistique LM est alors donnée par :

$$LM = \frac{1}{\hat{\sigma}_\varepsilon^2} \sum_{t=1}^T S_t^2 \quad (1.2.14)$$

où S_t est la somme partielle des résidus e_t définie par :

$$S_t = \sum_{i=1}^t e_i, \quad t = 1, \dots, T$$

et $\hat{\sigma}_\varepsilon^2$ est un estimateur de la variance des résidus e_t .

Il est important de noter que la statistique LM a été dérivée sous l'hypothèse que les erreurs ε_t sont indépendantes et identiquement distribuées selon une loi normale $\mathcal{N}(0, \sigma_u^2)$. Cependant, cette hypothèse est souvent irréaliste et le test KPSS suppose que les ε_t satisfont aux conditions de régularité de Phillips et Perron (1988)([21]), afin de permettre des formes générales de dépendance temporelle.

Comme dans le test de Phillips-Perron (1988)([21]), la variance à long terme est définie par :

$$\sigma^2 = \lim_{T \rightarrow \infty} \frac{E(S_T^2)}{T}.$$

Un estimateur de σ^2 est donné par :

$$\hat{\sigma}_{Tl}^2 = \frac{1}{T} \sum_{t=1}^T e_t^2 + 2 \sum_{s=1}^l \left(1 - \frac{s}{l+1}\right) \sum_{t=s+1}^T e_t e_{t-s}.$$

Comme on peut le voir dans (1.2.12), le dénominateur de la statistique LM est $\hat{\sigma}_\varepsilon^2$. Cependant, lorsque les erreurs ne sont pas i.i.d., le dénominateur approprié est un estimateur de σ^2 au lieu de σ_ε^2 .

Dans ce cas, il existe deux statistiques de test

Considérons d'abord le cas stationnaire au niveau. La statistique de test est donnée par :

$$\eta_\mu = \frac{T^2}{\hat{\sigma}_{Tl}^2} \sum_{t=1}^T S_t^2$$

la distribution limite de cette statistique est la suivante

$$\eta_\mu \rightarrow \int_0^1 V^2(s) ds$$

où $V(s)$ est un pont brownien standard.

Considérons maintenant le cas de tendance stationnaire. La statistique de test s'écrit :

$$\eta_\tau = \frac{T^2}{\hat{\sigma}_{Tl}^2} \sum_{t=1}^T S_t^2$$

avec

$$\eta_\tau \rightarrow \int_0^1 V_2^2(s) ds$$

où $V_2(s)$ est le second pont brownien, donné par :

$$V_2(s) = W(s) + (2s - 3s^2)W(1) + (-6s + 6s^2) \int_0^1 W(x) dx$$

où $W(s)$ étant un mouvement brownien standard.

Kwiatkowski *et al.* (1992)([16]) ont tabulé les valeurs critiques pour les deux statistiques de test η_μ et η_τ .

Chapitre 2

La cointégration linéaire

2.1 Introduction

Le comportement des données économiques en séries temporelles a été étudié de diverses manières à l'aide de différents outils statistiques et économiques. La plupart des données économiques sont collectées sous forme de séries chronologiques et présentent souvent des tendances non stationnaires, ce qui complique leur analyse et peut conduire à des résultats statistiques trompeurs, tels que les régressions fallacieuses. Une régression fallacieuse se produit lorsqu'il semble exister une relation statistiquement significative entre des variables, alors qu'en réalité ces variables ne sont pas liées. Une solution classique pour traiter la non-stationnarité est la différenciation des séries, mais cette transformation peut entraîner une perte d'informations importantes sur les données, ce qui constitue un problème. Pour surmonter cette difficulté, la notion de cointégration a été développée et est aujourd'hui un outil fondamental en économétrie des séries temporelles, introduite par Granger (1981) et développé par Engle et Granger (1987) ([8]), la cointégration permet de détecter et de modéliser des relations d'équilibre à long terme entre des variables non stationnaires. Ce concept repose sur l'idée que plusieurs séries, bien que non stationnaires individuellement, peuvent former une combinaison linéaire stable et stationnaire, traduisant ainsi une relation économique durable. La cointégration offre ainsi un cadre rigoureux pour analyser les interactions dynamiques entre variables économiques, tout en évitant les écueils des analyses classiques appliquées à des données non stationnaires.

Définition 2.1.1. *Deux séries non stationnaires Y_t et X_t , intégrées d'ordre 1 (notées $I(1)$), sont dites **cointégrées** s'il existe une combinaison linéaire de ces variables qui soit stationnaire (i.e., $I(0)$).*

$$Y_t - aX_t - b = \varepsilon_t \sim I(0) \quad (2.1.1)$$

Dans ce cas on note

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

De manière générale, si X_t et Y_t sont deux séries $I(d)$, alors il est possible que la combinaison linéaire

$$\varepsilon_t = Y_t - aX_t - b$$

ne soit pas $I(d)$ mais $I(d - b)$ où b est un entier positif ($0 < b \leq d$). Le vecteur $(1, -a, -b)$ est appelé vecteur de cointégration. Les séries sont alors cointégrées ($X_t, Y_t \sim CI(d, b)$).

Exemple 2.1.1. Soient deux séries temporelles X_t et Y_t telles que :

$$\begin{aligned} X_t &= X_{t-1} + u_t \quad \text{avec} \quad u_t \sim \mathcal{N}(0, 1) \\ Y_t &= 2X_t + 5 + \varepsilon_t \quad \text{avec} \quad \varepsilon_t \sim \mathcal{N}(0, 0.5) \end{aligned}$$

X_t et Y_t sont des marches aléatoire, elle sont intégrées d'ordre 1. Cependant, la combinaison linéaire suivante :

$$Y_t - 2X_t - 5 = \varepsilon_t \sim I(0)$$

est stationnaire, donc X_t et Y_t sont cointégrées de vecteur cointégrant $(1, -2, -5)$.

2.2 Modèle à correction d'erreur

Un modèle à correction d'erreur (MCE) est un modèle économétrique utilisé pour analyser des séries temporelles non stationnaires mais cointégrées, c'est-à-dire liées par une relation d'équilibre de long terme. 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) .

Le modèle à correction d'erreur peut s'écrire sous la forme suivante :

$$\Delta Y_t = \gamma \Delta X_t + \delta(Y_{t-1} - aX_{t-1} - b) + \nu_t \quad (2.2.1)$$

avec

ν_t est un bruit blanc

et $\delta < 0$ pour assurer un retour vers l'équilibre de long terme.

La dynamique de court terme s'écrit :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + \nu_t \quad (2.2.2)$$

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

$$Y_t = aX_t + b + \varepsilon_t \quad (2.2.3)$$

car à long terme on a $Y_{t-1} = Y_t$, $X_{t-1} = X_t$ et la dynamique de court terme devient à long terme,

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_t + \alpha_2 X_t + \alpha_3 X_t + \nu_t \\ (1 - \alpha_1)Y_t &= (\alpha_2 + \alpha_3)X_t + \alpha_0 + \nu_t \end{aligned}$$

On obtient alors,

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

Où,

$$a = \frac{\alpha_2 + \alpha_3}{1 - \alpha_1}, \quad b = \frac{\alpha_0}{1 - \alpha_1}, \quad \varepsilon_t = \frac{\nu_t}{1 - \alpha_1}.$$

Le MCE s'obtient à partir de la dynamique de court terme :

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + \nu_t \\ Y_t - Y_{t-1} &= \alpha_0 + \alpha_1 Y_{t-1} - Y_{t-1} + \alpha_2 X_t - \alpha_2 X_{t-1} + \alpha_3 X_{t-1} + \nu_t \end{aligned}$$

$$\Delta Y_t = (\alpha_1 - 1)Y_{t-1} + \alpha_2(X_t - X_{t-1}) + \alpha_0 + (\alpha_2 + \alpha_3)X_{t-1} + v_t$$

$$\Delta Y_t = (\alpha_1 - 1)Y_{t-1} + \alpha_2(X_t - X_{t-1}) + \alpha_0 + (\alpha_2 + \alpha_3)X_{t-1} + v_t$$

$$\Delta Y_t = -(1 - \alpha_1) \left(Y_{t-1} - \frac{\alpha_2 + \alpha_3}{1 - \alpha_1} X_{t-1} - \frac{\alpha_0}{1 - \alpha_1} \right) + \alpha_2 \Delta X_t + v_t$$

$$\Delta Y_t = \gamma \Delta X_t + \delta (Y_{t-1} - aX_{t-1} - b) + v_t$$

Où,

$$\alpha_2 = \gamma, \quad \delta = -(1 - \alpha_1), \quad a = \frac{\alpha_2 + \alpha_3}{1 - \alpha_1}, \quad b = \frac{\alpha_0}{1 - \alpha_1}$$

2.3 Tests de cointégration

Pour tester l'existence ou non de la cointégration entre séries, la littérature économétrique fournit plusieurs tests ou approches dont le test de Engle et Granger (1987)([8]) ceux de Johansen (1988,1991)([10],[12]) et Johansen et Juselius (1990)([11]) , et celui de Pesaran et al. (1997)([19]) et Pesaran et al. (2001)([20]). Dans ce qui suit, nous allons décrire certains de ces tests usuels.

2.3.1 Test de cointégration de Engle et Granger (1987)

Le test de Engle et Granger (1987)([8]) est une méthode statistique utilisée pour vérifier s'il existe une relation d'équilibre de long terme entre deux séries temporelles non stationnaires mais intégrées du même ordre (généralement $I(1)$), ce test s'effectue en deux étapes.

— **Étape 1** : Vérification de l'ordre d'intégration des deux variables :

Les séries doivent être intégrées de même ordre (par exemple $I(1)$). Il convient donc de vérifier l'ordre d'intégration des chroniques étudiées le test de Dickey-Fuller (simple et augmenté) est recommandé pour cette étape. Si les séries ne sont pas intégrées de même ordre, elles ne peuvent pas être cointégrées, donc la procédure s'arrête à cette première étape.

— **Étape 2** : Estimation de la relation à long terme :

Si les séries sont $I(1)$:

$$X_t \rightarrow I(1), \quad Y_t \rightarrow I(1)$$

La relation de long terme est estimée par les moindres carrés ordinaire (MCO) :

$$Y_t = a X_t + b + \varepsilon_t$$

On extrait les résidus $\hat{\varepsilon}_t$, puis on applique un test de racine unitaire (comme le test de Dickey-Fuller augmenté, ADF) sur les résidus estimés.

Hypothèses de test

- Hypothèse nulle (H_0) : Les résidus $\hat{\varepsilon}_t$ sont non stationnaires (présence d'une racine unitaire) – il n'y a pas de cointégration.
- Hypothèse alternative (H_1) : Les résidus sont stationnaires, donc il existe une relation de cointégration.

Pour que les séries soient cointégrées, alors les résidus $\hat{\varepsilon}_t$ doivent être stationnaires, On a

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

On remarque ici que 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. Il faut regarder ici les tables de MacKinnon.

Si les résidus sont stationnaires, un modèle à correction d'erreur (MCE) peut être établi, intégrant les variables à la fois en niveaux et en variations. Le MCE permet d'obtenir des prévisions plus fiables que la simple relation de long terme, grâce à sa capacité à corriger les effets de non-stationnarité.

2.3.2 Test de cointégration de Johansen (1988)

Le test de cointégration de Johansen (1988) ([10]) est une méthode économétrique utilisée pour déterminer le nombre de relation de cointégration entre plusieurs séries temporelles. Contrairement à l'approche d'Engle et Granger, qui ne permet de vérifier la cointégration qu'entre deux séries. Le test de Johansen peut identifier plusieurs vecteurs de cointégration, ce qui en fait un outil puissant pour l'analyse multivariée.

Avant de procéder au test de cointégration, il est nécessaire de vérifier que les séries temporelles sont intégrées d'ordre 1. Si le test de stationnarité montre que les séries sont intégrées d'un même ordre, il y a alors risque de cointégration. On peut envisager l'estimation d'un modèle vectoriel à correction d'erreur (VECM). Pour ce faire, on commence par déterminer le nombre de retards p du modèle VAR(p) à l'aide des critères d'information (Akaike et Schwarz).

Considérons un vecteur X_t contenant N variables intégrées d'ordre 1 (I(1)). La représentation VAR(p) de X_t est :

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + \varepsilon_t \quad (2.3.1)$$

Exemple 2.3.1. un modèle VAR(2) est composé de deux variables (X_{1t}, X_{2t}) pour X_t avec $p = 2$ retards :

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t \quad (2.3.2)$$

où :

$X_t \in \mathbb{R}^2$ est un vecteur de deux séries temporelles

A_1, A_2 sont des matrices 2×2

ε_t est un bruit blanc multivarié

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} a_{14} & a_{15} \\ a_{24} & a_{25} \end{pmatrix} \begin{pmatrix} x_{1,t-2} \\ x_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

On peut réécrire le modèle VAR(2) sous forme d'un VECM ("Vector Error Correction Model"). On d'abord écrit le modèle VAR(2) en différence première.

On soustrait X_{t-1} des deux côtés :

$$\begin{aligned} X_t - X_{t-1} &= A_1 X_{t-1} + A_2 X_{t-2} - X_{t-1} + \varepsilon_t \\ \Delta X_t &= (A_1 - I) X_{t-1} + A_2 X_{t-2} + \varepsilon_t \end{aligned}$$

Puis on exprime X_{t-2} en fonction de ΔX_{t-1} ,

$$X_{t-2} = X_{t-1} - \Delta X_{t-1}$$

Donc

$$\begin{aligned} \Delta X_t &= (A_1 - I) X_{t-1} + A_2 (X_{t-1} - \Delta X_{t-1}) + \varepsilon_t \\ &= (A_1 + A_2 - I) X_{t-1} - A_2 \Delta X_{t-1} + \varepsilon_t \end{aligned}$$

De manière générale, si on a la représentation VAR(p) suivante pour X_t :

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + \varepsilon_t \quad (2.3.3)$$

le modèle VECM va s'écrire comme suit

$$\Delta X_t = B_1 \Delta X_{t-1} + \dots + B_{p-1} \Delta X_{t-p+1} + \Pi X_{t-1} + \varepsilon_t \quad (2.3.4)$$

où $B_i = \sum_{j=i+1}^p -A_j$ avec $i = 1, \dots, k-1$ et $\Pi = A_1 + \dots + A_k - I$.

On pose $\Pi = \alpha \beta'$ avec α une matrice (N, r) avec $r < N$ contenant les vitesses d'ajustement pour chacun des vecteurs de cointégration et β' une matrice (r, N) comprenant les r relations de cointégration. Donc pour pouvoir estimer un modèle VECM, il faut que $Rg(\Pi) = Rg(\alpha \beta') = r$ ce qui implique que Π a "r" valeurs propres non nulles.

Or $Rg(\Pi) = Rg(\alpha \beta') \leq \min(Rg(\alpha), Rg(\beta'))$. Donc il faut que $Rg(\beta') = r$ pour que $Rg(\Pi) = r$. Donc il faut que β' possède r valeurs propres non nulles.

Trois cas peuvent se présenter :

- $Rg(\Pi) = 0$, donc $r = 0$, il n'existe pas de relation de cointégration. On ne peut pas estimer un modèle VECM. En revanche, il est possible d'estimer un modèle VAR sur ΔX_t .
- $Rg(\Pi) = r$, il existe r relations de cointégration. Un modèle VECM peut alors être estimé.
- $Rg(\Pi) = N$, il n'existe pas de relation de cointégration. Un modèle VAR peut être estimé directement sur X_t .

Pour déterminer le nombre de relations r de cointégration, Johansen (1988)([10]) se base sur la maximisation de la log-vraisemblance suivante :

$$\log L(\alpha, \beta, B_1, \dots, B_{p-1}, \Sigma) = -\frac{NT}{2} \log(2\pi) - \frac{T}{2} \log[\det(\Sigma)] - \frac{1}{2} \sum_{t=1}^T \varepsilon_t' \Sigma^{-1} \varepsilon_t \quad (2.3.5)$$

où T est le nombre d'observations, N le nombre de variables dans X et $\det(\Sigma)$ représente le déterminant de la matrice de variance-covariance.

Pour obtenir l'estimateur du maximum de vraisemblance de β , il faut résoudre l'équation suivante :

$$\text{Det} [\lambda S_{pp} - S_{p0} S_{00}^{-1} S_{0p}] = 0 \quad (2.3.6)$$

où : $S_{ij} = \frac{1}{T} \sum_{t=1}^T e_{it} e'_{ij}$ pour $i, j = 0, p$. Les termes e_{0t} et e_{pt} sont les résidus des deux régressions suivantes :

- **Première régression** : on régresse ΔX_t (la variation de X_t) sur les variables $\Delta X_{t-1}, \dots, \Delta X_{t-p+1}$. Cela donne les résidus e_{0t} .
- **Deuxième régression** : on régresse X_{t-p} (le niveau de X décalé de p) sur les mêmes variables $\Delta X_{t-1}, \dots, \Delta X_{t-p+1}$. Cela donne les résidus e_{pt} .

La résolution de l'équation (2.3.6) nous donne les N valeurs propres $\hat{\lambda}_1, \dots, \hat{\lambda}_N$ avec $\hat{\lambda}_1 > \dots > \hat{\lambda}_N$ et les N vecteurs propres associés. Pour déterminer les r valeurs propres non nulles qui vont nous donner les r relations de cointégration, Johansen a proposé deux tests.

1. Test de la trace

Ce test repose sur l'hypothèse nulle qu'il existe au plus r relations de cointégration (ce qui signifie qu'il y a r valeurs propres non nulles et $N - r$ valeurs propres nulles) :

$$\begin{aligned} & H_0 : r = 0 \quad \text{contre} \quad H_1 : r > 0 \\ \text{Si l'on rejette } H_0, \text{ alors : } & H_0 : r = 1 \quad \text{contre} \quad H_1 : r > 1 \\ \text{Si l'on rejette encore } H_0, \text{ alors : } & H_0 : r = 2 \quad \text{contre} \quad H_1 : r > 2 \\ & \vdots \\ & \text{jusqu'à } H_0 : r = N - 1 \quad \text{contre} \quad H_1 : r = N \end{aligned}$$

La statistique de test est la suivante :

$$\text{Trace}(r) = -T \sum_{i=r+1}^k \ln(1 - \lambda_i)$$

2. **Test de la valeur propre maximale (maximum eigenvalue test)** Ce test permet de tester l'hypothèse nulle d'existence de r relations de cointégration contre l'hypothèse alternative de $r + 1$ relations de cointégration. Autrement dit :

$$\begin{aligned} & H_0 : r = 0 \quad (\text{aucune relation de cointégration}) \\ & \quad \text{contre} \quad H_1 : r > 0 \quad (\text{au moins une relation de cointégration}) \\ & \text{Si on rejette } H_0, \text{ on passe à l'hypothèse suivante :} \\ & H_0 : r = 1 \quad (\text{une relation de cointégration}) \\ & \quad \text{contre} \quad H_1 : r > 1 \quad (\text{au moins deux relations}) \\ & \text{Si on rejette encore } H_0, \text{ on continue :} \\ & H_0 : r = 2 \quad (\text{il existe deux relations de cointégration}) \\ & \quad \text{contre} \quad H_1 : r > 2 \\ & \quad \vdots \\ & H_0 : r = k - 1 \quad \text{contre} \quad H_1 : r = k \end{aligned}$$

La statistique de test est la suivante :

$$\text{Max}(r, r + 1) = -T \ln(1 - \lambda_{r+1})$$

Les valeurs critiques de ces tests sont tabulées par Johansen et Juselius (1990)([11]) et par Osterwald-Lenum (1992)([18]). On accepte H_0 lorsque la valeur de la statistique de test est inférieure à sa valeur critique.

2.3.3 Test de cointégration de Phillips-Ouliaris (1990)

Le test de Phillips-Ouliaris (1990)([22]) est un test statistique utilisé en économétrie pour détecter la cointégration entre deux ou plusieurs séries temporelles non stationnaires. Il repose sur l'analyse des résidus d'une régression de long terme entre ces séries, et teste la stationnarité de ces résidus. Contrairement au test d'Engle-Granger, il corrige les problèmes d'autocorrélation et d'hétéroscédasticité des résidus en utilisant des estimateurs robustes de la variance à long terme, sans exiger la spécification préalable d'un nombre de retards dans le modèle.

Voici les étapes principales pour réaliser le test de Phillips-Ouliaris.

- **Étape 1** : On vérifie que toutes les séries sont intégrées du même ordre $I(1)$ à l'aide du test de Dickey-Fuller augmenté (ADF)
- **Étape 2** : Estimation de la relation à long terme :
On estime la relation d'équilibre de long terme entre les séries par une régression en moindres carrés ordinaires (OLS), par exemple :

$$Y_t = \alpha + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t \quad (2.3.7)$$

où ε_t est les résidus de la régression

On extrait les résidus $\hat{\varepsilon}_t$, puis on applique un test de racine unitaire pour vérifier si les résidus sont stationnaires ou non.

Hypothèses de test

- **Hypothèse nulle** (H_0) : Les résidus $\hat{\varepsilon}_t$ sont non stationnaires (présence d'une racine unitaire) il n'y a pas de cointégration.
- **Hypothèse alternative** (H_1) : Les résidus sont stationnaires, donc il existe une relation de cointégration.

Remarque 2.3.1. *Le test de Phillips-Ouliaris n'utilise pas les tests ADF pour vérifier la stationnarité des résidus. En effet, Phillips et Ouliaris ont développé leurs propres tests, spécifiquement adaptés à ce contexte.*

Les statistiques de test

Le test de Phillips-Ouliaris propose deux statistiques principales, notées Z_α et Z_t .

1. La statistique Z_α est définie par :

$$Z_\alpha = T(\hat{\rho} - 1) - \frac{1}{2} \cdot \frac{\hat{\sigma}_\eta^2 - \hat{s}_\eta^2}{\frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{t-1}^2} \quad (2.3.8)$$

où :

T : nombre d'observations

$\hat{\rho}$: estimateur MCO de l'autorégression d'ordre 1 des résidus

$$\hat{\varepsilon}_t = \hat{\rho}\hat{\varepsilon}_{t-1} + u_t$$

$\hat{\sigma}_\eta^2$: estimateur de la variance de long terme de u_t

\hat{s}_η^2 : estimateur de la variance instantanée de u_t

2. La statistique Z_t est définie par :

$$Z_t = \sqrt{\frac{\hat{s}_\eta^2}{\hat{\sigma}_\eta^2}} \cdot t_{\hat{\rho}} - \frac{1}{2 \cdot \hat{\sigma}_\eta \sqrt{T}} \cdot \frac{\hat{\sigma}_\eta^2 - \hat{s}_\eta^2}{\frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_{t-1}^2} \quad (2.3.9)$$

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

Remarque 2.3.2. *Les statistiques sont comparées aux valeurs critiques tabulées dans Phillips & Ouliaris (1990) ([22]). Ces tables tiennent compte du fait que les résidus proviennent d'une régression entre séries intégrées d'ordre un, $I(1)$.*

Elles dépendent notamment :

- du nombre de variables explicatives
- de la présence d'une constante ou d'une tendance dans la régression.

Interprétation

- Si la statistique de test est inférieure à la valeur critique, alors on rejette l'hypothèse nulle H_0 :
 \Rightarrow Cointégration : les résidus sont stationnaires.
- Si la statistique de test est supérieure à la valeur critique (ou si la p-valeur est élevée), alors on ne rejette pas l'hypothèse nulle H_0 :
 \Rightarrow Pas de cointégration : les résidus ont une racine unitaire (non stationnaires).

2.4 Comparaison des tests de cointégration

| Critère | Engle-Granger | Johansen | Phillips-Ouliaris |
|----------------------------|-----------------------------|------------------------------|---------------------------------|
| Nombre de variables | 2 seulement | Multiplés ($n \geq 2$) | Multiplés |
| Approche | Univariée (résidus) | Multivariée (rang de Π) | Univariée/multivariée (résidus) |
| Robustesse | Sensible à la spécification | Robuste (VAR structurel) | Correction autocorrélation |
| Usage | Simple, rapide | Systèmes complexes | Alternative à Engle-Granger |

TABLE 2.1 – Comparaison des tests de cointégration : Engle-Granger, Johansen et Phillips-Ouliaris

Ce tableau compare les principales caractéristiques des trois tests de cointégration les plus utilisés : Engle-Granger, Johansen et Phillips-Ouliaris.

Le test d'Engle-Granger est simple et adapté à l'analyse de deux variables, basé sur une approche univariée des résidus, mais il est sensible à la spécification du modèle.

Le test de Johansen, quant à lui, est plus robuste et permet d'étudier des systèmes multivariés complexes en s'appuyant sur la structure du vecteur d'erreur-correction.

Enfin, le test de Phillips-Ouliaris constitue une alternative qui corrige l'autocorrélation dans les résidus et peut être appliqué aussi bien en contexte univarié que multivarié.

Le choix du test dépend donc du nombre de variables considérées, de la complexité du modèle et de la précision souhaitée.

Chapitre 3

Les modèles Auto Regressive Distributed Lag (ARDL)

3.1 Introduction

L'émergence des modèles ARDL (Autoregressive Distributed Lag) s'inscrit dans une quête méthodologique visant à surmonter les limites des techniques traditionnelles d'analyse des séries temporelles, notamment en matière de cointégration. Dès les années 1930, les modèles autorégressifs (AR) ont permis d'exploiter les valeurs passées d'une variable pour expliquer son évolution présente, mais ils ne prenaient pas en compte les effets différés des variables explicatives ni les relations d'équilibre à long terme entre séries non stationnaires. Dans les années 1970, les économistes Davidson, Hendry, Srba et Yeo (1978) ont introduit le modèle à correction d'erreur (ECM), qui intègre la notion d'ajustement vers un équilibre de long terme, marquant une avancée majeure dans la modélisation dynamique. Cependant, les méthodes classiques de cointégration, comme celles de Engle et Granger (1987) ([8]) ou Johansen (1988) ([10]), imposaient une contrainte forte : toutes les séries temporelles étudiées devaient être intégrées du même ordre, généralement $I(1)$. Or, dans la pratique, les données économiques ne respectent pas toujours cette condition, certaines séries pouvant être stationnaires en niveau ($I(0)$) et d'autres en première différence ($I(1)$). Pour répondre à cette problématique, Pesaran, Shin et Smith (2001) ([20]) ont développé le modèle ARDL et le test de cointégration aux bornes, qui permettent d'estimer des relations de long terme entre variables intégrées d'ordres différents ($I(0)$ ou $I(1)$) sans nécessiter un pré-test rigide d'intégration uniforme. Cette flexibilité méthodologique est essentielle, car elle évite les erreurs liées à une mauvaise classification des séries et permet d'utiliser efficacement des petits échantillons, fréquents en économétrie appliquée. Le modèle ARDL fournit ainsi une solution robuste aux limites des méthodes antérieures, en offrant des estimations non biaisées dans les analyses à court et long terme, tout en s'adaptant aux caractéristiques réelles des données économiques.

Définition 3.1.1. *Les modèles "Auto Regressive Distributed Lag" ou "modèles autorégressifs à retards échelonnés distribués", sont des modèles dynamiques qui intègrent la dynamique temporelle pour améliorer les prévisions et l'efficacité des politiques. Ces modèles permettent d'analyser à la fois les effets à court terme et à long terme des variables sur la variable dépendante. Le modèle ARDL combine les caractéristiques de deux modèles :*

• **Les modèles autorégressifs (AR) :** ce sont des modèles où l'on trouve parmi les variables explicatives X_t les valeurs passées de la variable à expliquer Y_t . Ils se représentent comme suit (forme explicite) :

$$\mathbf{Y}_t = f(\mathbf{X}_t, \mathbf{Y}_{t-p}) \quad (3.1.1)$$

• **Les modèles à retards échelonnés (DL)** : ce sont des modèles dont les variables explicatives sont X_t et ses valeurs passées. Leur forme explicite est :

$$\mathbf{Y}_t = f(\mathbf{X}_t, \mathbf{X}_{t-q}) \quad (3.1.2)$$

Ainsi, les modèles autorégressifs à retards échelonnés (ARDL) sont des modèles où l'on trouve, parmi les variables explicatives X_t , la variable dépendante retardée Y_{t-p} ainsi que les valeurs passées de la variable exogène X_{t-q} . Leur forme explicite est :

$$\mathbf{Y}_t = f(\mathbf{X}_t, \mathbf{Y}_{t-p}, \mathbf{X}_{t-q}) \quad (3.1.3)$$

De manière plus développée, le modèle ARDL s'écrit classiquement comme suit :

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=0}^q \gamma_j X_{t-j} + \varepsilon_t \quad (3.1.4)$$

Ou encore :

$$Y_t = \alpha + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \gamma_0 X_t + \dots + \gamma_q X_{t-q} + \varepsilon_t \quad (3.1.5)$$

Où Y_t , X_t et ε_t désignent respectivement la variable endogène (dépendante), la variable exogène (indépendante), et le terme d'erreur.

▷ **Forme générale du modèle ARDL (multi-exogène) :**

$$Y_t = c_0 + c_1 t + \sum_{i=1}^p b_{y,i} Y_{t-i} + \sum_{j=1}^k \sum_{l=0}^{q_j} b_{j,l} x_{j,t-l} + \varepsilon_t \quad (3.1.6)$$

Où

K représente le nombre total de variables explicatives exogènes dans le modèle ARDL.

c_0 est la constante ;

$c_1 t$ est une tendance temporelle ;

$b_{y,i}$ sont les coefficients autorégressifs de Y_t ;

$x_{j,t-l}$ représente la j -ème variable explicative avec l -ème retard ;

$b_{j,l}$ est le coefficient associé à $x_{j,t-l}$;

ε_t est le terme d'erreur aléatoire.

Exemple 3.1.1. L'écriture du modèle ARDL(3, 1, 3, 2)

$$Y_t = c_0 + c_1 t + \sum_{i=1}^3 b_{Y_i} Y_{t-i} + \sum_{j=1}^3 \sum_{l=0}^{q_j} b_{j,l} X_{j,t-l} + \varepsilon_t \quad (3.1.7)$$

$$Y_t = c_0 + c_1 t + b_{Y_1} Y_{t-1} + b_{Y_2} Y_{t-2} + b_{Y_3} Y_{t-3} + \sum_{l=0}^{q_1} b_{1,l} X_{1,t-l} + \sum_{l=0}^{q_2} b_{2,l} X_{2,t-l} + \sum_{l=0}^{q_3} b_{3,l} X_{3,t-l} + \varepsilon_t \quad (3.1.8)$$

$$Y_t = c_0 + c_1 t + b_{Y_1} Y_{t-1} + b_{Y_2} Y_{t-2} + b_{Y_3} Y_{t-3} + \sum_{l=0}^1 b_{1,l} X_{1,t-l} + \sum_{l=0}^3 b_{2,l} X_{2,t-l} + \sum_{l=0}^2 b_{3,l} X_{3,t-l} + \varepsilon_t \quad (3.1.9)$$

$$Y_t = c_0 + c_1 t + b_{Y_1} Y_{t-1} + b_{Y_2} Y_{t-2} + b_{Y_3} Y_{t-3} + b_{10} X_{1,t} + b_{11} X_{1,t-1} + b_{20} X_{2,t} + b_{21} X_{2,t-1} + b_{22} X_{2,t-2} + b_{23} X_{2,t-3} + b_{30} X_{3,t} + b_{31} X_{3,t-1} + b_{32} X_{3,t-2} + \varepsilon_t \quad (3.1.10)$$

Remarque 3.1.1. — *Variables endogènes* : ce sont des variables dont la valeur est déterminée par le modèle.

— *Variables exogènes* : ce sont les variables du modèle qu'on considère comme déterminées à l'extérieur du modèle.

3.1.1 Propriétés du modèle ARDL

- Les modèles ARDL comme tous les modèles dynamiques utilisent des critères d'informations (AIC , SIC et HQ) pour déterminer les décalages optimaux.
- Le modèle ARDL permet d'examiner à la fois la dynamique de court terme et les effets à long terme d'une ou plusieurs variables explicatives sur une variable à expliquer.
- Les modèles ARDL souffrent généralement de problèmes d'autocorrélation d'erreurs et de multi-colinéarité , ce qui complique l'estimation des paramètres par les moindres carrés ordinaires.
- Le modèle ARDL peut être appliqué lorsque les variables sont stationnaires en niveau $I(0)$, en première différence $I(1)$, ou une combinaison des deux, sans nécessiter que toutes soient intégrées au même ordre. En revanche, il ne doit pas y avoir de variables intégrées d'ordre 2 ou plus $I(2)$.
- Les variables dans les modèles ARDL doivent être stationnaires pour éviter des régression fallacieuses.
- Le modèle ARDL est souvent associé à l'approche de test aux bornes pour vérifier la cointégration.

3.1.2 Critères de sélection de modèle

Pour déterminer l'ordre optimal (p ou q) du modèle, on utilise des critères d'information qui équilibrent la qualité d'ajustement et la parcimonie. Les principaux critères sont :

1. Critère d'Information d'Akaike (AIC)

Le critère AIC vise à maximiser la vraisemblance tout en pénalisant la complexité du modèle :

$$AIC(p) = \ln |\hat{\Sigma}_p| + \frac{2}{T} \cdot (n \cdot p) \quad (3.1.11)$$

où :

$\hat{\Sigma}_p$ est la matrice de variance-covariance des résidus

T est le nombre d'observations

n est le nombre de variables endogènes

p est l'ordre du retard (lag)

2. Critère d'Information Bayésien (BIC)

Également appelé SIC (Schwarz Information Criterion), le BIC impose une pénalité plus forte pour la complexité :

$$BIC(p) = \ln |\hat{\Sigma}_p| + \frac{\ln T}{T} \cdot (n \cdot p) \quad (3.1.12)$$

Comparaison des critères

| Critère | Pénalité | Usage recommandé |
|---------|-------------------------------------|-------------------------------|
| AIC | $\frac{2}{T} \cdot (n \cdot p)$ | Prévision à court terme |
| BIC | $\frac{\ln T}{T} \cdot (n \cdot p)$ | Identification du vrai modèle |

TABLE 3.1 – Comparaison des critères AIC et BIC

- Le modèle optimal est celui qui minimise le critère choisi
- Le BIC tend à sélectionner des modèles plus parcimonieux que l'AIC
- Pour $T \geq 8$, la pénalité du BIC est plus sévère

3.2 L'approche du test aux bornes ou test de cointégration de Pesaran et al. (2001)

Pour vérifier l'existence d'une relation de cointégration entre plusieurs séries temporelles, la littérature économétrique propose plusieurs tests, parmi lesquels les tests d'Engle et Granger, et ceux de Johansen et Juselius. Le test d'Engle et Granger est adapté au cas bivarié, c'est-à-dire pour deux séries intégrées d'ordre 1, mais il devient moins efficace lorsque le nombre de variables augmente. Le test de Johansen, quant à lui, permet d'étudier la cointégration dans un cadre multivarié grâce à un modèle vectoriel à correction d'erreur (VECM), mais il requiert que toutes les séries soient intégrées du même ordre, généralement d'ordre 1. Cependant, en pratique, il arrive fréquemment que les séries temporelles ne soient pas toutes intégrées du même ordre. Pour pallier cette limitation, Pesaran, Shin et Smith (2001) ([20]) ont développé le test de cointégration aux bornes (ou *bounds test*), fondé sur le modèle auto-régressif à retards échelonnés (ARDL). Ce test présente l'avantage de pouvoir être appliqué à des variables intégrées d'ordre zéro $I(0)$ ou d'ordre un $I(1)$, sans exiger qu'elles soient toutes du même ordre, à condition qu'aucune ne soit intégrée d'ordre supérieur.

Principe du test ARDL bounds

Le test repose sur la spécification suivante d'un modèle ARDL (p, q) :

$$\Delta Y_t = \alpha_0 + \pi_y Y_{t-1} + \pi_x X_{t-1} + \sum_{i=1}^{p-1} \psi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \gamma_j \Delta X_{t-j} + \varepsilon_t \quad (3.2.1)$$

où :

y_t est la variable dépendante,

x_t la ou les variables explicatives,

π_y, π_x sont les coefficients de long terme,

ψ_i, γ_j sont les coefficients de court terme,

ε_t est le terme d'erreur.

Hypothèses du test

- **Hypothèse nulle** (H_0) : absence de cointégration, soit

$$\pi_y = \pi_x = 0$$

- **Hypothèse alternative** (H_1) : présence de cointégration, c'est-à-dire

$$\pi_y \neq 0 \quad \text{ou} \quad \pi_x \neq 0$$

La statistique de test

Dans le cadre du test ARDL Bounds deux types de statistiques peuvent être utilisés pour tester l'existence d'une relation de cointégration entre les variables :

1. Test F (ou Bounds Test)

Ce test vérifie la significativité conjointe des coefficients des variables retardées au niveau (π_x et π_y).

La statistique F est calculée à partir de la formule suivante :

$$F = \frac{(SCR_0 - SCR_1)/m}{SCR_1/(T - k)}$$

où

SCR_0 somme des carrés des résidus du modèle contraint (sans les termes au niveau)

SCR_1 somme des carrés des résidus du modèle non contraint (avec les termes au niveau)

m nombre de restrictions (ici, $m = 2$, car on teste $\pi_y = \pi_x = 0$)

T nombre total d'observations

k nombre total de paramètres dans le modèle non contraint

2. Test t

Le test t porte uniquement sur le coefficient π_y , associé à la variable dépendante retardée y_{t-1} . Il est utile lorsque x_t est supposée exogène. Ce test est plus limité que le test F, mais peut servir de test complémentaire ou de test unilatéral.

Distribution et Régions critiques

Le test ARDL Bounds repose sur une distribution *non standard* de la statistique F , influencée par le nombre de variables explicatives, leur ordre d'intégration ($I(0)$ ou $I(1)$), ainsi que par la présence de termes déterministes (constante, tendance). Contrairement aux tests classiques fondés sur la loi de Fisher-Snedecor, la présence simultanée de variables stationnaires ($I(0)$) et non stationnaires ($I(1)$) rend les tables statistiques standards inapplicables.

Pour résoudre ce problème, Pesaran, Shin et Smith (2001)([20]) ont simulé par la méthode de Monte Carlo les distributions de la statistique F pour différents cas. Ils ont ensuite tabulé deux bornes critiques pour chaque niveau de signification :

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

Ces bornes permettent de couvrir tous les cas possibles, y compris les situations mixtes où certaines variables sont $I(0)$ et d'autres $I(1)$.

Règle de décision

La décision dépend de la comparaison entre la valeur calculée de la statistique F et les deux bornes critiques :

- Si $F_{calculée} > \text{Borne supérieure} \Rightarrow$ Rejet de H_0 (cointégration)
- Si $F_{calculée} < \text{Borne inférieure} \Rightarrow$ Non-rejet de H_0 (pas de cointégration)
- Si $\text{Borne inférieure} < F_{calculée} < \text{Borne supérieure} \Rightarrow$ Test **non conclusif** (analyse approfondie nécessaire)

Exemple de valeurs critiques (cas avec constante, sans tendance)

| Niveau de signification | Borne inférieure (I(0)) | Borne supérieure (I(1)) |
|-------------------------|-------------------------|-------------------------|
| 1% | 4,94 | 5,58 |
| 5% | 3,62 | 4,16 |
| 10% | 3,02 | 3,51 |

(Ces valeurs dépendent du nombre de régressions et de la spécification du modèle.)

Résumé des régions critiques

- **Région de rejet** : $F >$ borne supérieure.
- **Région d'acceptation** : $F <$ borne inférieure.
- **Zone grise (non conclusive)** : F entre les deux bornes.

Remarque importante :
 En cas de résultat non conclusif, il faut :

- Vérifier l'ordre d'intégration des variables via des tests ADF/PP
- Augmenter la taille de l'échantillon si possible
- Considérer d'autres spécifications du modèle (ajout/suppression de tendances)

Étapes pratiques du test ARDL Bounds

Étape 1 : Estimation du modèle

• **Modèle non contraint (Unrestricted Model)** : Estimer le modèle suivant et obtenir la somme des carrés des résidus (SCR_0).

$$\Delta Y_t = \alpha_0 + \pi_y Y_{t-1} + \pi_x X_{t-1} + \sum_{i=1}^{p-1} \psi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \gamma_j \Delta X_{t-j} + \varepsilon_t$$

- **Modèle contraint (sous H_0) :**

$$\Delta Y_t = \alpha_0 + \sum_{i=1}^{p-1} \psi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \gamma_j \Delta X_{t-j} + u_t$$

Estimer ce modèle par MCO et obtenir la somme des carrés des résidus (SCR_0).

Étape 2 : Calcul de la statistique F

$$F = \frac{(SCR_0 - SCR_1)/m}{SCR_1/(T - k)}$$

Étape 3 : Interprétation des résultats

- Si $F >$ Borne supérieure \Rightarrow **Cointégration**
- Si $F <$ Borne inférieure \Rightarrow **Pas de cointégration**
- Si Borne inf $< F <$ Borne sup \Rightarrow **Test non concluant**

Conditions d'utilisation du test F dans le cadre ARDL

Pour que les résultats du test ARDL Bounds soient valides et fiables, certaines conditions doivent être respectées :

- **Ordre d'intégration des variables :** Les variables du modèle doivent être intégrées d'ordre zéro ($I(0)$) ou d'ordre un ($I(1)$). Aucune variable ne doit être intégrée d'ordre deux ($I(2)$), sinon les tables de Pesaran ne sont plus valides.
- **Absence d'autocorrélation des résidus :** Les résidus du modèle estimé ne doivent pas présenter d'autocorrélation. Cela se vérifie à l'aide de tests comme le test de Breusch-Godfrey ou le test de Ljung-Box.
- **Homoscédasticité des résidus :** La variance des résidus doit être constante (homoscédasticité). On peut le vérifier avec le test de Breusch-Pagan.
- **Normalité des résidus :** Les résidus doivent suivre une distribution normale, surtout si la taille de l'échantillon est faible (test de Jarque-Bera ou de Shapiro-Wilk).
- **Stabilité des coefficients :** Les paramètres du modèle doivent être stables dans le temps (tests de stabilité comme CUSUM ou CUSUMSQ).
- **Taille de l'échantillon suffisante :** Les résultats du test ARDL Bounds sont plus fiables avec un échantillon de taille raisonnable (généralement, au moins 30 à 50 observations).

Remarque 3.2.1. *Si ces conditions ne sont pas respectées, il est conseillé de corriger le modèle (par exemple, en transformant les variables, en ajustant la spécification, ou en utilisant des méthodes robustes) avant d'interpréter les résultats du test.*

Exemple 3.2.1. *Supposons les valeurs suivantes :*

$$SCR_0 = 50, \quad SCR_1 = 40, \quad T = 100, \quad k = 5 \quad (\text{incluant constante et retards}), \quad m = 2 \quad (\text{restrictions sur } \pi_y \text{ et } \pi_x).$$

La statistique F est donnée par :

$$F = \frac{(SCR_0 - SCR_1)/m}{SCR_1/(T - k)} = \frac{(50 - 40)/2}{40/(100 - 5)} = \frac{10/2}{40/95}$$

$$F = \frac{5}{0,421} \approx 11,88$$

Si la borne supérieure à 5% est 4,94, alors :

$$F = 11,88 > 4,94$$

Donc, on rejette l'hypothèse nulle H_0 : une relation de cointégration est détectée.

3.3 Estimation dans les modèles ARDL

L'estimation dans les modèles ARDL (AutoRegressive Distributed Lag) est une méthode puissante pour analyser les relations dynamiques entre variables économiques, notamment lorsque ces variables sont intégrées d'ordres différents, c'est-à-dire $I(0)$ et/ou $I(1)$. La procédure d'estimation d'un modèle ARDL suit plusieurs étapes essentielles afin d'assurer la validité statistique des résultats et leur interprétation correcte. Ces étapes comprennent :

Étape 1 : Vérification de l'ordre d'intégration des variables

Avant d'estimer le modèle ARDL, il est indispensable de tester le degré d'intégration des variables à l'aide de tests de racine unitaire (ADF, PP, KPSS, etc.). Ces tests permettent de déterminer si les séries temporelles Y_t et X_t sont $I(0)$ ou $I(1)$.

Remarque 3.3.1. *Il est important de souligner que si une variable est intégrée d'ordre deux $I(2)$ ou plus, alors l'utilisation du modèle ARDL n'est pas appropriée, car ce modèle requiert que les variables soient au maximum intégrées d'ordre un $I(1)$ afin de garantir la validité des tests de cointégration et des estimations.*

Étape 2 : Sélection des ordres de retard (p, q)

Il est ensuite nécessaire de déterminer les ordres optimaux de retard pour (p et q), Cette sélection peut se faire à l'aide de critères d'information tels que :

- Critère d'information d'Akaike (AIC)
- Critère de Schwarz (BIC)
- Critère de Hannan-Quinn (HQ) : ce critère est un compromis entre AIC et BIC.

Une mauvaise spécification des retards peut entraîner une perte d'information ou une sur-paramétrisation du modèle.

Étape 3 : Estimation du modèle ARDL

Une fois les ordres de retard déterminés, le modèle ARDL peut être estimé par plusieurs méthodes, notamment les moindres carrés ordinaires (MCO) et la méthode du maximum de vraisemblance.

Étape 4 : Reformulation du modèle en ECM (Erreur Correction)

Une fois le modèle ARDL estimé, il peut être reformulé sous la forme d'un modèle à correction d'erreur (ECM). Cette étape est nécessaire car le test aux bornes de Pesaran et al. (2001)([20]) repose sur la version ECM du modèle. Le modèle ECM permet d'évaluer conjointement les relations de court terme (à travers les différences) et de long terme (à travers les niveaux retardés) :

$$\Delta Y_t = \alpha_0 + \pi_y Y_{t-1} + \pi_x X_{t-1} + \sum_{i=1}^{p-1} \psi_i \Delta Y_{t-i} + \sum_{j=0}^{q-1} \delta_j \Delta X_{t-j} + \varepsilon_t$$

Où

- π_y et π_x : Coefficients de long terme.
- ψ_i et δ_j : Coefficients de court terme.

Étape 5 : Application du test de cointégration aux bornes

Le test aux bornes s'applique à la forme ECM du modèle. Il teste l'hypothèse nulle selon laquelle il n'existe pas de relation de long terme entre les variables :

$$H_0 : \pi_y = \pi_x = 0$$

La statistique F issue du test est comparée aux seuils critiques fournis par Pesaran et al. (2001)([20]) :

- Si $F < I(0)$: pas de cointégration ;
- Si $F > I(1)$: cointégration confirmée ;
- Si $I(0) < F < I(1)$: la conclusion est incertaine (zone d'indécision).

Étape 6 : Validation des résidus

Après l'estimation du modèle ARDL, il est crucial de vérifier la validité des hypothèses classiques sur les résidus pour garantir la qualité des estimations.

Les résidus doivent vérifier :

1. Stationnarité (test ADF).
2. Absence d'autocorrélation (Test de Breusch-Godfrey).
3. Homoscédasticité (Test de Breusch-Pagan ou test de White).
4. Normalité des résidus (Test de Jarque-Bera ou Shapiro-Wilk).

Étape 7 : Interprétation des résultats

Une fois le modèle estimé et la cointégration confirmée, il est possible d'interpréter économiquement les coefficients du modèle sous sa forme à correction d'erreur (ECM). Cette interprétation se fait à trois niveaux :

Effets de court terme

Les effets de court terme sont représentés par les coefficients des différences premières des variables. Les effets de court terme sont représentés par les coefficients des différences premières des variables dans le modèle à correction d'erreur (ECM). Ces coefficients capturent l'impact immédiat ou différé des variations des variables explicatives sur la variable dépendante.

- δ_j : coefficients associés aux termes ΔX_{t-j} , représentant l'effet immédiat ou différé d'une variation de la variable explicative X sur la variable dépendante Y .
- ψ_i : coefficients associés aux termes ΔY_{t-i} , reflétant la dynamique propre de la variable dépendante.

Effets de long terme

Les coefficients de long terme sont obtenus à partir des coefficients des variables de niveau dans le modèle ECM.

$$\text{Long-term coefficient} = \frac{\sum_{j=0}^q \beta_j}{1 - \sum_{i=1}^p \phi_i}$$

Vitesse d'ajustement vers l'équilibre

Le coefficient du terme d'erreur Y_{t-1} , noté généralement π_y , mesure la vitesse à laquelle le système retourne vers son équilibre de long terme après un choc.

- Ce coefficient doit être négatif et significatif.
- Plus sa valeur absolue est élevée plus le retour à l'équilibre est rapide.

3.4 Méthodes d'estimation du modèle ARDL

Après avoir présenté les différentes étapes d'estimation du modèle ARDL, il est utile de mentionner les deux principales méthodes utilisées pour estimer les coefficients du modèle. Ces méthodes sont :

3.4.1 Estimation du modèle ARDL par la méthode des moindres carrés ordinaires (MCO)

L'estimation des modèles ARDL (AutoRegressive Distributed Lag) par la méthode des moindres carrés ordinaires (MCO) constitue une approche fondamentale en économétrie des séries temporelles. Cette méthode vise à déterminer les paramètres du modèle en minimisant la somme des carrés des résidus (SCR). Voici la formulation mathématique et les étapes de calcul.

Formulation générale du modèle ARDL (p, q)

On considère un modèle ARDL(p, q) sous la forme :

$$Y_t = \alpha_0 + \sum_{i=1}^p \delta_i Y_{t-i} + \sum_{j=0}^q \beta_j X_{t-j} + \varepsilon_t \quad (3.4.1)$$

L'objectif de la méthode MCO est de minimiser la somme des carrés des résidus (SCR) suivante :

$$\text{SCR} = \sum_{t=p+1}^T \left(Y_t - \alpha_0 - \sum_{i=1}^p \delta_i Y_{t-i} - \sum_{j=0}^q \beta_j X_{t-j} \right)^2 \quad (3.4.2)$$

Formules explicites des coefficients

On minimise SCR par rapport aux paramètres du modèle $\theta = (\alpha_0, \delta_1, \dots, \delta_p, \beta_0, \dots, \beta_q)$. Pour cela, on dérive la fonction SCR par rapport à chaque paramètre, puis on résout le système d'équations obtenu :

$$\begin{aligned} \frac{\partial \text{SCR}}{\partial \alpha_0} &= -2 \sum_t \alpha_0 \left(Y_t - \alpha_0 - \sum_{i=1}^p \phi_i Y_{t-i} - \sum_{j=0}^q \beta_j X_{t-j} \right) = 0 \\ \frac{\partial \text{SCR}}{\partial \phi_k} &= -2 \sum_t Y_{t-k} \left(Y_t - \alpha_0 - \sum_{i=1}^p \phi_i Y_{t-i} - \sum_{j=0}^q \beta_j X_{t-j} \right) = 0 \\ \frac{\partial \text{SCR}}{\partial \beta_m} &= -2 \sum_t X_{t-l} \left(Y_t - \alpha_0 - \sum_{i=1}^p \phi_i Y_{t-i} - \sum_{j=0}^q \beta_j X_{t-j} \right) = 0 \end{aligned}$$

Solution en notation matricielle

$$\begin{pmatrix} \sum X_t^2 & \sum X_t Y_{t-1} & \cdots & \sum X_t X_{t-q} \\ \sum Y_{t-1} X_t & \sum Y_{t-1}^2 & \cdots & \sum Y_{t-1} X_{t-q} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{t-q} X_t & \sum X_{t-q} Y_{t-1} & \cdots & \sum X_{t-q}^2 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \phi_1 \\ \vdots \\ \beta_q \end{pmatrix} = \begin{pmatrix} \sum X_t Y_t \\ \sum Y_{t-1} Y_t \\ \vdots \\ \sum X_{t-q} Y_t \end{pmatrix}$$

Écriture matricielle du modèle ARDL

Pour simplifier le calcul, on réécrit le modèle en notation matricielle :

$$Y = X\theta + \varepsilon$$

où :

$\mathbf{Y} = [Y_{p+1}, \dots, y_T]'$ vecteur colonne de taille $((T - p) \times 1)$

$\mathbf{X} = [\mathbf{1}, Y_{-1}, \dots, Y_{-p}, X_0, \dots, X_{-q}]$ matrice d'ordre $(T - p) \times (1 + p + q + 1)$.

$\boldsymbol{\theta} = [\alpha_0, \phi_1, \dots, \phi_p, \beta_0, \dots, \beta_q]'$ vecteur des paramètres à estimer.

$\boldsymbol{\varepsilon} = [\varepsilon_{p+1}, \dots, \varepsilon_T]'$ vecteur des résidus.

Estimateur MCO

L'estimateur des moindres carrés s'obtient par la formule classique :

$$\hat{\theta} = (X'X)^{-1}X'Y \quad (3.4.3)$$

où :

- $\hat{\theta}$: vecteur des coefficients estimés,
- $X'X$: matrice des produits croisés,
- $X'Y$: vecteur des produits croisés.

Propriétés des estimateurs MCO dans l'ARDL**1. Consistance (convergence) :**

- Si les variables sont I(0) ou cointegrées, $\hat{\theta}$ converge vers la vraie valeur.
- Si certaines variables sont I(1) mais non cointegrées, l'estimateur peut être biaisé.

2. Normalité Asymptotique :

- Sous la stationnarité, $\sqrt{T}(\hat{\theta} - \theta) \rightarrow_d \mathcal{N}(0, Q^{-1}\sigma^2)$, où $Q = \lim\left(\frac{X'X}{T}\right)$.

3. Efficacité :

- Si les résidus sont homoscedastiques et non autocorrélés, MCO est le meilleur estimateur linéaire sans biais.

Exemple Numérique

Supposons un ARDL(1,1) :

$$Y_t = \alpha_0 + \phi_1 Y_{t-1} + \beta_0 x_t + \beta_1 X_{t-1} + \varepsilon_t$$

Matrice X et Vecteur Y

$$Y = \begin{pmatrix} y_3 \\ y_4 \\ y_5 \end{pmatrix} \quad X = \begin{pmatrix} 1 & y_2 & x_3 & x_2 \\ 1 & y_3 & x_4 & x_3 \\ 1 & y_4 & x_5 & x_4 \end{pmatrix}$$

Estimation MCO

$$\hat{\theta} = (X'X)^{-1}X'Y$$

6. Implémentation sous logiciel R

Le package utilisé : `dynlm` (série de package "dyn")

```
modele <- dynlm(y ~ L(y, 1) + x + L(x, 1), data = data)
```

Conclusion

L'estimation des moindres carrés ordinaires consiste à résoudre un système d'équations linéaires obtenu par l'inversion de la matrice $X'X$ ce que l'on peut faire également avec la commande `ardl` du package `ARDL`. Cette approche permet d'obtenir des estimateurs cohérents, notamment sous l'hypothèse de cointégration entre les variables. Cependant, il est important de noter que l'estimateur MCO dans les modèles ARDL reste sensible à certaines violations des hypothèses classiques, telles que l'autocorrélation et l'hétéroscédasticité des erreurs. Ces phénomènes peuvent biaiser les résultats et affecter la validité des inférences statistiques.

3.4.2 Estimation d'un modèle ARDL par la méthode du maximum de vraisemblance (MV)

L'estimation par maximum de vraisemblance (MV) pour les modèles ARDL est une méthode puissante qui offre des estimateurs efficaces sous l'hypothèse de normalité des résidus. Cette méthode est particulièrement utile pour capturer les relations dynamiques à court et long terme entre variables économiques.

Conditions d'application

— **Normalité des résidus :**

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Condition essentielle pour l'efficacité des estimateurs. En cas de non-normalité, les estimateurs restent convergents mais perdent en efficacité.

— **Spécification correcte du modèle :**

— Absence de variables omises

— les retards (p et q) sont bien choisis via (AIC, BIC)

— **Absence d'autocorrélation et d'hétéroscédasticité :** Condition nécessaire pour éviter un biais dans la matrice de variance-covariance des estimateurs.

Formulation de la log-vraisemblance

Pour un modèle ARDL(p, q) :

$$Y_t = \alpha_0 + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=0}^q \beta_j X_{t-j} + \varepsilon_t \quad (3.4.4)$$

avec

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2) \quad \Rightarrow \quad f(\varepsilon_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma^2}\right) \quad (3.4.5)$$

Or, on a :

$$\varepsilon_t = Y_t - \mu_t \quad \text{avec} \quad \mu_t = \alpha_0 + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=0}^q \beta_j X_{t-j} \quad (3.4.6)$$

Ainsi, on obtient :

$$f(\varepsilon_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_t - \mu_t)^2}{2\sigma^2}\right) \quad (3.4.7)$$

En supposant que les observations Y_t sont conditionnellement indépendantes, la fonction de vraisemblance est le produit des densités :

$$L(\boldsymbol{\theta}) = \prod_{t=T_0+1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_t - \mu_t)^2}{2\sigma^2}\right) \quad (3.4.8)$$

où $T_0 = \max(p, q)$ correspond au nombre de retards pour éviter les valeurs initiales manquantes. En prenant le logarithme, on obtient :

$$\ln L(\boldsymbol{\theta}) = \sum_{t=T_0+1}^T \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_t - \mu_t)^2}{2\sigma^2}\right) \right] \quad (3.4.9)$$

Ce qui se développe en :

$$\ln L(\boldsymbol{\theta}) = \sum_{t=T_0+1}^T \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2) - \frac{(Y_t - \mu_t)^2}{2\sigma^2} \right] \quad (3.4.10)$$

En factorisant les termes constants, on obtient la forme finale :

$$\boxed{\ln L(\boldsymbol{\theta}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=T_0+1}^T \left(Y_t - \alpha_0 - \sum_{i=1}^p \phi_i Y_{t-i} - \sum_{j=0}^q \beta_j X_{t-j} \right)^2} \quad (3.4.11)$$

où $n = T - T_0$ est la taille effective de l'échantillon et

$$\boldsymbol{\theta} = (\alpha_0, \phi_1, \dots, \phi_p, \beta_0, \dots, \beta_q, \sigma^2)$$

est le vecteur des paramètres à estimer

Étapes de l'estimation par maximum de vraisemblance

Étape 1 : Initialisation

Avant de maximiser la log-vraisemblance, on fournit des valeurs initiales aux paramètres. Celles-ci sont généralement obtenues en estimant le modèle par la méthode des moindres carrés ordinaires (MCO).

Étape 2 : Maximisation numérique

La fonction $\ln L(\boldsymbol{\theta})$ est maximisée numériquement à l'aide d'algorithme de Newton-Raphson, car il n'existe pas de solution analytique directe.

Étape 3 : Calcul des erreurs standards

Les erreurs standards des estimateurs sont obtenues à partir de la matrice d'information de Fisher, qui mesure la précision des estimations :

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E} \left[\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] \quad (3.4.12)$$

L'inverse de cette matrice donne la matrice de variance-covariance des estimateurs :

$$\text{Var}(\hat{\boldsymbol{\theta}}) = \mathcal{I}^{-1}(\boldsymbol{\theta}) \quad (3.4.13)$$

Les erreurs standards sont alors les racines carrées des éléments diagonaux de cette matrice.

Comparaison de la méthode du maximum de vraisemblance vs la méthode des moindres carrés ordinaires

| Critère | Maximum de Vraisemblance (MV) | Moindres Carrés Ordinaires (MCO) |
|------------------------|--|--------------------------------------|
| Hypothèse de normalité | Requise (sinon biais asymptotique) | Non requise |
| Efficacité | Optimale si $\varepsilon_t \sim \mathcal{N}$ | Moins efficace que MV sous normalité |
| Complexité | Calcul intensif (optimisation numérique) | Simple (formule analytique) |
| Applications | Modèles non linéaires, GARCH, VAR | Modèles linéaires standards |

TABLE 3.2 – Comparaison de la méthode du maximum de vraisemblance vs la méthode des moindres carrés ordinaires

Interprétation du tableau comparatif Le tableau présente une comparaison entre la méthode du maximum de vraisemblance (MV) et celle des moindres carrés ordinaires (MCO), deux approches fondamentales en estimation économétrique.

Tout d’abord, la méthode MV repose sur l’hypothèse que les erreurs sont normalement distribuées. Cette hypothèse est cruciale car, en son absence, les estimateurs peuvent être biaisés de manière asymptotique. En revanche, la méthode MCO ne requiert pas la normalité des erreurs pour fournir des estimateurs sans biais, tant que les conditions classiques du modèle linéaire sont respectées.

Ensuite, en termes d’efficacité, la MV est optimale lorsque l’hypothèse de normalité est vérifiée, ce qui signifie qu’elle fournit des estimateurs ayant une variance plus faible. Cependant, cet avantage disparaît si la distribution des erreurs s’écarte de la normalité, auquel cas MCO peut être préférée pour sa robustesse.

En ce qui concerne la mise en œuvre, la MV implique des calculs plus complexes, nécessitant des méthodes d’optimisation numérique, tandis que la MCO repose sur une solution analytique simple et directe.

Enfin, la MV est largement utilisée pour l’estimation de modèles non linéaires, tels que les modèles GARCH ou VAR, où les MCO ne sont pas applicables. À l’inverse, les MCO restent la méthode de référence pour les modèles linéaires standards.

Ainsi, le choix entre MV et MCO dépend à la fois des propriétés statistiques recherchées, des hypothèses sur les erreurs, et de la nature du modèle utilisé.

Implémentation pratique

Sous R, on utilise le package `stats` :

```
library(stats)
# Exemple : ARDL(2,1) avec variable exogène x
model <- arima(y, order = c(2, 0, 0), xreg = x)
summary(model)
```

Ce programme fournit les estimateurs du maximum de vraisemblance grâce à la commande `summary`.

TABLE 3.3 – Structure de l’objet `summary(model)` dans un modèle ARDL(2,1)

| Composant | Longueur | Classe | Mode | Interprétation |
|------------------------|----------|--------|-----------|---|
| <code>coef</code> | 4 | -none- | numeric | Coefficients estimés (AR(1), AR(2), exogène, constante) |
| <code>sigma2</code> | 1 | -none- | numeric | Variance de l’erreur (bruit blanc) |
| <code>var.coef</code> | 16 | -none- | numeric | Matrice 4×4 des variances-covariances des coefficients |
| <code>mask</code> | 4 | -none- | logical | Indique si chaque coefficient a été estimé (TRUE/FALSE) |
| <code>loglik</code> | 1 | -none- | numeric | Log-vraisemblance maximale du modèle |
| <code>aic</code> | 1 | -none- | numeric | Critère d’information d’Akaike (AIC) |
| <code>arma</code> | 7 | -none- | numeric | Structure du modèle ARMA (p, d, q, etc.) |
| <code>residuals</code> | 100 | ts | numeric | Résidus (erreurs) du modèle sur 100 observations |
| <code>call</code> | 4 | -none- | call | Appel original de la fonction <code>arma()</code> |
| <code>series</code> | 1 | -none- | character | Nom de la série temporelle modélisée |
| <code>code</code> | 1 | -none- | numeric | Code de convergence (0 = succès) |
| <code>n.cond</code> | 1 | -none- | numeric | Nombre d’observations conditionnelles |
| <code>nobs</code> | 1 | -none- | numeric | Nombre total d’observations utilisées |
| <code>model</code> | 10 | -none- | list | Liste contenant les composantes internes du modèle |

Éléments de sortie :

- Coefficients ($\alpha_0, \phi_i, \beta_j$)
- Erreurs standards et tests t
- Log-vraisemblance et critères AIC/BIC

Avantages et limitations

Ces estimateurs sont efficaces sous l’hypothèse de normalité, mais deviennent sensibles aux écarts à cette hypothèse et présentent une complexité numérique accrue.

Recommandations pratiques

1. Utiliser l’estimateur du maximum de vraisemblance (MV) lorsque la normalité est plausible (à vérifier par des tests QQ-plot ou Shapiro-Wilk)
2. Privilégier les moindres carrés ordinaires (MCO) pour les grands échantillons ou en cas de doute sur la normalité.
3. Toujours vérifier les diagnostics de résidus (autocorrélation, hétéroscédasticité).

Conclusion

En conclusion, la méthode du maximum de vraisemblance est utilisable pour estimer un modèle ARDL, mais elle exige des résidus gaussiens pour une inférence valide. L’un de ses atouts est l’efficacité des estimateurs, et la méthode reste adaptable aux modèles complexes. Cependant, elle est sensible aux hypothèses, et les calculs sont très complexes. Il est recommandé d’utiliser la méthode des moindres carrés ordinaires (MCO) si la normalité est douteuse. En revanche, si nous disposons d’un petit échantillon et que la normalité est plausible, alors la méthode du maximum de vraisemblance est tout à fait adéquate.

3.5 Inférence Bayésienne dans les Modèles ARDL et Extension des Bounds Tests

Les modèles ARDL (Autoregressive Distributed Lag) constituent un cadre flexible pour analyser les relations dynamiques entre variables économiques. L'approche bayésienne, telle que décrite par Koop (2008) ([15]), offre des avantages substantiels dans l'estimation de ces modèles, particulièrement en présence d'échantillons de petite taille ou lorsque des informations a priori sont disponibles.

Spécification du Modèle

Un modèle ARDL(p, q) s'écrit :

$$Y_t = \alpha + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=0}^q \beta_j X_{t-j} + \varepsilon_t \quad (3.5.1)$$

où :

Y_t est la variable dépendante

X_t est la variable explicative

$\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ est le terme d'erreur

p et q représentent les ordres de retard pour Y_t et X_t respectivement

3.5.1 Inférence Bayésienne dans les Modèles ARDL

L'approche bayésienne consiste, dans cette catégorie de modèles, à tout d'abord spécifier des distributions *a priori* sur les paramètres ($\phi_i, \beta_j, \sigma^2$). On combine par la suite ces lois *a priori* avec la fonction de vraisemblance des données via le théorème de Bayes, afin d'obtenir les distributions *a posteriori*, souvent par des méthodes MCMC (Markov Chain Monte Carlo).

Étapes clés :

1. Choix des lois a priori :

Les coefficients (ϕ, β) ont pour lois a priori des lois gaussiennes (par exemple, $\mathcal{N}(0, 10^2)$). Pour la variance (σ^2), on la modélise généralement par une loi *Inverse-Gamma* ou *Half-Cauchy*.

Pour les modèles non stationnaires, on utilise des distributions favorisant la stationnarité (par exemple, contraintes sur ϕ_i).

2. Calcul de la vraisemblance :

La fonction de vraisemblance est de forme gaussienne si le terme d'erreur ε_t suit une distribution normale.

3. Distribution a posteriori :

La distribution a posteriori des paramètres est généralement obtenue par des méthodes MCMC (comme l'algorithme de Gibbs ou Metropolis-Hastings).

Remarque 3.5.1. 1. Dans le cas de la non-stationnarité, il existe un risque d'obtenir des distributions a posteriori impropres si les lois a priori ne contraignent pas suffisamment la stationnarité.

2. En cas de calcul intensif, les méthodes MCMC peuvent être lentes lorsque les ordres de retard p et q sont élevés.
3. Dans les modèles ARDL, les résultats sont très sensibles au choix des lois a priori, ce qui nécessite une attention particulière lors de leur spécification.

Extension Bayésienne des Bounds Tests

L'extension bayésienne des *Bounds Tests* est une adaptation de l'approche classique de Pesaran, Shin et Smith (2001) ([20]) pour tester la cointégration dans les modèles ARDL, en utilisant une inférence bayésienne. Cette méthode permet d'évaluer l'existence d'une relation de long terme entre variables, même en présence de non-stationnarité (séries $I(0)$ ou $I(1)$), tout en intégrant des informations *a priori* et en fournissant des probabilités directes sur l'hypothèse de cointégration.

Rappel : Bounds Test Classique

Le test standard repose sur une approche fréquentiste et suit les étapes suivantes :

1. Estimation d'un modèle ARDL en incluant des retards de la variable dépendante et des variables explicatives.
2. Test de la significativité conjointe des termes au niveau (effets de long terme) à l'aide d'un test F ou d'un test de Wald.
3. Comparaison de la statistique F aux valeurs critiques tabulées (bornes inférieure et supérieure), qui dépendent du nombre de variables et de leur ordre d'intégration.

Problème :

Les valeurs critiques sont asymptotiques et peuvent être peu fiables en petits échantillons. De plus, l'approche fréquentiste ne fournit pas de probabilité directe concernant l'hypothèse de cointégration.

Approche Bayésienne

L'approche bayésienne résout ces limites en utilisant des distributions *a posteriori* pour les paramètres de long terme, et en calculant directement la probabilité que le terme de correction d'erreur (ECM) soit significatif (c'est-à-dire, la probabilité de cointégration). Voici les différentes étapes décrites ci-dessous.

a. Spécification du Modèle ARDL-ECM

Le modèle est réécrit sous forme à correction d'erreur (ECM) pour isoler la relation de long terme :

$$\Delta Y_t = \alpha + \lambda(Y_{t-1} - \theta X_{t-1}) + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \beta_j \Delta X_{t-j} + \varepsilon_t$$

λ représente la vitesse d'ajustement vers l'équilibre de long terme (si $\lambda < 0$, on a la cointégration).

θ est le paramètre de long terme.

b. Choix des lois a Priori

Pour le paramètre λ , la distribution a priori considérée doit être centrée sur zéro avec une

queue gauche pour autoriser des valeurs négatives. Par exemple :

$$\lambda \sim \mathcal{N}(-0,5, 0,5)$$

Le paramètre θ est modélisé par une loi a priori peu informative, par exemple :

$$\theta \sim \mathcal{N}(0, 10^2)$$

c. Inférence bayésienne

L'estimation s'effectue via les méthodes MCMC (par exemple, Gibbs sampling) pour obtenir la distribution a posteriori de λ . La probabilité de cointégration est la proportion des tirages MCMC où λ est significativement négatif, c'est-à-dire :

$$P(\lambda < 0 \mid \text{données})$$

d. Décision

Étant donné le cadre bayésien, si par exemple $P(\lambda < 0) > 0,95$, on conclut à une cointégration avec une forte probabilité.

Avantages de l'approche bayésienne

L'un des avantages de l'approche bayésienne est son interprétation intuitive. En effet, avec le calcul de la probabilité directe sur l'hypothèse de cointégration, nous pouvons accepter fortement, modérément ou faiblement l'hypothèse de cointégration, contrairement aux p-values qui ne permettent qu'un rejet ou non-rejet selon un seuil arbitraire.

Un autre avantage notable est la robustesse en petits échantillons. Une meilleure performance est souvent constatée par rapport à l'approche fréquentiste lorsque T , représentant la taille de l'échantillon, est faible.

Enfin, l'approche bayésienne évite le recours aux seuils critiques fixes. Les bornes usuelles du test de cointégration sont remplacées par des *intervalles de crédibilité*, calculés directement à partir de la distribution a posteriori des paramètres.

Exemple de simulation sous langage R avec le package *Bayesian Regression Models using Stan*

Voici un exemple simple des étapes de simulation dans le cadre de l'approche bayésienne.

```
library(brms)

# Données simulées
data <- data.frame(y = ..., x = ...)

# Modèle ARDL(1,1) sous forme ECM
formula <- bf(
  diff(y) ~ 1 + error_correction_term + lag(diff(y)) + diff(x),
  error_correction_term = (y - theta * x)[-1], # Terme ECM
```

```

theta ~ 1,
nl = TRUE
)

# Priors
priors <- c(
  prior(normal(-0.5, 0.5), nlpar = "error_correction_term"),
  prior(normal(0, 10), nlpar = "theta")
)

# Estimation
fit <- brm(formula, data = data, prior = priors, chains = 4)

# Interprétation
posterior_samples(fit, "error_correction_term") %>%
  summarise(prob_coint = mean(error_correction_term < 0))

```

Comparaison de l'extension bayésienne avec le Bounds Test Classique

TABLE 3.4 – Comparaison entre Bounds Test Classique et Extension Bayésienne

| Critère | Bounds Test Classique | Extension Bayésienne |
|----------------------|----------------------------|--|
| Approche | Fréquentiste (test F) | Probabiliste (posterior de λ) |
| Interprétation | Comparaison à seuils fixes | Probabilité directe (e.g., 95%) |
| Taille d'échantillon | Sensible en petits T | Plus robuste |
| Priors | Non utilisés | Intégration d'informations a priori |

Remarque 3.5.2. Dans le contexte des modèles ARDL avec correction d'erreur (ECM) et de l'inférence bayésienne, la probabilité *a posteriori* de λ (le paramètre de vitesse d'ajustement dans le terme de correction d'erreur) est conditionnée à l'ensemble des données observées, c'est-à-dire à la fois Y et X .

Explication Détaillée

Conditionnement dans le modèle bayésien : La distribution *a posteriori* de λ est obtenue via le théorème de Bayes :

$$P(\lambda \mid \text{data}) \propto P(\text{data} \mid \lambda) \cdot P(\lambda),$$

où :

- $\text{data} = \{Y_t, X_t\}_{t=1}^T$ inclut toutes les observations de Y et X ,
- $P(\text{data} \mid \lambda)$ est la vraisemblance du modèle ARDL-ECM,
- $P(\lambda)$ est la loi a priori sur λ .

Par conséquent, λ dépend conjointement de Y et X , car le terme de correction d'erreur ($Y_{t-1} - \theta X_{t-1}$) dans l'ECM lie les deux variables. De plus, le modèle ARDL-ECM est multivarié : il capture une relation dynamique entre Y et X .

La vraisemblance $P(\text{data} \mid \lambda)$ intègre la structure des retards (ΔY_t , ΔX_t , et le terme de long terme $Y_{t-1} - \theta X_{t-1}$). Ainsi, λ est estimé en tenant compte de toutes les dépendances entre les variables.

Exemple avec un modèle ARDL(1,1) bayésien : Soit le modèle ECM :

$$\Delta Y_t = \lambda(Y_{t-1} - \theta X_{t-1}) + \beta \Delta X_t + \varepsilon_t$$

La distribution a posteriori de λ :

La distribution $P(\lambda \mid Y, X)$ est estimée via MCMC en utilisant toutes les données (Y et X). Si $\lambda < 0$ (avec une probabilité élevée, par exemple 95%), on conclut à une cointégration.

Implications pratiques :

1. *Interprétation* : Une probabilité a posteriori $P(\lambda < 0 \mid Y, X) = 0,98$ signifie qu'il y a 98% de chances que la relation de cointégration existe, compte tenu de l'ensemble des données.
2. *Comparaison avec les MCO* : Dans le cadre fréquentiste, le test de Pesaran repose sur la statistique F calculée à partir de Y et X , mais sans probabilité directe. L'approche bayésienne fournit une quantification intuitive de l'incertitude.

Conclusion

L'inférence bayésienne dans les ARDL combine la richesse des modèles dynamiques avec la rigueur probabiliste, idéale pour les problèmes où l'incertitude est élevée ou les données sont limitées. Elle nécessite cependant une réflexion approfondie sur les priors et des outils computationnels adaptés.

3.6 Application sur données réelles

3.6.1 Présentation des données “Denmark”

L'article de Johansen et Juselius (1990), intitulé *Maximum Likelihood Estimation and Inference on Cointegration. With Applications to the Demand for Money*, constitue une contribution fondamentale à l'analyse des séries temporelles et à la théorie de la cointégration. Les données utilisées dans cette étude, souvent appelées *jeu de données Denmark*, comprennent des séries macroéconomiques relatives à l'économie danoise.

Description des données

Les données utilisées sont des séries trimestrielles couvrant la période allant du premier trimestre 1974 au troisième trimestre 1987, soit un total de 51 observations, et comprennent les variables suivantes :

- **LRM** : logarithme de la masse monétaire réelle (Real Money Supply), mesurée comme M2 divisée par l'indice des prix à la consommation (IPC). Elle représente la demande de monnaie en termes réels.
- **LRY** : logarithme du revenu réel (Real Income), mesuré par le PIB réel ajusté de l'inflation, servant de proxy pour l'activité économique.

- **IBO** : taux d'intérêt sur les obligations (Bond Rate), représentant le rendement des obligations à long terme, exprimant le coût d'opportunité de détenir de la monnaie plutôt que des actifs financiers.
- **IDE** : taux d'intérêt sur les dépôts (Deposit Rate), représentant le rendement des comptes d'épargne ou dépôts bancaires, mesurant le rendement des actifs monétaires.

Objectif de l'étude

L'objectif principal de Johansen et Juselius (1990) était d'étudier la demande de monnaie au Danemark en utilisant une approche de cointégration. Ils cherchaient à identifier des relations d'équilibre à long terme entre ces variables, tout en tenant compte des dynamiques de court terme.

Méthodologie appliquée

- **Test de cointégration (approche de Johansen)** : utilisation du test de trace et du test de la valeur propre maximale pour déterminer le nombre de relations de cointégration.
- **Modèle à correction d'erreur (VECM)** : estimation d'un modèle vectoriel à correction d'erreur pour analyser les ajustements dynamiques entre les variables.

Nous allons suivre une approche étape par étape, depuis l'exploration des données jusqu'à l'estimation du modèle à correction d'erreur (VECM).

1. Chargement et Exploration des Données

Les données utilisées sont les suivantes :

- **LRM** : Logarithme de la masse monétaire réelle (M2 / IPC).
- **LRY** : Logarithme du revenu réel (PIB / déflateur).
- **IBO** : Taux d'intérêt des obligations (long terme).
- **IDE** : Taux d'intérêt des dépôts (court terme).

Code R pour importer et visualiser les données

```
# Installer les packages nécessaires (si ce n'est pas déjà fait)
install.packages("urca")      # Pour les tests de cointégration
install.packages("vars")     # Pour le VECM
install.packages("ggplot2")  # Pour les graphiques
install.packages("ARDL")     # Pour les modèles ARDL

# Charger les packages
library(urca)
library(vars)
library(ggplot2)
library(ARDL)
library(patchwork)

# Charger les données Denmark (disponibles dans le package 'urca' ou
```

```

dans le package 'ARDL')
data(denmark)
df <- denmark

# Aperçu des données
head(df)
summary(df)

```

Le tableau suivant résume les principales statistiques descriptives des variables utilisées dans l'analyse, que l'on obtient grâce à la commande "summary".

TABLE 3.5 – Statistiques descriptives (1974Q1–1987Q3)

| | LRM | LRY | LPY | IBO | IDE | Observations |
|--------------------------|-------|-------|--------|--------|--------|--------------|
| Min. | 11.58 | 5.787 | -0.619 | 0.0977 | 0.0675 | 51 |
| 1 ^{er} quartile | 11.63 | 5.910 | -0.276 | 0.1336 | 0.0775 | |
| Médiane | 11.70 | 5.940 | 0.0333 | 0.1611 | 0.0880 | |
| Moyenne | 11.75 | 5.952 | 0.0095 | 0.1564 | 0.0904 | |
| 3 ^e quartile | 11.82 | 5.986 | 0.324 | 0.1794 | 0.1020 | |
| Max. | 12.08 | 6.099 | 0.468 | 0.2108 | 0.1210 | |

Visualisation des séries

Le code R ci-dessous permet de visualiser simultanément les quatre principales séries de données macroéconomiques (LRM, LRY, IBO et IDE) sur une même fenêtre graphique :

```

1 library(ggplot2)
2 library(gridExtra)
3
4 p1 <- ggplot(df, aes(x = seq_along(LRM), y = LRM)) +
5   geom_line() +
6   ggtitle("LRM (Log Real Money)")
7
8 p2 <- ggplot(df, aes(x = seq_along(LRY), y = LRY)) +
9   geom_line() +
10  ggtitle("LRY (Log Real Income)")
11
12 p3 <- ggplot(df, aes(x = seq_along(IBO), y = IBO)) +
13   geom_line() +
14   ggtitle("IBO (Bond Rate)")
15
16 p4 <- ggplot(df, aes(x = seq_along(IDE), y = IDE)) +
17   geom_line() +
18   ggtitle("IDE (Deposit Rate)")
19
20 grid.arrange(p1, p2, p3, p4, ncol = 2)

```

Comme on peut le voir à la figure 3.1

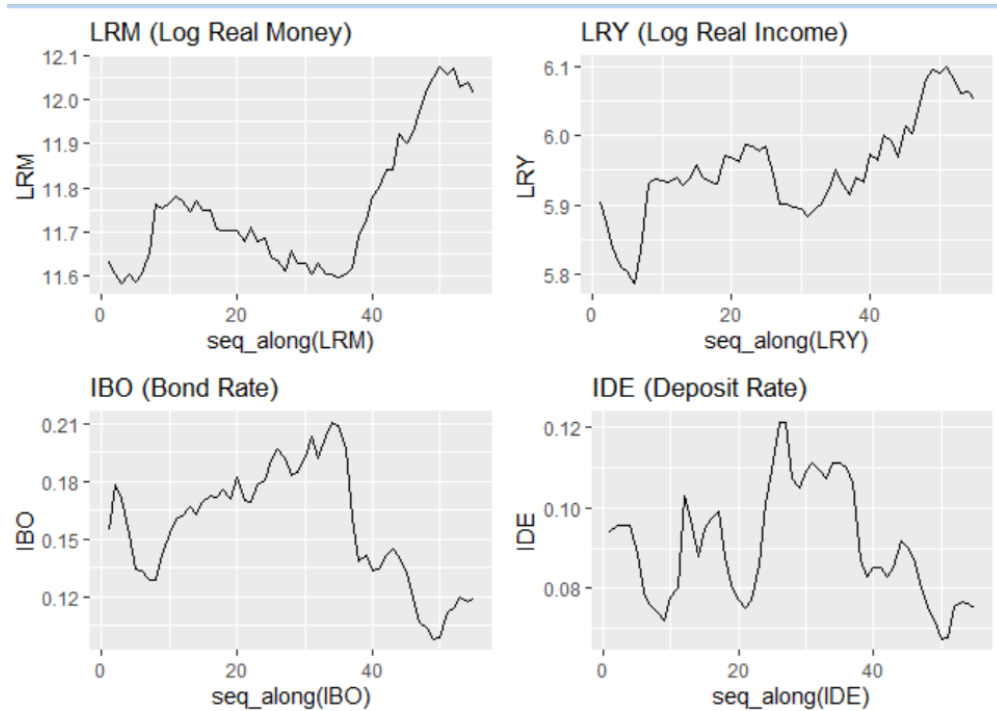


FIGURE 3.1 – Représentation graphique de quatre séries

Analyse des graphes

Les deux séries LRM (logarithme de la monnaie réelle) et LRY (logarithme du revenu réel) présentent une tendance croissante claire sur la période observée, ce qui signifie qu'elles ne sont pas stationnaires en niveau.

La série IBO fluctue autour d'une moyenne, mais présente des périodes de hausse et de baisse marquées, notamment une chute brutale après le point $t=40$, ce qui suggère que la série pourrait ne pas être stationnaire.

IDE fluctue autour d'une moyenne, avec des pics et des creux, mais sans tendance claire à long terme. Cette série semble plus proche de la stationnarité, car elle oscille autour d'une valeur moyenne sans tendance évidente, bien que cette hypothèse doive être vérifiée.

Pour une analyse rigoureuse, il est nécessaire de vérifier la stationnarité de ces séries à l'aide de tests statistiques, tels que le test de Dickey-Fuller augmenté (ADF).

Application du test (ADF) sur chaque série

```

1
2 adf_lrm <- ur.df(df$LRM, type = "trend", lags = 4)
3 adf_lry <- ur.df(df$LRY, type = "trend", lags = 4)
4 adf_ibo <- ur.df(df$IBO, type = "drift", lags = 4)
5 adf_ide <- ur.df(df$IDE, type = "drift", lags = 4)

```

afficher les résultats

```

1 summary(adf_lrm)
2 summary(adf_lry)
3 summary(adf_ibo)
4 summary(adf_ide)

```

Résumé des résultats du test Augmented Dickey-Fuller (ADF)

| Série | Type de test | Statistique τ | Valeur critique (5%) | Stationnarité |
|-------|----------------------|--------------------|----------------------|------------------|
| LRM | Avec tendance (tau3) | -2.0913 | -3.45 | Non stationnaire |
| LRY | Avec tendance (tau3) | -2.2126 | -3.45 | Non stationnaire |
| IBO | Avec dérive (tau2) | -1.0924 | -2.89 | Non stationnaire |
| IDE | Avec dérive (tau2) | -1.8223 | -2.89 | Non stationnaire |

TABLE 3.6 – Résumé des résultats du test ADF sur les séries étudiées.

Interprétation :

Pour chaque série, la statistique τ est supérieure à la valeur critique au seuil de 5%, ce qui signifie que l'hypothèse nulle de racine unitaire (non stationnarité) ne peut pas être rejetée. Ainsi, toutes les séries sont non stationnaires en niveau et nécessitent une transformation (différenciation) pour devenir stationnaires.

Étant donné que les séries temporelles sont non stationnaires en niveau, il convient d'utiliser le test de cointégration de Johansen pour analyser leurs relations à long terme.

3.6.2 Test de Johansen (avec constante dans la relation de cointégration)

```

1 variables <- c("LRM", "LRY", "IBO", "IDE")
2 data_vecm <- df[, variables]
3
4
5 sum(is.na(data_vecm))
6
7 johansen_test <- ca.jo(data_vecm, type = "eigen", ecdet = "const", K = 2)
8
9 summary(johansen_test)

```

Résumé des résultats du test de cointégration de Johansen

Test de Johansen utilisant la statistique du maximum propre (maximal eigenvalue test), sans tendance linéaire ni constante dans la relation de cointégration.

Valeurs propres (Eigenvalues)

| Rang | 1 | 2 | 3 | 4 | 5 |
|-------------------------|--------|--------|--------|--------|-------------|
| Valeur propre λ | 0.4697 | 0.1742 | 0.1181 | 0.0422 | ≈ 0 |

Statistiques de test et valeurs critiques

| Hypothèse sur le rang r | Statistique de test | 10% | 5% | 1% |
|---------------------------|---------------------|-------|-------|-------|
| $r \leq 3$ | 2.29 | 7.52 | 9.24 | 12.97 |
| $r \leq 2$ | 6.66 | 13.75 | 15.67 | 20.20 |
| $r \leq 1$ | 10.15 | 19.77 | 22.00 | 26.81 |
| $r = 0$ | 33.62 | 25.56 | 28.14 | 33.24 |

Interprétation : La statistique de test est supérieure à la valeur critique uniquement pour $r = 0$, ce qui permet de rejeter l'hypothèse nulle d'absence de cointégration. Les autres hypothèses ne sont pas rejetées. Ainsi, on conclut à l'existence d'une **relation de cointégration unique** (rang $r = 1$) entre les séries.

Relations de cointégration (vecteurs propres normalisés)

Les vecteurs propres normalisés, représentant les relations de cointégration, sont donnés dans le tableau ci-dessous :

| Variable | Relation 1 | Relation 2 | Relation 3 | Relation 4 | Constante |
|-----------|------------|------------|------------|------------|-----------|
| LRM.12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| LRY.12 | -0.9691 | -1.3760 | -2.2592 | -0.1439 | -0.4096 |
| IBO.12 | 5.4028 | -0.3352 | 1.6408 | -11.7074 | 2.3548 |
| IDE.12 | -4.1403 | 9.0059 | -3.7193 | 7.5375 | -1.1374 |
| Constante | -6.4781 | -4.3298 | 1.7861 | -10.2593 | -9.5376 |

Ces vecteurs indiquent la combinaison linéaire des variables qui forme une relation stationnaire à long terme.

Résumé des résultats du modèle VECM

Équation 1 : ΔLRM

| Variable | Coefficient | p-value | Interprétation |
|----------|-------------|--------------|--|
| ect1 | -0.29978 | 0.0001 | Ajustement significatif au déséquilibre de long terme (vitesse de correction d'environ 30%). |
| LRM.d11 | -0.51983 | 0.003 | Effet négatif significatif du retard de LRM sur sa propre variation. |
| LRY.d11 | 0.36751 | 0.046 | Impact positif significatif du retard de LRY sur LRM. |
| IBO.d11 | -1.44128 | 0.002 | Impact négatif fort du retard de IBO sur LRM. |
| IDE.d11 | -0.11657 | 0.863 | Non significatif. |
| | | $R^2 = 0.40$ | Le modèle explique 40% de la variance de ΔLRM . |

Conclusion : LRM réagit fortement aux déséquilibres de long terme et est influencée à court terme par les variations passées de LRY et IBO.

Équation 2 : ΔLRY

| Variable | Coefficient | p-value | Interprétation |
|----------|-------------|--------------|---|
| ect1 | 0.02694 | 0.664 | Ajustement non significatif au déséquilibre de long terme. |
| LRM.d11 | 0.29421 | 0.046 | Impact positif significatif du retard de LRM sur LRY. |
| LRY.d11 | -0.04730 | 0.761 | Non significatif. |
| IBO.d11 | 0.01768 | 0.963 | Non significatif. |
| IDE.d11 | -0.90331 | 0.124 | Non significatif. |
| | | $R^2 = 0.21$ | Le modèle explique 21% de la variance de ΔLRY . |

Conclusion : LRY ne corrige pas les déséquilibres de long terme. Seul le retard de LRM influence significativement LRY à court terme.

Équation 3 : ΔIBO

| Variable | Coefficient | p-value | Interprétation |
|--------------|-------------|---------|--|
| ect1 | 0.00392 | 0.864 | Ajustement non significatif au déséquilibre de long terme. |
| LRM.dl1 | 0.00662 | 0.901 | Non significatif. |
| LRY.dl1 | 0.14629 | 0.013 | Impact positif significatif du retard de LRY sur IBO. |
| IBO.dl1 | 0.37769 | 0.009 | Effet positif significatif de son propre retard. |
| IDE.dl1 | 0.02748 | 0.898 | Non significatif. |
| $R^2 = 0.29$ | | | Le modèle explique 29% de la variance de ΔIBO . |

Conclusion : IBO est influencé par ses propres variations passées et celles de LRY. Aucune correction significative des déséquilibres de long terme.

Équation 4 : ΔIDE

| Variable | Coefficient | p-value | Interprétation |
|--------------|-------------|---------|--|
| ect1 | 0.02000 | 0.186 | Ajustement non significatif au déséquilibre de long terme. |
| LRM.dl1 | 0.04396 | 0.213 | Non significatif. |
| LRY.dl1 | 0.01405 | 0.709 | Non significatif. |
| IBO.dl1 | 0.40212 | <0.001 | Impact positif très significatif du retard de IBO sur IDE. |
| IDE.dl1 | 0.05077 | 0.718 | Non significatif. |
| $R^2 = 0.37$ | | | Le modèle explique 37% de la variance de ΔIDE . |

Conclusion : IDE est fortement influencé par les variations passées de IBO. Aucune réaction significative aux déséquilibres de long terme.

Synthèse générale

Seule la variable LRM corrige significativement les déséquilibres de long terme, avec un coefficient négatif indiquant une vitesse de correction d'environ 30%. Les autres variables réagissent principalement aux variations passées sans ajuster l'équilibre à long terme. Les dynamiques à court terme montrent des influences croisées importantes, notamment entre LRM, LRY, IBO et IDE.

3.6.3 Présentation de la fonction `ardl()` et de ses composantes

Dans cette étude, nous utilisons la fonction `ardl()` du package `ARDL` sous le logiciel R pour estimer des modèles ARDL (Autoregressive Distributed Lag). Cette fonction permet de spécifier de manière souple des modèles dynamiques, en autorisant des ordres de retard différents pour chaque variable explicative.

Syntaxe générale

```
1 ardl(formula, data, order, start = NULL, end = NULL, ...)
```

Les principaux arguments sont :

- `formula` : La formule du modèle à estimer, de la forme $y \sim x_1 + x_2 + \dots$, où y est la variable dépendante, et x_1 , x_2 , etc., sont les variables explicatives.
- `data` : Le jeu de données à utiliser, qui doit être au format temporel (`ts`, `zoo`, ou `data.frame`). Dans notre cas, nous utilisons les données `denmark` fournies avec le package.
- `order` : Un vecteur numérique indiquant les ordres de retard. Le premier élément correspond aux retards de la variable dépendante, les suivants à ceux des variables explicatives, dans l'ordre spécifié par `formula`.
- `start` (optionnel) : Permet de fixer le début de la période d'estimation, par exemple "1975 Q4" ou `c(1975, 4)` pour démarrer au quatrième trimestre 1975.
- `end` (optionnel) : Permet de fixer la fin de la période d'estimation.

Exemple utilisé dans cette étude

```
1 data(denmark)
2 ardl_3132 <- ardl(LRM ~ LRY + IBO + IDE, data = denmark, order = c(3,1,3,2))
```

Ce modèle estime l'effet dynamique du revenu réel (`LRY`), des taux obligataires (`IBO`) et des taux de dépôt (`IDE`) sur la demande réelle de monnaie (`LRM`).

Estimation du modèle ARDL(3,1,3,2)

L'estimation du modèle ARDL avec les retards spécifiés se fait avec la commande suivante :

```
1 ardl_3132 <- ardl(LRM ~ LRY + IBO + IDE, data = denmark, order = c(3,1,3,2))
2 summary(ardl_3132)
```

Interprétation des résultats

Les principaux résultats sont présentés dans le tableau ci-dessous :

| Variable | Coefficient | Std. Error | t-Stat | p-value |
|-------------|-------------|------------|--------|----------|
| (Intercept) | 2.6202 | 0.5678 | 4.615 | 4.19e-05 |
| LRM_{t-1} | 0.3192 | 0.1367 | 2.336 | 0.0247 |
| LRM_{t-2} | 0.5326 | 0.1324 | 4.024 | 0.0003 |
| LRM_{t-3} | -0.2687 | 0.1021 | -2.631 | 0.0121 |
| LRY | 0.6728 | 0.1312 | 5.129 | 8.32e-06 |
| LRY_{t-1} | -0.2574 | 0.1472 | -1.749 | 0.0881 |
| IBO | -1.0785 | 0.3217 | -3.353 | 0.0018 |
| IBO_{t-1} | -0.1062 | 0.5858 | -0.181 | 0.8571 |
| IBO_{t-2} | 0.2877 | 0.5691 | 0.505 | 0.6161 |
| IBO_{t-3} | -0.9947 | 0.3925 | -2.534 | 0.0154 |
| IDE | 0.1255 | 0.5545 | 0.226 | 0.8222 |
| IDE_{t-1} | -0.3280 | 0.7213 | -0.455 | 0.6518 |
| IDE_{t-2} | 1.4079 | 0.5520 | 2.550 | 0.0148 |

- Les variables retardées de LRM sont toutes significatives, car leurs p-valeurs sont inférieures à 0,05. Cela indique une dynamique d'ajustement persistante dans le temps.
- LRY a un effet positif significatif sur LRM , en accord avec la théorie de la demande de monnaie.
- L'effet de IBO est globalement négatif et significatif aux périodes t et $t - 3$, indiquant un effet de substitution.
- IDE n'a d'effet significatif que sous la forme retardée à $t - 2$, ce qui pourrait traduire un effet différé ou un manque de sensibilité immédiate.

Création d'une variable muette (dummy) fixe

```
1 d_74Q1_75Q3 <- ifelse(time(denmark) >= 1974 & time(denmark) <= 1975.5, 1, 0)
2 d_74Q1_75Q3_ <- ifelse(time(denmark) >= "1974 Q1" & time(denmark) <= "1975 Q3", 1, 0)
```

Cette commande crée une variable muette `d_74Q1_75Q3` qui vaut 1 pour les trimestres entre 1974 T1 et 1975 T3 (inclus), et 0 sinon.

Cela permet de capter un effet fixe (choc, événement particulier) sur cette période.

Ajout de la variable dummy au jeu de données

La commande suivante combine la variable `d_74Q1_75Q3` avec le jeu de données `denmark` en utilisant la fonction `cbind()`, qui permet d'ajouter des colonnes à un jeu de données. Le résultat est stocké dans un nouveau jeu de données appelé `den`.

```
1 den <- cbind(denmark, d_74Q1_75Q3)
```

Autrement dit, `cbind()` concatène horizontalement les colonnes de `denmark` et de `d_74Q1_75Q3` pour former un nouveau tableau de données.

Estimation d'un modèle ARDL avec la variable dummy

La commande suivante estime un modèle ARDL en incluant la variable dummy `d_74Q1_75Q3` comme régresseur :

```
ardl_3132_d <- ardl(LRM ~ LRY + IBO + IDE | d_74Q1_75Q3,
                   data = den, order = c(3,1,3,2))
```

La formule du modèle est spécifiée comme suit :

$$\text{LRM} \sim \text{LRY} + \text{IBO} + \text{IDE} \mid \text{d_74Q1_75Q3}$$

où la barre verticale `|` sert à séparer les variables qui entrent dans la relation de cointégration (ici `LRM`, `LRY`, `IBO`, `IDE`) de la variable dummy `d_74Q1_75Q3`, qui est traitée comme une variable exogène fixe dans le modèle. Les résultats sont présentés dans le tableau ci-dessous :

| Variable | Coefficient |
|--------------------------|-------------|
| (Intercept) | 4.19447 |
| $L(\text{LRM}, 1)$ | 0.16657 |
| $L(\text{LRM}, 2)$ | 0.49903 |
| $L(\text{LRM}, 3)$ | -0.11428 |
| LRY | 0.44941 |
| $L(\text{LRY}, 1)$ | -0.21907 |
| IBO | -1.44100 |
| $L(\text{IBO}, 1)$ | -0.23277 |
| $L(\text{IBO}, 2)$ | 0.17368 |
| $L(\text{IBO}, 3)$ | -0.91649 |
| IDE | 0.41237 |
| $L(\text{IDE}, 1)$ | -0.69134 |
| $L(\text{IDE}, 2)$ | 1.38546 |
| <code>d_74Q1_75Q3</code> | -0.05955 |

La variable dummy `d_74Q1_75Q3`, qui correspond à la période spécifique entre le premier trimestre 1974 et le troisième trimestre 1975, a un coefficient légèrement négatif (-0,05955). Cela indique un effet négatif faible mais notable sur `LRM` durant cette période, probablement lié à un choc économique ou un événement particulier affectant la série à ce moment.

3.6.4 Présentation de la fonction `auto_ardl()` et de ses composantes

`auto_ardl` recherche la meilleure spécification d'ordre d'un modèle ARDL, selon le critère choisi, tout en respectant les contraintes éventuelles.

Valeur retournée

La fonction `auto_ardl` renvoie une liste contenant :

| Argument | Description |
|----------------------|-----------------------------|
| <code>formula</code> | Spécification de la formule |
| <code>data</code> | Jeu de données |

TABLE 3.7 – Description des arguments

- **best_model** : un objet de classes `c("dynlm", "lm", "ardl")` représentant le meilleur modèle trouvé.
- **best_order** : un vecteur numérique donnant l'ordre retenu.
- **top_orders** : un `data.frame` listant les 20 meilleures spécifications.

Usage

```

1 auto_ardl(
2   formula,
3   data,
4   max_order,
5   fixed_order = -1,
6   starting_order = NULL,
7   selection = "AIC",
8   selection_minmax = c("min", "max"),
9   grid = FALSE,
10  search_type = c("horizontal", "vertical"),
11  start = NULL,
12  end = NULL,
13  ...)

```

Exemple d'utilisation de `auto_ardl()` avec différentes options

```

# Sélection du modèle ARDL avec auto_ardl() -----
library(ARDL)
data(denmark)

# Modèle de base avec recherche horizontale et critère AIC
model1 <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                  max_order = c(5,4,4,4))
print(model1$top_orders)

# Recherche verticale du modèle
model1_h <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                   max_order = c(5,4,4,4),
                   search_type = "vertical")
model1_h

# Recherche de l'optimum global (grid = TRUE)
model_grid <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                     max_order = c(5,4,4,4), grid = TRUE)
model_grid

```

```

# Critères de sélection différents
modell_b <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                    max_order = c(5,4,4,4), selection = "BIC")
modell_b

# Critère personnalisé : R2 ajusté
adjr2 <- fonction(x) { summary(x)$adj.r.squared }
modell_adjr2 <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                       max_order = c(5,4,4,4),
                       selection = "adjr2",
                       selection_minmax = "max")
modell_adjr2

# Configuration des ordres de départ
modell_so <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                    max_order = c(5,4,4,4),
                    starting_order = c(1,1,3,2))
modell_so

# Contraintes sur les ordres des variables
modell_ibo2 <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                      max_order = c(5,4,4,4),
                      fixed_order = c(-1,-1,2,-1))
modell_ibo2

# Début de la régression à partir de "1976 Q1"
modell_76q1 <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark,
                      max_order = c(5,4,4,4),
                      start = "1976 Q1")
modell_76q1

```

Interprétation des résultats

avec le critère AIC : Le modèle ARDL(3,1,3,2) ressort toujours comme le meilleur selon le critère AIC, avec la même valeur (-251,03), quelle que soit la méthode de recherche utilisée (horizontale, verticale ou exhaustive) ce qui confirme de la robustesse et de la fiabilité de la procédure de sélection du modèle ARDL.

Avec le critère BIC : Le modèle ARDL sélectionné selon le critère BIC est très parcimonieux, avec un ordre (1,0,0,0), c'est-à-dire un seul retard pour la variable dépendante LRM et aucun pour les variables explicatives. Ce choix reflète la nature du BIC, qui pénalise fortement les modèles complexes afin de limiter le sur-ajustement.

Avec le critère du R² ajusté : Le modèle ARDL(3,1,3,2) est celui qui explique le mieux les données, avec un R² ajusté de 0,984. Cela veut dire qu'il explique environ 98,4% des

variations de la variable LRM. Ce très bon résultat confirme que ce modèle est très adapté pour représenter la relation entre les variables, comme l'avait déjà montré le critère AIC.

En restreignant la période d'analyse : Le meilleur modèle obtenu est ARDL(2,0,4,3), estimé à partir de 1976 Q1. Les coefficients diffèrent de ceux obtenus sur l'ensemble de la période : par exemple, l'effet du revenu réel (LRY) devient plus faible (0,35009), tandis que l'effet du taux d'intérêt (IBO) devient plus négatif (-1,60481). De plus, la valeur de l'AIC (-232,34) est moins bonne que celle obtenue sur toute la période, ce qui s'explique par la réduction du nombre d'observations. Ce résultat montre que le choix de la période d'estimation influence la dynamique capturée par le modèle.

Fixation de l'ordre de retard dans un modèle ARDL : Fixer l'ordre de retard d'une ou plusieurs variables dans un modèle ARDL consiste à imposer des contraintes basées sur des hypothèses théoriques ou des objectifs d'interprétation. Cette méthode limite la recherche du meilleur modèle aux configurations respectant ces contraintes. Par exemple, en fixant l'ordre de retard de la variable IBO à 2, le meilleur modèle trouvé est ARDL(2,2,2,2) avec un AIC de -245,77, proche du modèle libre. En fixant aussi l'ordre de LRM à 3, le modèle ARDL(3,1,2,2) est sélectionné avec un AIC de -245,10.

Coefficients du meilleur modèle ARDL(3,1,3,2)

| Variable | (Intercept) | L(LRM,1) | L(LRM,2) | L(LRM,3) | LRY | L(LRY,1) | IBO |
|-------------|-------------|----------|----------|----------|----------|----------|---------|
| Coefficient | 2.6202 | 0.3192 | 0.5326 | -0.2687 | 0.6728 | -0.2574 | -1.0785 |
| Variable | L(BO,1) | L(BO,2) | L(BO,3) | IDE | L(IDE,1) | L(IDE,2) | |
| Coefficient | -0.1062 | 0.2877 | -0.9947 | 0.1255 | -0.3280 | 1.4079 | |

TABLE 3.8 – Coefficients estimés du modèle ARDL(3,1,3,2)

Test de cointégration ARDL : *Bounds Wald-test for no cointegration*

3.6.5 Test des bornes de cointégration avec `bounds_f_test`

La fonction `bounds_f_test` dans R permet d'effectuer le test de cointégration par les bornes développé par Pesaran, Shin et Smith (2001). Ce test est basé sur un test de Wald appliqué aux coefficients d'un modèle à correction d'erreur non restreint (UECM). Il peut produire une statistique de type F ou Chi^2 .

Syntaxe de la fonction

```
bounds_f_test(
  object,
  case,
  alpha = NULL,
  pvalue = TRUE,
  exact = FALSE,
```

```

R = 40000,
test = c("F", "Chisq"),
vcov_matrix = NULL
)

```

Arguments

- `object` : un objet de classe `ardl` ou `uecm`.
- `case` : un entier de 1 à 5, ou une chaîne de caractères, indiquant la structure de l'équation de long terme (voir ci-dessous).
- `alpha` : niveau de signification (par exemple 0.01 ou 0.05).
- `pvalue` : booléen indiquant si la *p-value* doit être renvoyée (TRUE par défaut).
- `exact` : booléen indiquant si les bornes doivent être calculées pour la taille exacte de l'échantillon (FALSE par défaut = bornes asymptotiques).
- `R` : nombre d'itérations Monte Carlo si `exact = TRUE`.
- `test` : choix entre les statistiques "F" ou "Chisq".

Les 5 cas possibles (argument `case`)

- **Cas 1** : Pas d'intercept, pas de tendance (1 ou "n").
- **Cas 2** : Intercept restreint, pas de tendance (2 ou "rc").
- **Cas 3** : Intercept non restreint, pas de tendance (3 ou "uc").
- **Cas 4** : Intercept non restreint, tendance restreinte (4 ou "ucrt").
- **Cas 5** : Intercept et tendance non restreints (5 ou "ucut").

Construction d'un modèle ARDL(3,1,3,2) avec différents termes déterministes

Construction des modèles ARDL

Modèle sans constante

```

1 ardl_3132_n <- ardl(LRM ~ LRY + IBO + IDE -1, data = denmark, order = c(3,1,3,2))
2 bounds_f_test(ardl_3132_n, case = 1)
3 ou
4 bounds_f_test(ardl_3132_n, case = "n")

```

Listing 3.1 – ARDL sans constante (case 1)

Modèle avec constante :

```

1 ardl_3132_c <- ardl(LRM ~ LRY + IBO + IDE, data = denmark, order = c(3,1,3,2))
2 Case 2 : avec constante restreinte
3 bounds_f_test(ardl_3132_c, case = 2)
4 ou
5 bounds_f_test(ardl_3132_c, case = "rc")
6 Case 3 : avec constante non restreinte
7 bounds_f_test(ardl_3132_c, case = 3)
8 ou
9 bounds_f_test(ardl_3132_c, case = "uc")

```

Listing 3.2 – ARDL avec constante

```

Modèle avec constante et tendance :
1 ardl_3132_ct <- ardl(LRM ~ LRY + IBO + IDE + trend(LRM), data = denmark, order = c
  (3,1,3,2))
2 Case 4 : constante non restreinte, tendance restreinte
3 bounds_f_test(ardl_3132_ct, case = 4)
4 ou
5 bounds_f_test(ardl_3132_ct, case = "ucrt")
6 Case 5 : constante et tendance non restreintes
7 bounds_f_test(ardl_3132_ct, case = 5)
8 ou
9 bounds_f_test(ardl_3132_ct, case = "ucut")

```

Listing 3.3 – ARDL avec constante et tendance (case 4 et 5)

TABLE 3.9 – Résultats du test de cointégration par la méthode des bornes selon différentes spécifications déterministes

| Cas | Spécification | Syntaxe R | F-statistique | p-value | Interprétation préliminaire |
|-----|--|--------------------|---------------|----------|-----------------------------|
| 1 | Sans constante | case = 1 ou "n" | 0.7109 | 0.8765 | Pas de cointégration |
| 2 | Constante restreinte (long terme) | case = 2 ou "rc" | 5.1168 | 0.004418 | Cointégration probable |
| 3 | Constante non restreinte (court terme) | case = 3 ou "uc" | 6.2059 | 0.004146 | Cointégration forte |
| 4 | Constante non restreinte, tendance restreinte (LT) | case = 4 ou "ucrt" | 5.4306 | 0.007069 | Cointégration présente |
| 5 | Constante + tendance non restreintes (CT) | case = 5 ou "ucut" | 6.7853 | 0.005303 | Cointégration très forte |

Test de cointégration par la méthode des bornes de Pesaran et al. (2001)

Afin de vérifier l'existence d'une relation d'équilibre à long terme entre les variables de notre modèle ARDL(3,1,3,2) avec constante restreinte (cas 2), nous avons appliqué le test des bornes proposé par Pesaran et al. (2001). Ce test permet de détecter la cointégration en comparant la statistique de test à des valeurs critiques calculées selon différents niveaux de signification.

Programme R utilisé

```

# Test asymptotique au seuil 5%
bft <- bounds_f_test(ardl_3132_c, case = 2, alpha = 0.05)
print(bft)
print(bft$tab)
print(bft$PSS2001parameters)

# Test au seuil strict 0.5%
bounds_f_test(ardl_3132_c, case = 2, alpha = 0.005)

# Test exact avec taille d'échantillon réelle (simulation)

```

```

set.seed(2020)
bounds_f_test(ardl_3132_c, case = 2, alpha = 0.01, exact = TRUE)

# Comparaison entre statistique F et Khi-deux au seuil 1%
bounds_f_test(ardl_3132_c, case = 2, alpha = 0.01)
bounds_f_test(ardl_3132_c, case = 2, alpha = 0.01, test = "Chisq")

```

Résultats obtenus

| Test | Statistique | Borne I(0) | Borne I(1) | Alpha | p-valeur |
|---------------------------|-------------|------------|------------|-------|----------|
| Asymptotique (F) | 5.1168 | 2.7750 | 3.6595 | 0.05 | 0.0044 |
| Asymptotique (F) | 5.1168 | 4.0491 | 5.0372 | 0.005 | 0.0044 |
| Exact (F) | 5.1168 | 4.1335 | 5.2902 | 0.01 | 0.01285 |
| Asymptotique (F) | 5.1168 | 3.6744 | 4.6104 | 0.01 | 0.0044 |
| Asymptotique (χ^2) | 25.584 | 18.372 | 23.052 | 0.01 | 0.0044 |

TABLE 3.10 – Résultats du test des bornes de cointégration pour le modèle ARDL(3,1,3,2) (cas 2)

Interprétation

Les résultats montrent que la statistique de test (F ou Khi-deux) est systématiquement supérieure à la borne critique supérieure I(1) pour les niveaux de signification usuels (5%, 1%, 0,5%). Cela signifie que l'on rejette l'hypothèse nulle d'absence de cointégration, et donc que les variables étudiées partagent une relation d'équilibre stable à long terme.

Le test exact, qui prend en compte la taille réelle de l'échantillon via une simulation Monte Carlo, donne une p-valeur légèrement plus élevée (0,01285) mais reste significative au seuil de 5%. Cette différence souligne l'importance de considérer la taille d'échantillon dans l'interprétation des tests.

Enfin, la comparaison entre la statistique F et la statistique Khi-deux confirme la robustesse des résultats, les deux tests menant à la même conclusion.

Conclusion

Le test des bornes a confirmé l'existence d'une relation de cointégration entre la variable dépendante *LRM* et les variables explicatives *LRY*, *IBO*, et *IDE* dans le modèle ARDL(3,1,3,2) avec constante restreinte. Cette cointégration signifie qu'il existe un équilibre à long terme stable liant ces variables, malgré leurs fluctuations à court terme.

3.6.6 Test des bornes basé sur la statistique t (`bounds_t_test`)

Pour compléter le test des bornes basé sur la statistique F, nous utilisons également le test des bornes basé sur la statistique t. Ce test permet de vérifier la présence d'une relation de cointégration en se focalisant spécifiquement sur le coefficient retardé de la variable dépendante dans le modèle ARDL.

Principe Le test évalue si ce coefficient est significativement différent de zéro, ce qui indique une relation d'équilibre à long terme entre les variables.

Cas de spécification Comme pour le test F, différents cas sont possibles selon la présence ou non de constante et/ou tendance dans la relation de long terme :

- Cas 1 : sans constante ni tendance
- Cas 3 : constante non restreinte, pas de tendance
- Cas 5 : constante et tendance non restreintes

Le test t ne s'applique pas aux cas 2 et 4.

Utilisation dans R

```
bounds_t_test(ardl_3132_c, case = 3, alpha = 0.05)
```

Application du test des bornes basé sur la statistique t (bounds_t_test)

Programme R utilisé

```
# Test asymptotique au seuil 5%
btt <- bounds_t_test(ardl_3132_c, case = 3, alpha = 0.05)
print(btt)
print(btt$tab)
print(btt$PSS2001parameters)

# Test asymptotique au seuil 0.5%
bounds_t_test(ardl_3132_c, case = 3, alpha = 0.005)

### Test exact avec taille d'échantillon réelle (simulation)
set.seed(2020)
bounds_t_test(ardl_3132_c, case = 3, alpha = 0.01, exact = TRUE)
```

Résultats obtenus

TABLE 3.11 – Résultats du test bounds_t_test pour différents seuils

| Test | α | Statistique t | I(0) | I(1) | Décision |
|--------------|----------|---------------|---------|---------|-------------------------------------|
| Asymptotique | 5% | -4.5479 | -2.8650 | -3.7713 | Rejet de H_0 ($p = 0.0055$) |
| Asymptotique | 0.5% | -4.5479 | -3.6595 | -4.5708 | Non-rejet de H_0 ($p = 0.0055$) |
| Exact (T=52) | 1% | -4.5479 | -3.5572 | -4.5414 | Rejet de H_0 ($p = 0.0099$) |

Interprétation

Les résultats du test bounds_t_test montrent que :

- Au seuil de 5%, la statistique t est plus négative que la borne supérieure I(1), et la p-valeur est inférieure à 0.05. On rejette donc l'hypothèse nulle d'absence de cointégration, ce qui indique la présence d'une relation d'équilibre à long terme entre les variables du modèle ARDL.

- Au seuil très strict de 0.5%, la statistique t ne dépasse pas la borne critique $I(1)$ et la p -valeur est légèrement supérieure à 0.005, ce qui ne permet pas de rejeter l'hypothèse nulle. La preuve de cointégration est donc insuffisante à ce niveau très exigeant.
- Le test exact, qui prend en compte la taille réelle de l'échantillon ($T=52$) par simulation, confirme la présence de cointégration au seuil de 1%, avec une statistique t plus négative que la borne critique exacte et une p -valeur inférieure à 0.01.

Ces résultats soulignent l'importance de considérer la taille de l'échantillon dans l'interprétation des tests de cointégration. Le test exact est plus fiable pour des échantillons de taille modeste.

Conclusion

Les résultats obtenus à travers les tests des bornes (bounds F-test et bounds t-test) confirment l'existence d'une relation de long terme stable entre la variable dépendante **LRM** et les variables explicatives **LRY**, **IBO** et **IDE** dans notre modèle ARDL. Ces conclusions sont cohérentes avec celles des travaux classiques de Johansen et Juselius, qui, par leur méthode basée sur le test de rang de cointégration, identifient généralement au moins un vecteur de cointégration dans des données économiques comparables.

Application 2

3.7 Exemple appliqué sur des données simulées d'un modèle ARDL avec détails sur l'estimation

Problématique : Nous souhaitons étudier l'impact à court et long terme de la consommation d'électricité (CONSO) et du prix de l'électricité (PRIX) sur la demande industrielle (DEM), en utilisant des données trimestrielles (2000–2020).

Étapes de l'estimation ARDL

1. Vérification des ordres d'intégration (Test de racine unitaire)

Avant d'appliquer le modèle ARDL, il faut vérifier que les variables ne sont pas $I(2)$, sinon le modèle n'est pas valide.

Tests ADF (Augmented Dickey-Fuller) :

- DEM : est $I(1)$ (non stationnaire en niveau, stationnaire en différences premières).
- CONSO : est $I(1)$.
- PRIX : est $I(0)$ (stationnaire en niveau).

Le modèle ARDL est applicable car il y a une mixité de $I(0)$ et $I(1)$.

2. Sélection des Retards Optimaux(p, q)

Le modèle $ARDL(p, q_1, q_2)$ doit être estimé avec des retards optimaux pour éviter le surajustement.

Critères d'information (AIC / BIC / HQ) : On teste différentes combinaisons de retards (p pour DEM, q_1 pour CONSO, q_2 pour PRIX). Le meilleur modèle, selon l'AIC, est $ARDL(2, 1, 0)$.

3. Estimation du Modèle ARDL : Spécification du modèle

$$DEM_t = \alpha + \beta_1 DEM_{t-1} + \beta_2 DEM_{t-2} + \gamma_0 CONSO_t + \gamma_1 CONSO_{t-1} + \delta_0 PRIX_t + \varepsilon_t \quad (3.7.1)$$

Résultats (exemple simulé) :

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|-----------|-------------|------------|-------------|-------|
| C | 10.5* | 2.1 | 5.0 | 0.000 |
| DEM(-1) | 0.6* | 0.1 | 6.0 | 0.000 |
| DEM(-2) | -0.3* | 0.1 | -3.0 | 0.004 |
| CONSO | 0.8* | 0.2 | 4.0 | 0.000 |
| CONSO(-1) | -0.5* | 0.2 | -2.5 | 0.016 |
| PRIX | -0.2* | 0.05 | -4.0 | 0.000 |

* significatif à 5%.

4. Test de Cointégration (Bounds Test)

L'hypothèse nulle (H_0 : pas de cointégration) est testée via les seuils critiques de Pesaran et al. (2001).

- La statistique F calculée est égale à 6.2.
- Les seuils critiques (à 5%) sont :
 - Lower Bound (I(0)) : 3.2
 - Upper Bound (I(1)) : 4.5

Puisque $F = 6.2 > 4.5$, c'est à dire la valeur calculée de la statistique F est supérieure à la valeur de la borne inférieure, on rejette H_0 : par conséquent la cointégration est confirmée.

5. Estimation de l'Équation de Long Terme

On extrait les coefficients de long terme du modèle :

$$DEM_t = \Theta_0 + \theta_1 CONSO_t + \theta_2 PRIX_t + \eta_t$$

Calcul :

$$\theta_1 = -\frac{\gamma_0 + \gamma_1}{1 - \beta_1 - \beta_2} = -\frac{0.8 - 0.5}{1 - 0.6 + 0.3} = 0.43$$

$$\theta_2 = -\frac{\delta_0}{1 - \beta_1 - \beta_2} = -\frac{-0.2}{1 - 0.6 + 0.3} = -0.29$$

Interprétation :

- À long terme, une hausse de 1% de **CONSO** accroît **DEM** de **0.43%**.
- À long terme, une hausse de 1% de **PRIX** réduit **DEM** de **0.29%**.

6. Correction d'erreur (ECM : Error Correction Model)

Pour analyser la dynamique de court terme, on formule le modèle à correction d'erreur :

$$\begin{aligned} \Delta DEM_t = & \alpha + \sum_{i=1}^{p-1} \phi_i \Delta DEM_{t-i} + \sum_{j=1}^{q_1-1} \psi_j \Delta CONSO_{t-j} \\ & + \sum_{k=0}^{q_2-1} \delta_k \Delta PRIX_{t-k} + \rho ECT_{t-1} + u_t \end{aligned} \quad (3.7.2)$$

Résultat :

$$\rho = -0.4 \quad (\text{p-value} = 0,01)$$

La vitesse d'ajustement est de 40% du déséquilibre corrigé chaque trimestre.
Le coefficient ρ doit être **négatif et significatif** (ajustement vers l'équilibre).

7. Validation du Modèle**Tests de diagnostic**

- **Test d'autocorrélation (Breusch-Godfrey)** : $p > 0,05$, on accepte H_0 .
- **Hétéroscédasticité (White)** : $p > 0,05$
- **Normalité des résidus (Test Jarque-Bera)** : $p > 0,05$

Remarques :

- À court terme, la consommation et le prix ont un impact immédiat sur la demande.
- À long terme, il existe une relation d'équilibre : le modèle est stable de correction.
- L'ajustement vers l'équilibre se fait à 40% par période.

Implémentation R : En utilisant le package ARDL du langage R, on procède comme suit :

```
model <- ardl(DEM ~ CONSO + PRIX, data = data, order = c(2,1,0))
bounds_test(model)
```

Voici un programme R complet qui simule les données CONSO, PRIX et DEM utilisées dans l'exemple, puis reproduit toutes les étapes de l'analyse ARDL.

1. Simulation des données du premier trimestre 2000 au quatrième trimestre 2020

```
set.seed(123) # Pour la reproductibilité

# Créer un index temporel trimestriel (Calendrier)
```

```

dates <- seq(as.Date("2000-01-01"), as.Date("2020-12-31"),
  by = "quarter")
n <- length(dates)

# Simuler CONSO (I(1))
CONSO <- cumsum(rnorm(n, mean = 0.5, sd = 1)) + 100 # Trend + bruit

# Simuler PRIX (I(0))
PRIX <- arima.sim(model = list(ar = 0.7), n = n, mean = 50, sd = 2)

# Simuler DEM (dépend de CONSO et PRIX avec retard)
DEM <- rep(NA, n)
DEM[1:2] <- 200 # Initialisation for t in 3:n
for (t in 3:n) {
  DEM[t] <- 10 + 0.6*DEM[t-1] - 0.3*DEM[t-2] + 0.8*CONSO[t] -
  0.5*CONSO[t-1] - 0.2*PRIX[t] + rnorm(1, sd = 3)
}

# Créer un dataframe
data <- data.frame(
  Date = dates,
  DEM = DEM,
  CONSO = CONSO,
  PRIX = PRIX
)

# Visualisation
library(ggplot2)
ggplot(data, aes(x = Date)) +
  geom_line(aes(y = DEM, color = "DEM")) +
  geom_line(aes(y = CONSO/2, color = "CONSO (scaled)")) + # Mise à
  l'échelle pour la lisibilité
  geom_line(aes(y = PRIX*2, color = "PRIX (scaled)")) +
  scale_color_manual(values = c("DEM" = "blue",
  "CONSO (scaled)" = "red", "PRIX (scaled)" = "green")) +
  labs(title = "Données Simulées (2000-2020)", y = "Valeur",
  color = "Variable") + theme_minimal()

```

2. Vérification des ordres d'intégration (ADF Test)

```

library(urca)

# Test ADF pour DEM (avec constante et tendance)
summary(ur.df(data$DEM, type =
  "trend", lags = 4)) # Niveau, c à d Non stationnaire

# Test ADF pour DEM (différenciée)

```

```
summary(ur.df(diff(data$DEM), type = "drift", lags = 4))
# Différence c à d Stationnaire =>I(1)

# Test ADF pour CONSO
summary(ur.df(data$CONSO, type = "trend", lags = 4)) # I(1)

# Test ADF pour PRIX
summary(ur.df(data$PRIX, type = "drift", lags = 4)) # I(0)
```

3. Estimation ARDL (Package 'ARDL')

```
library(ARDL)

# Sélection automatique des retards (AIC)
model_scl <- auto.ardl(DEM ~ CONSO + PRIX, data = data, max_order = c(3,3,3))
bcst_ordcr <- model_scl$bcst_ordcr # Retient ARDL(2,1,0)

# Estimation du modèle
model <- ardl(DEM ~ CONSO + PRIX, data = data, order = c(2,1,0))
summary(model)
```

4. Test de Cointégration (Bounds Test)

```
bounds.test(model)
# Résultat attendu : F-stat > seuil I(1) alors il y a cointégration
```

5. Calcul des Coefficients de Long Terme

```
multiplier(model, type = "lr") # Output:
# 1 CONSO 0.43
# 2 PRIX -0.29
```

6. Modèle à correction d'erreur (ECM)

```
ecm <- ECM(model)
summary(ecm)
# Vérifier que le coefficient ECT est négatif et significatif (ex: -0.4)
```

7. Validation du modèle

```
# Autocorrélation (H0 : pas d'autocorrélation)
# Dans le package lmtest, on utilise la commande bgtest.
bgtest(model, order = 2) # p > 0.05, il n'y a pas autocorrélation.

# Hétéroscédasticité (H0 : homoscedasticité)
# Dans le package lmtest, on utilise la commande bptest.
bptest(model) # p > 0.05, on accepte H0.

# Normalité des résidus
```

Dans le package `tseries`, on utilise la commande `jarque.bera.test.jarque.bera.test(residuals(model))` # $p > 0.05$, on accepte H_0 .

Explication de la méthodologie

Les séries sont simulées comme suit :

CONSO : Simulée comme une marche aléatoire (I(1)) avec tendance.

PRIX : Simulé comme un processus AR(1) stationnaire (I(0)).

La variable **DEM** est générée selon le modèle ARDL(2,1,0) avec :

- Effets retardés de DEM : $(0.6 \times DEM_{t-1} - 0.3 \times DEM_{t-2})$.
- Impact contemporain et retardé de CONSO : $(0.8 \times CONSO_t - 0.5 \times CONSO_{t-1})$.
- Impact contemporain de PRIX : $(-0.2 \times PRIX_t)$.

Choix des paramètres

Le choix des coefficients est cohérent avec :

1. La théorie économique (élasticité prix négative, revenu positive).
2. Les propriétés des séries (I(0)/I(1)).
3. La stabilité du modèle ARDL (racines du polynôme retard < 1).

Explication détaillée des coefficients de long terme (θ_1 et θ_2)

Dans un modèle ARDL, les coefficients de long terme (θ) représentent l'effet d'équilibre stable entre les variables. Leur calcul vient de la relation d'équilibre de long terme dérivée du modèle dynamique estimé.

Origine de la formule : Réécriture du modèle ARDL en équilibre de long terme

Prenons un modèle ARDL(2,1,0) général :

$$DEM_t = \alpha + \beta_1 DEM_{t-1} + \beta_2 DEM_{t-2} + \gamma_0 CONSO_t + \gamma_1 CONSO_{t-1} + \delta_0 PRIX_t + \varepsilon_t$$

Étape 1 : Supposons l'équilibre de long terme (steady state) En long terme, les variables convergent vers des valeurs stables :

$$\begin{aligned} DEM_t &= DEM_{t-1} = DEM_{t-2} = \overline{DEM} \\ CONSO_t &= CONSO_{t-1} = \overline{CONSO} \\ PRIX_t &= \overline{PRIX} \quad \text{et} \quad \varepsilon_t = 0 \end{aligned}$$

Le modèle devient :

$$\overline{DEM} = \alpha + \beta_1 \overline{DEM} + \beta_2 \overline{DEM} + \gamma_0 \overline{CONSO} + \gamma_1 \overline{CONSO} + \delta_0 \overline{PRIX}$$

Étape 2 : Regrouper les termes Factorisons \overline{DEM} et \overline{CONSO} :

$$\overline{DEM}(1 - \beta_1 - \beta_2) = \alpha + (\gamma_0 + \gamma_1) \overline{CONSO} + \delta_0 \overline{PRIX}$$

Étape 3 : Isoler \overline{DEM} pour obtenir la relation de long terme

$$\overline{DEM} = \frac{\alpha}{1 - \beta_1 - \beta_2} + \frac{\gamma_0 + \gamma_1}{1 - \beta_1 - \beta_2} \overline{CONSO} + \frac{\delta_0}{1 - \beta_1 - \beta_2} \overline{PRIX}$$

Ce qui donne la relation de long terme :

$$\overline{DEM} = \theta_0 + \theta_1 \overline{CONSO} + \theta_2 \overline{PRIX}$$

Avec :

$$\begin{aligned} \theta_0 &= \frac{\alpha}{1 - \beta_1 - \beta_2} && \text{(terme constant)} \\ \theta_1 &= \frac{\gamma_0 + \gamma_1}{1 - \beta_1 - \beta_2} && \text{(effet de long terme de CONSO)} \\ \theta_2 &= \frac{\delta_0}{1 - \beta_1 - \beta_2} && \text{(effet de long terme de PRIX)} \end{aligned}$$

Application numérique (d'après l'exemple précédent)

Reprenons les coefficients estimés :

$$\begin{aligned} \beta_1 &= 0.6 && \text{(coefficient de } DEM_{t-1}) \\ \beta_2 &= -0.3 && \text{(coefficient de } DEM_{t-2}) \\ \gamma_0 &= 0.8 && \text{(coefficient de } CONSO_t) \\ \gamma_1 &= -0.5 && \text{(coefficient de } CONSO_{t-1}) \\ \delta_0 &= -0.2 && \text{(coefficient de } PRIX_t) \end{aligned}$$

Calcul de θ_1 (effet de CONSO en long terme)

$$\theta_1 = \frac{\gamma_0 + \gamma_1}{1 - \beta_1 - \beta_2} = \frac{0.3}{0.7} \approx 0.43$$

Calcul de θ_2 (effet de PRIX en long terme)

$$\theta_2 = \frac{\delta_0}{1 - \beta_1 - \beta_2} = \frac{-0.2}{0.7} \approx -0.29$$

Interprétation économique

$\theta_1 = 0.43$: Une augmentation permanente de 1% de la consommation d'électricité (CONSO) entraîne une hausse de 0.43% de la demande industrielle (DEM) à long terme.

$\theta_2 = -0.29$: Une augmentation permanente de 1% du prix de l'électricité (PRIX) réduit la demande industrielle de 0.29% à long terme.

Résultat

Ce calcul est nécessaire dans la mesure où :
Les coefficients $\gamma_0, \gamma_1, \delta_0$ ne donnent que des effets à court terme.

Le modèle ARDL permet de décomposer ces effets en :

- (1) Dynamique de court terme (via les différences premières et le mécanisme de correction d'erreur).
- (2) Relation de long terme (via θ_1, θ_2).

Remarques importantes

1. Dénominateur $(1 - \beta_1 - \beta_2)$:
 - Représente la vitesse d'ajustement vers l'équilibre.
 - Doit être positif et inférieur à 1 pour assurer la stabilité.
2. Cas des retards supérieurs :
 - Si CONSO avait un retard $q = 2$, on sommerait $\gamma_0 + \gamma_1 + \gamma_2$.
 - Si PRIX avait un retard $q = 1$, on sommerait $\delta_0 + \delta_1$.
3. Une erreur souvent commise est d'oublier de diviser par $(1 - \sum \beta_i)$, ce qui fausse l'interprétation des élasticités de long terme.

Résumé des résultats obtenus

| Effet | Formule | Calcul (exemple) |
|-------------|---|------------------------------|
| Court terme | $\gamma_0, \gamma_1, \delta_0$ | $CONSO_t = 0.8$ |
| Long terme | $\theta = \frac{\sum \gamma}{1 - \sum \beta_i}$ | $\theta_1 = \frac{0.3}{0.7}$ |

Conclusion

Les θ sont obtenus en normalisant les coefficients de court terme par la persistance de la variable dépendante $(1 - \sum \beta_i)$.

Conclusion générale

Ce mémoire a permis d'analyser en profondeur les séries chronologiques, depuis les concepts fondamentaux de stationnarité jusqu'aux modèles avancés de cointégration et ARDL, en passant par les tests statistiques essentiels. Sur le plan théorique, les tests de racine unitaire (ADF, PP, KPSS) se sont révélés indispensables pour déterminer la nature des séries étudiées, tandis que les modèles à correction d'erreur (ECM) et les tests de cointégration (Engle-Granger, Johansen) ont permis de capturer les relations de long terme entre variables économiques. L'approche ARDL, plus flexible, a démontré son efficacité pour traiter des séries intégrées d'ordres différents, surmontant ainsi les limitations des méthodes traditionnelles.

Sur le plan empirique, l'application sur les données danoises a confirmé la robustesse du *bounds test* pour identifier des relations de cointégration, même avec des échantillons de taille modérée.

Cette recherche ouvre néanmoins plusieurs perspectives pour de futurs travaux. D'un point de vue méthodologique, l'utilisation de modèles ARDL non linéaires (NARDL) ou bayésiens permettrait de mieux appréhender les éventuelles ruptures structurelles ou asymétries dans les relations économiques. Des applications élargies à d'autres contextes nationaux ou sectoriels pourraient en vérifier la portée générale, tandis que l'intégration de données à plus haute fréquence offrirait une granularité d'analyse accrue.

En définitive, cette étude met en lumière l'importance cruciale du choix des outils économétriques dans l'analyse des séries temporelles. Les différentes méthodes testées constituent un cadre d'analyse cohérent pour étudier les phénomènes économiques dynamiques, comme l'a illustré notre application empirique sur le cas danois. Ces approches, combinées aux développements méthodologiques futurs, présentent des perspectives particulièrement prometteuses pour la recherche appliquée en macroéconomie et en finance, tout en offrant des pistes concrètes pour le travail des décideurs politiques et économiques.

Bibliographie

- [1] Akaike, H. (1974). A new look at the statistical model identification.
- [2] Arnaud, R. (2022). *Séries chronologiques*. Université de Bourgogne.
- [3] Box, G. E. P. et Jenkins, G. M. (1976). *Time Series Analysis : Forecasting and Control*. Revised Edition.
- [4] Daudin, J.-J., Duby, C., Robin, S. & Trécourt, P. (1996). *Analyse de Séries Chronologiques*. INA-PG, Mathématiques, mai 1996.
- [5] Dickey, D. A. et Fuller, W. A. (1979). *Distribution of the Estimators for Autoregressive Time Series With a Unit Root*. *Journal of the American Statistical Association*, **74**(366), 427–431.
- [6] Dickey, D. A. et Fuller, W. A. (1981). *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. *Econometrica*, **49**(4), 1057–1072.
- [7] Dufrenot, G. et Mignon, V. (2002). *Recent Developments in Nonlinear Cointegration with Applications to Macroeconomics and Finance*. Springer-Science.
- [8] Engle, R. F. et Granger, C. W. J. (1987). *Co-Integration and Error Correction : Representation, Estimation, and Testing*. *Econometrica*, **55**(2), 251–276.
- [9] Hamisultane, H. (2002). *Modèle à correction d’erreur (MCE) et applications*. Mémoire de Master, France.
- [10] Johansen, S. (1988). *Statistical Analysis of Cointegration Vectors*. *Journal of Economic Dynamics and Control*, **12**(2-3), 231–254.
- [11] Johansen, S. & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, **52**(2), 169–210.
- [12] Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, **59**(6), 1551-1580.
- [13] Kibala Kuma, J. (2018). *Modélisation ARDL, test de cointégration aux bornes et approche de Toda-Yamamoto : éléments de théorie et pratiques sur logiciels*. Mémoire de licence, République Démocratique du Congo.
- [14] Koop, G. (2003). *Bayesian Econometrics*. Wiley, Chichester.
- [15] Koop, G., Leon-Gonzalez, R., & Strachan, R. (2008). Bayesian inference in a cointegrating panel data model. In *Bayesian Econometrics* (Vol. 23, pp. 433–469). Emerald Group Publishing Limited.
- [16] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. et Shin, Y. (1992). *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 ? Journal of Econometrics*, **54**(1-3), 159–178.
- [17] Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**(3), 703-708.
- [18] Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics*, **54**(3), 461-472.
- [19] Pesaran, M. H., & Pesaran, B. (1997). *Working with Microfit 4.0 : Interactive econometric analysis*. Oxford University Press.
- [20] Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). *Bounds testing approaches to the analysis of level relationships*. *Journal of Applied Econometrics*, **16**(3), 289–326.
- [21] Phillips, P. C. B. et Perron, P. (1988). *Testing for a Unit Root in a Time Series Regression*. *Biometrika*, **75**(2), 335–346.
- [22] Phillips, P. C. B. & Ouliaris, S. (1990). *Asymptotic properties of residual based tests for cointegration*. *Econometrica*, **58**(1), 165–193.
- [23] Schwarz, G. (1978). Estimating the dimension of a model.
- [24] Schwert, G. W. (1989). Tests for unit roots : A Monte Carlo investigation. *Journal of Business & Economic Statistics*, **7**(2), 147-159

Résumé

Ce mémoire étudie les modèles ARDL (Autoregressive Distributed Lag) pour l'analyse des relations entre séries temporelles économiques, en particulier dans les cas où les variables présentent des ordres d'intégration différents ($I(0)$ et $I(1)$). Après avoir présenté les concepts fondamentaux des séries chronologiques et les tests de stationnarité (ADF, Phillips-Perron et KPSS), nous développons une méthodologie complète incluant les approches de cointégration traditionnelles (Engle-Granger et Johansen) et le test aux bornes de Pesaran. Nous comparons ensuite différentes méthodes d'estimation (moindres carrés ordinaires, maximum de vraisemblance et approche bayésienne), en soulignant leurs avantages respectifs. Une application empirique sur les données macroéconomiques du Danemark illustre l'utilisation des modèles ARDL pour modéliser la demande de monnaie, démontrant leur flexibilité pour capturer simultanément les dynamiques de court et long terme tout en évitant les problèmes de régressions fallacieuses. Les résultats confirment la robustesse de cette approche pour identifier des relations stables entre variables économiques, avec des perspectives prometteuses pour les extensions bayésiennes. L'implémentation pratique sous R avec le package ARDL est également discutée.

Abstract

This thesis examines ARDL models (Autoregressive Distributed Lag) for analyzing relationships between economic time series, particularly when variables have different integration orders ($I(0)$ and $I(1)$). After presenting fundamental concepts of time series and stationarity tests (ADF, Phillips-Perron and KPSS), we develop a comprehensive methodology including traditional cointegration approaches (Engle-Granger and Johansen) and Pesaran's bounds test. We then compare different estimation methods (ordinary least squares, maximum likelihood and Bayesian approach), highlighting their respective advantages. An empirical application on Danish macroeconomic data illustrates the use of ARDL models to analyze money demand, demonstrating their flexibility in capturing both short- and long-term dynamics while avoiding spurious regression problems. The results confirm the robustness of this approach for identifying stable relationships between economic variables, with promising prospects for Bayesian extensions. Practical implementation in R using the ARDL package is also discussed.

Mots-clés

Modèles ARDL, cointégration, séries temporelles, test de Pesaran, bounds test, inférence bayésienne, économétrie appliquée.