CHAPITRE 2 : LA COÏNTEGRATION
(TESTS DE RACINES UNITAIRES)

Section 1 : Le point de départ : l'ECM

La coïntégration rime avec séries temporelles (notamment avec la notion d'ARMA [Cf. Box et jenkins (1976)]) et stationnarité.

A l'origine pour estimer une équation dynamique on avait à notre disposition différentes méthodes (Cf. Chapitre 1) mais une seule c'est vraiment généralisée et imposée : le modèle à correction d'erreur (ECM) [cf. Davidson-Hendry-Srba-Yeo (1978)].

Toutefois l'utilité de cette méthode est purement intuitive :

On met en présence des variables que la théorie économique nous impose (exemple la consommation dépend du revenu disponible) mais rien n'est dit ou précisé sur le comportement (stationnarité) de chaque variable.

En principe en économie toutes les variables ne sont pas stationnaires.

La stratégie de la coïntégration permet de rationaliser la modélisation dynamique à court et à long terme.

Dans la première stratégie, la formalisation de type ECM était posée <u>a priori</u>, ici a contrario, Engle & Granger (1987) fondent la coïntégration sur les <u>propriétés stochastiques</u> des variables à modéliser.

Il s'agit de ce fait de modéliser le comportement de variables économiques tendancielles or la non-stationnarité peut provoquer des problèmes dans les résultats des estimations. Avec l'approche ECM ce fait n'est pas pris en compte.

Cette stratégie repose sur la notion de séries coïntégrées et sur le théorème de Granger qui montre l'équivalence entre système coïntégré et modèle à correcteur d'erreurs.

Sur le plan formel on peut introduire la coïntégration en se rappelant que les développements précédents reposaient sur l'hypothèse que $|a| < 1$. Il n'y avait donc pas de racine unitaire.

Si cette hypothèse n'est plus vérifiée (c'est le cas malheureusement pour la plus grande partie des séries économiques) qu'elles vont être les conséquences sur les estimations ?

La présence de racine unitaire dans les données a principalement deux conséquences très importantes :

↳ 1) les propriétés asymptotiques générales des estimateurs ne tiennent plus (vitesse de convergence, normalité asymptotique.........). Il faut avoir recours à une théorie asymptotique spéciale,

↳ 2) la présence de régresseurs comportant une racine unitaire dans une régression peut conduire à estimer des régressions apparemment très bonnes entre des variables qui sont totalement indépendantes entre elles. C'est le problème des régressions factices (spurious regression).

Section 2 : Définition de la coïntégration

L'analyse de la coïntégratrion permet d'identifier la relation véritable entre deux variables (ou n variables) en recherchant l'existence d'un vecteur de coïntégration et en éliminant son effet.

- **On appelle variables intégrées d'ordre d une variable X telle que sa différence dme soit stationnaire et possède une représentation ARMA inversible et on note :**

$$X \sim I(d)$$

Sur le plan formel on a :

$$X \sim I(d) \Leftrightarrow \Delta d\, X = (1 - L)d\, X$$

- **Deux séries X et Y sont dites coïntégrées d'ordre d, b pour $0 < b \leq d$, si :**

1) $X \sim I(d)$ et $Y \sim I(d)$
2) il existe $(\alpha,\beta) \neq 0$ tel que $Z = \alpha X + \beta Y \sim I(d\text{-}b)$

(α,β) est dit vecteur coïntégrant. On adopte la notation : $(X,Y) \sim CI(d,b)$. Cette définition peut être généralisée à plus de deux séries.

Dans le cas particulier d = b = 1, la coïntégration traduit le fait que la combinaison linéaire Z ne s'éloigne jamais très longtemps de sa moyenne, bien que les séries X et Y présentent des mouvements tendanciels.

Dans <u>le cas de deux séries, il y a au plus un vecteur coïntégrant</u> mais il peut y avoir <u>(n-1) vecteurs coïntégrant un ensemble de n séries</u>. Dans le vocabulaire des ECM, le vecteur Z va jouer le rôle de l'écart à la cible de long terme. Le lien précis entre ces deux approches est établi par le théorème de représentation de Granger.

Section 3 : Les différents aspects de la coïntégration

3.1. Les processus TS et les processus DS

Nous recherchons ici la bonne manière de stationnariser une série. Deux types de processus sont distingués :

- **Les processus TS (Trend Stationnary)**

qui représentent une non-stationnarité de type déterministe. Si on suppose une relation linéaire polynomiale de degré 1, le processus TS s'écrit :

$$Y_t = cste + a*t + \varepsilon_t$$

Avec t le trend, ε_t un processus bruit blanc.

Pour stationnariser ce processus il faut estimer le processus par OLS (on obtient $\hat{cste}$ + $\hat{a}*t$).

- **Les processus DS (Differency Stationnary)**

qui représentent une non-stationnarité de type aléatoire. On peut rendre stationnaire ces processus par l'utilisation d'un filtre aux différences de la forme :

$$(1-L)^d Y_t = cste + \varepsilon_t$$

Avec L l'opérateur décalage, d l'ordre du filtre, ε_t un processus bruit blanc.

Dans la pratique d prend généralement la valeur 1. Dans ce cas deux processus particuliers voient le jour :

- **1) un processus DS sans dérive :**

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

(la constante est nulle). Dans la littérature on appelle un tel modèle une <u>marche aléatoire</u>. La nouvelle série stationnaire sera $\Delta Y_t = Y_t - Y_{t-1}$.

2) un processus DS avec dérive :

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

(la constante n'est pas nulle). La nouvelle série stationnaire sera $\Delta Y_t = Y_t - Y_{t-1} - cste$.

Pour résumer : dans un processus TS l'effet produit par un choc aléatoire à un instant t est transitoire alors que dans un processus DS l'effet est permanent (il se répercute à l'infini de manière décroissante).

Sur le plan pratique comment pouvons-nous reconnaître une série TS d'une série DS ? Tout simplement à l'aide des tests de Dickey-Fuller et Dickey-Fuller augmenté [dans ce contexte ces tests s'appellent les tests de stationnarité].

➲ Les tests de Dickey-Fuller (DF) :

Ce sont des tests paramétriques composés de la forme :

$$\begin{cases} H_0 : \phi_1 = 1 & \text{(pas stationnaire)} \\ H_1 : |\phi_1| < 1 & \end{cases}$$

Que l'on applique dans les trois modèles suivants :

[1] $X_t = \phi_1 X_{t-1} + \varepsilon_t$ (DS sans dérive)

[2] $X_t = \phi_1 X_{t-1} + cste + \varepsilon_t$ (DS avec dérive)

[3] $X_t = \phi_1 X_{t-1} + cste + b^*t + \varepsilon_t$ (TS)

Si H_0 est vérifiée alors la série X_t n'est pas stationnaire (i.e. il existe une racine unitaire) quel que soit le modèle. Pour des raisons statistiques le T-ratio calculé pour ϕ_1 n'est plus comparé au t_α lu dans une normale (0,1) (ou Student) mais dans des tables spéciales tabulées par Dickey-Fuller (voir infra). Si T-ratio > t_α alors on accepte H_0.

Dans le modèle [3] si b est significativement différent de zéro ($|$T-ratio$|$ > t_α; t_α lu dans une loi normale traditionnelle ou Student) alors on a affaire à un processus TS.

Dans le modèle [2] si cste est significativement différent de zéro, on a affaire à un modèle DS avec dérive.

➲ Les tests de Dickey-Fuller augmentés (ADF) :

Les tests ADF supposent que les résidus ne sont plus des bruits blancs mais soient autocorrélés. Ces tests se déroulent de manière similaire aux tests DF (seules les tables statistiques sont différentes) pour les modèles suivants :

[1’] $\Delta X_t = \rho X_{t-1} - \Sigma \phi_j \Delta X_{t-j+1} + \varepsilon_t$
[2’] $\Delta X_t = \rho X_{t-1} - \Sigma \phi_j \Delta X_{t-j+1} + cste + \varepsilon_t$
[3’] $\Delta X_t = \rho X_{t-1} - \Sigma \phi_j \Delta X_{t-j+1} + cste + b^*t + \varepsilon_t$

Les sommes varient généralement de 1 à 4.

Le test à réaliser se construit de la manière suivante:

$$\begin{cases} H_0 : \rho = 0 & \text{(pas stationnaire)} \\ H_1 : \rho \neq 0 \end{cases}$$

➲ Le test de Phillips et Perron (1988) :

Ce test est construit pour prendre en compte des erreurs hétéroscédastiques. C’est un test alternatif au test ADF. Pour des raisons purement statistiques (il n’y a pas d’ajout de régresseurs supplémentaires) ces tests sont plus puissants que les tests ADF.

3.2. La coïntégration traditionnelle

3.2.1. Les racines unitaires et l’ordre d’intégration des séries

Le test d'une racine unitaire (d'une seule série) peut être vu de deux façons différentes mais qui aboutissent au même résultat :

Cas 1 : on teste :

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

sur le modèle : $Y_t = \rho\, Y_{t-1}\ \{+ \text{cste} + \text{time}\} + \varepsilon_t$

Si $\rho = 1$ alors il existe une racine unitaire $\Rightarrow$ $Y_t \sim I(1)$ et ΔY_t est $I(0)$ donc stationnaire. On dit ici que Y_t est stationnaire en différence première. Si ρ est différent de 1 alors Y_t est $I(0)$ i.e. stationnaire en niveau.

Cas 2 : on teste :

$$\begin{cases} H_0 : \beta = 0 \\ H_1 : \beta \neq 0 \end{cases}$$

sur le modèle : $\Delta Y_t = \beta\, Y_{t-1}\ \{+ \text{cste} + \text{time}\} + \varepsilon_t$

Si $\beta = 0$ alors il existe une racine unitaire $\Rightarrow$ $Y_t \sim I(1)$.

Pour rechercher l'ordre de stationnarité on utilise les tests DF et ADF.

<u>Remarque</u> : pour tester la présence de plusieurs racines unitaires on fait le test n°1 (ou n°2) autant de fois que nécessaire afin d'obtenir une variable stationnaire. Par exemple si le test conclue que Y_t en niveau n'est pas stationnaire, on va refaire le test sur la différence première c'est à dire avec le modèle :

$$\Delta y_t = \rho\, \Delta y_{t-1} \{+ \text{cste} + \text{time}\} + \varepsilon_t$$

Si le test conclue à une non-stationnarité (de ΔY_t), on effectue le test sur la différence seconde c'est à dire avec le modèle :

$$\Delta\Delta y_t = \rho\, \Delta\Delta y_{t-1} \{+ \text{cste} + \text{time}\} + \varepsilon_t.$$

A ce stade, sur des données économiques, le test concluera à la stationnarité. On aura alors la séquence suivante :

- Y_t non-stationnaire,
- ΔY_t non-stationnaire,
- $\Delta^2 Y_t$ stationnaire (i.e. $y_t \sim I(2)$)

3.2.2. Les systèmes coïntégrés : aperçu de la méthode

3.2.2.1. La coïntégration entre deux variables

On présente la méthode d'estimation en deux étapes de systèmes coïntégrés d'ordre (1,1)[1] proposée par Engle et Granger en 1987. Il s'agit d'estimer dans une première étape les paramètres du vecteur coïntégrant puis, dans une deuxième étape, d'utiliser ceux-ci dans un modèle à correcteur d'erreurs. En d'autres termes, cette méthode consiste à estimer d'une part une équation statique de long terme et d'autre part une équation de court terme c'est à dire la spécification ECM. Sur le plan statistique on peut démontrer que les OLS fournissent une estimation convergente du paramètre coïntégrant. De même que l'estimateur des OLS de l'équation à ECM obtenu en utilisant la valeur estimée du paramètre coïntégrant est convergent et asymptotiquement équivalent à l'estimateur du maximum de vraisemblance qui utilise la vraie valeur du paramètre coïntégrant. En pratique on peut exposer la méthode comme suit :

↳ Etape n°1 : on estime l'équation de long terme :

$$Y_t = cste + a\ x_t + \varepsilon_t$$

On sauvegarde les résidus de cette équation (par exemple RES)

1 Seulement pour deux variables I(1). Pour d'autres ordre d'intégration il n'y a aucune méthode.

↳ Etape n°2 : on estime l'équation de court terme :

$$\Sigma\ a_i'\ \Delta Y_{t-i} = cste + \Sigma\ b_i'\ \Delta X_{t-i} + c\ RES_{t-1}$$

Cette dernière étape n'est pas automatique. Il faut en fait vérifier s'il existe réellement une relation de coïntégration entre les variables. Pour cela des tests de coïntégration sont appliqués sur l'équation de la première étape. Si les tests indiquent l'existence d'une relation de coïntégration alors, et seulement dans ce cas, on peut passer à la deuxième étape. Dans le cas contraire on arrête la procédure et on estime la relation par les ECM traditionnels.

Examinons d'un peu plus près ces tests de coïntégration. Utilisant les apports de Dickey et Fuller, Engle et Granger proposent plusieurs tests de coïntégration qui permettent de juger de la stationnarite d'une combinaison lineaire de variables qui sont individuellement non stationnaires. Engle et Granger ont defini sept tests qui portent tous sur l'étude des résidus de l'équation de long terme :

➊ test 1: test CRDW (coïntegrating regression DW)

Ce test est utilisé pour vérifier la stationnarité des résidus. On montre que sous l'hypothèse de non-coïntégration, la statistique DW tend asymptotiquement vers zéro, de ce fait les résidus ne sont pas stationnaires ; d'où le test CRDW qui rejette l'hypothèse nulle de non-coïntégration lorsque le DW est supèrieur à une valeur critique (on accepte la coïntégration dès lors que le DW est

assez grand). Les valeurs critiques ont été tabulées par Engle et Granger (1987). Il faut toutefois mentionner que ce test permet de tester la coïntégration uniquement dans le cas bivarié.

❷ tests 2/3: test DF (dickey-fuller) et test ADF(augmented dickey-fuller)

Ces deux tests permettent de vérifier la stationnarité des résidus en se basant directement sur les tests standard de racine unitaire de Dickey et Fuller. Ce sont les mêmes tests que la section précédente en remplaçant la variable X_t par la variable ε_t, le résidu. Engle et Yoo (1987) ont tabulé des valeurs critiques pour $n = 2, 3, 4, 5$ et $p = 0$ et 4. L'hypothèse nulle est la non-coïntégration que l'on rejette lorsque les statistiques DF et ADF sont supérieures aux valeurs critiques (i.e. on accepte la coïntégration pour des valeurs élevées de DF et ADF).

❸ tests 4/5: test RVAR (restricted VAR) et ARVAR (augmented RVAR)

Ces tests utilisent l'estimation en deux étapes des modèles à correcteur d'erreurs en testant la significativité des termes en niveau dans l'écriture d'un modèle à correction d'erreurs bivarié. Le test RVAR se fait sur le modèle sans terme retardé en ΔX_t et ΔY_t , alors que le test ARVAR introduit des termes en $\Delta X_t\text{-}1$ et $\Delta Y_t\text{-}1$.

❹ tests 6/7: test UVAR (unrestricted VAR) et test AUVAR (augmented UVAR)

Ces tests s'utilisent dans un modèle analogue au précédent, à savoir un modèle VAR en X et Y écrit dans une forme à correction d'erreurs mais dans lequel la forme des termes en niveau n'est pas contrainte par l'étape de la régression coïntégrante.

Notons que les tests 4/5 et 6/7 sont très peu employés.

A ce stade il faut préciser un point de détail. Les tests précédents ne sont pas la seule manière de vérifier si la relation de long terme est significative. Une autre statistique, que l'on oublie trop souvent, doit être aussi étudié : le R^2. Il faut aussi se fonder sur la valeur du R^2 pour juger si la relation étudiée est coïntégrante ou non. En théorie lorsque le R^2 est supérieur à .95 alors on peut penser qu'il existe une relation coïntégrante (toutefois en pratique on attend un R^2 proche de .99).

3.2.2.2. La coïntégration entre k variables

La généralisation de deux à k variables s'avère assez complexe du fait du nombre de possibilités de vecteurs de coïntégration : il peut y avoir k-1 vecteurs de coïntégration linéairement indépendants. Pour cette raison on se contente uniquement de rappeler les grandes lignes :

- **Si les variables sont de même ordre d'intégration (I(1) en pratique) alors l'existence d'un seul vecteur de coïntégration est possible. Nous pouvons employer la méthode en deux étapes de Engle et Granger définie dans la section précédente sans ambiguïté,**

- **Si les séries ne sont pas toutes intégrées du même ordre, alors nous pouvons être certains que le vecteur de coïntégration n'est pas unique. La méthode de Engle et Granger en deux étapes n'est plus valide. Nous devons faire appel à la représentation vectorielle à correction d'erreurs (VECM),**

- **Dans la pratique, pour tester une éventuelle coïntégration entre plusieurs variables, il convient tout d'abord de la tester sur l'ensemble des k variables, puis, en cas de coïntégration, de la tester par combinatoire entre les variables (en général deux à deux). Cf. test de Johansen.**

<u>Section 4 : Résumé, conclusion...</u>

Il y a plusieurs raisons pour s'intéresser à la présence d'une racine unitaire dans une série économique :

➲ Sur le plan <u>statistique</u>, la présence d'une racine unitaire impliquent que les propriétés asymptotiques générales des estimateurs (vitesse de convergence, normalité asymptotique,…) ne tiennent plus. Il faut donc avoir recours à une théorie asymptotique spéciale;

➲ Sur le plan <u>économique</u>, la présence d'une racine unitaire peut conduire à estimer des régressions apparemment très bonnes entre des variables qui sont totalement indépendantes entre elles (problèmes des régressions factices);

➲ Sur le plan <u>des simulations</u>, une série TS et une série DS se comportent de manière radicalement opposée dans le long terme :

- **Une série TS a tendance à se repositionner autour de son trend déterministe après un choc aléatoire,**

- **Une série DS ne revient pas autour de sa tendance après un choc.**

En macroéconomie la coïntégration prend tout son essort dans les domaine suivants :

- **(1) La théorie du cycle conjoncturel et la théorie de la croissance endogène (avec la variable production),**

- **(2) La théorie du revenu permanent (avec la variable consommation),**

- **(3) La théorie de l’hystérésis (avec la variable chômage).**

(1) La théorie du cycle conjoncturel implique une composante tendancielle déterministe dans l'évolution de la production. Les fluctuations cycliques n'affectent pas la tendance et la politique conjoncturelle parvient à stabiliser les fluctuations sans changer la tendance profonde. A l'inverse avec la théorie de la croissance endogène, la tendance est stochastique et les chocs (réels ou monétaires) ont un impact permanent sur celle-ci.

(2) La consommation peut être déterminée par le revenu courant (c'est la thèse keynésienne) ou par le revenu permanent (c'est la thèse de M. Friedman). Dans ce dernier cas, la consommation ne réagit pas à des changements transitoires du revenu, mais seulement par rapport à des changements du revenu anticipé de manière rationnelle. En conséquence, il

n'est pas possible de prévoir les changements dans la consommation et celle-ci suit une marche aléatoire.

(3) Depuis 1986 une nouvelle théorie du marché du travail selon laquelle le salaire est déterminé de manière à conserver un même niveau d'emploi pour les insiders de la firme, commence à prendre forme. Il en résulte un comportement particulier de la série d'emploi qui au temps t est égale à sa valeur au temps t-1 à un terme d'erreur près. L'emploi suit donc une marche aléatoire et si l'offre de travail est constante, le taux de chômage suit également une marche aléatoire. Il y a donc persistance du chômage.

Concrétement (Cf. TD) une procédure de coïntégration se décompose en 6 étapes principales :

Etape 1 : Recherche de saisonnalité en étudiant l'ACF.

Etape 2 : Recherche de la présence ou pas de la constance et/ou du trend dans l'équation suivante :

$$\text{Ln } Y = \text{cste} + a \text{ TREND} + \varepsilon$$

On utilisera :

le test DF3 si cste et trend significatifs,
le test DF2 si cste significatif, trend non significatif,
le test DF1 si cste et trend pas significatifs.

Etape 3 : recherche de l'ordre de stationnarité

Tests utilisés : ADF-PP-KPSS-WS-DFGLS...

avec retards automatiques (ou imposés)

Etape 4 : modélisation VAR

On doit rechercher le p max.

Etape 5 : recherche de relation de coïntégration 2*2 ou 3*3 ou 4*4...

Etape 6 : VECM + impulse fonction