

Économétrie des séries temporelles non-stationnaires

Chapitre 3: Modélisation avancée des séries non-stationnaires

Gilles de Truchis

Master 2 EIPMC

Les chapitres du cours

- | | |
|---|-------------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

Objectif du Chapitre : cointégration fractionnaire

- Le Ch. 2 offre un remède à la régression factice : la cointégration.

\Rightarrow La cointégration va pouvoir s'étendre aux cas des processus $I(\delta)$

- Soit $X_t = (x_{1t}, x_{2t})$ deux processus $I(\delta = \delta_1 = \delta_2 > 0.25)$
- S'il existe une **combinaison linéaire** $\beta'X_t \sim I(\gamma)$ telle que $\gamma < \delta$, alors x_{1t} et x_{2t} partagent une dynamique commune de long terme

$\Rightarrow x_{1t}$ et x_{2t} sont **fractionnellement cointégrés**

- Nous discuterons de techniques d'**estimation jointe** de β , δ et γ

Approche fréquentielle

- Les difficultés numériques liées à $\gamma_x(r - s; \vartheta)$ ont amené à d'autres pistes d'estimation
 - Une piste intéressante consiste à quitter le **domaine temporel**
 - l'idée est de considérer x_t comme un **signal**
 - un **signal** est un support informationnel dont l'évolution s'analyse
 - dans le **domaine temporel**
 - ou dans le **domaine des fréquences**
 - Dans le **domaine des fréquences**
 - on cherche à identifier quelle part du signal s'exprime sur une **fréquence donnée**
- ⇒ nous allons définir la notion de **fréquence**

Les fréquences

- Soit une fonction $\mathcal{F}(x)$ diffusant un signal x
- $n \in \mathbb{R}$ tel que $\mathcal{F}(x + n) = \mathcal{F}(x)$ définit la **période** de la fonction
 \Rightarrow le temps nécessaire pour que la fonction repasse en $\mathcal{F}(x)$
- On nomme **fréquence ordinaire** le nombre de fois qu'un signal périodique se reproduit par unité de temps

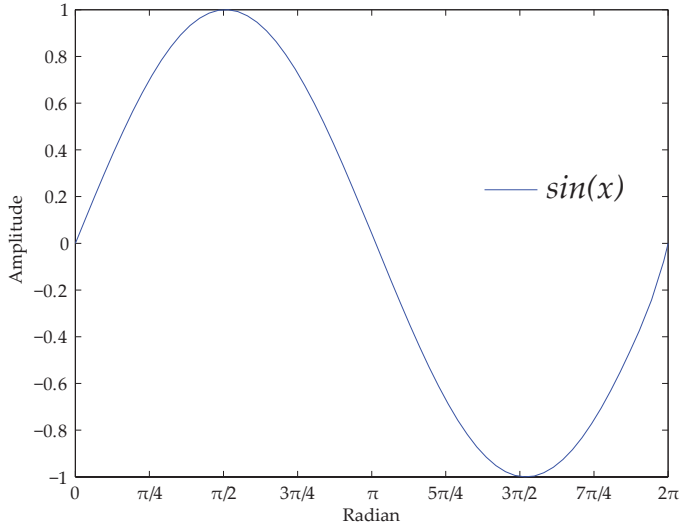
$$f = \frac{1}{n}$$

- En proportion d'un disque unité (dont la circonférence est 2π) on obtient la notion de **fréquence angulaire**

$$\lambda = 2\pi f = \frac{2\pi}{n}$$

Intuition graphique

- Par exemple, $\sin(x)$ est un signal 2π -périodique



Transformées de Fourier

- Soit une fonction $\mathcal{F}(t)$ intégrable opérant dans le domaine temporel
- Pour convertir $\mathcal{F}(t)$ en une fonction $\mathcal{F}(\lambda)$ opérant dans le domaine fréquentiel on utilise la **transformée de Fourier**

$$\mathcal{F}(\lambda) = \int_{-\infty}^{+\infty} \mathcal{F}(t) e^{-i\lambda t} dt$$

- Pour convertir $\mathcal{F}(\lambda)$ en une fonction $\mathcal{F}(t)$ opérant dans le domaine temporel on utilise la **transformée de Fourier inverse**

$$\mathcal{F}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{F}(\lambda) e^{i\lambda t} d\lambda$$

Note $e^{i\varphi} = \cos(\varphi) + i \sin(\varphi)$ en vertu de la **formule d'Euler**

- On notera **DFT** où **IDFT** les versions discrètes de ces équations

On voit qu'il n'y a pas une symétrie parfaite entre la transformée et la transformée inverse de Fourier (le facteur $1/(2\pi)$ apparaît). Afin d'obtenir une symétrie parfaite on peut parfois normaliser les deux formules par $1/(\sqrt{2\pi})$ ce qui donne

$$\mathcal{F}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathcal{F}(t) e^{-i\lambda t} dt$$

et

$$\mathcal{F}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathcal{F}(\lambda) e^{i\lambda t} d\lambda$$

Représentation spectrale

- Ces définitions tiennent pour des processus stochastiques
- Soit x_t un processus stationnaire d'autocovariance γ_x
- La représentation spectrale de x_t est alors donnée par

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} dM(\lambda)$$

avec $M(\lambda) \in [-\pi, \pi]$, une **mesure spectrale**

- $M(\lambda)$ spécifie entièrement x_t , à l'image d'une **loi de distribution**

Considérons un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ avec

- Ω l'ensemble **dénombrable** des éventualités
- \mathcal{F} un σ -algèbre représentant les événements
- \mathbb{P} une mesure de probabilité telle que $\mathbb{P}(A)$ donne la probabilité de l'événement $A \in \mathcal{A}$ et $\mathbb{P}(\Omega) = 1$

Une mesure spectral M définie sur \mathcal{M} , l'algèbre des fonctions complexes mesurables est, dans l'esprit de \mathbb{P} bien que de manière non-équivalente, une mesure dans l'espace mesuré (Ω, \mathcal{A}, M) où \mathcal{A} est un σ -algèbre dans l'espace des projections orthogonales d'un espace hilbertien.

Distribution spectrale

- $M(\lambda)$ est stochastique, défini sur le plan complexe \mathbb{C} et pour $\lambda \neq \nu$

$$\text{cov}(dM(\lambda), dM(\nu)) = \mathbb{E}(dM(\lambda)dM(\nu)^*) = 0$$

sachant que

$$\mathbb{E}(dM(\lambda)) = 0, \quad \mathbb{V}(dM(\lambda)) = \mathbb{E}(|dM(\lambda)|^2) = dF(\lambda)$$

avec $dF(\lambda)$ une fonction de **distribution spectrale**

- On parle de distribution car $dF(\lambda)$ assigne les masses de x_t sur les fréquences de l'intervall $[-\pi, \pi]$
- La masse assignée à la j -ième fréquence est la variance $\sigma_x^2(j)$ correspondant à cette fréquence

$$F(\lambda) = \sum_{j, \lambda_j \leq \lambda} \sigma_x^2(j)$$

Densité spectrale (théorème)

Theorem (1 : Conséquences d'Herglotz)

Réécrivons $dF(\lambda)$ comme l'équation infinitésimale $dF(\lambda) = f(\lambda)d\lambda$.
 $f(\lambda)$ est appelée **densité spectrale** et vérifie

$$\gamma_x(h) = \int_{-\pi}^{\pi} e^{ih\lambda} dF(\lambda) = \int_{-\pi}^{\pi} e^{ih\lambda} f_x(\lambda) d\lambda < \infty,$$

faisant apparaître que $f(\lambda)$ est l'équivalent de γ_x dans le domaine des fréquences. De plus la **DFT** normalisée par $(2\pi)^{-1}$ permet de passer de $\gamma_x(h)$ à $f_x(\lambda)$

$$f_x(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) e^{-ih\lambda}$$

- Utiliser $f_x(\lambda)$ va permettre d'éviter les problèmes numériques

⇒ il existe un algorithme très rapide pour la **transformée de Fourier**

Preuve du Théorème 1

- Supposons $f_x(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) e^{-ih\lambda}$
- Écrivons alors

$$\gamma_x(h) = \int_{-\pi}^{\pi} e^{ij\lambda} f_x(\lambda) d\lambda = \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) e^{i(j-h)\lambda} d\lambda$$

- Par application du théorème de Fubini on a

$$\int_{-\pi}^{\pi} e^{ij\lambda} f_x(\lambda) d\lambda = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) \int_{-\pi}^{\pi} e^{i(j-h)\lambda} d\lambda$$

- Le seul terme non-nul de la somme apparaît pour $j = h$ car

$$\int_{-\pi}^{\pi} e^{ik\lambda} d\lambda = 0 \text{ pour } k \neq 0 \text{ et } \int_{-\pi}^{\pi} e^{i(j-h)\lambda} d\lambda = \int_{-\pi}^{\pi} d\lambda = 2\pi \text{ si } j = h$$

⇒ On trouve bien

$$\int_{-\pi}^{\pi} e^{ih\lambda} f_x(\lambda) d\lambda = \frac{1}{2\pi} 2\pi \gamma_x(h) = \gamma_x(h)$$

$$\int_{-\pi}^{\pi} e^{ik\lambda} d\lambda = \left[ike^{ik\lambda} \right]_{-\pi}^{\pi} = ike^{-ik\pi} - ike^{ik\pi}$$

qui part l'identité d'Euler donne pour $k \neq 0$

$$\int_{-\pi}^{\pi} e^{ik\lambda} d\lambda = ik(\cos(-k\pi) + i\sin(-k\pi) - \cos(k\pi) - i\sin(k\pi))$$

avec $\sin(-k\pi) = \sin(k\pi) = 0$ pour $k \in \mathbb{Z}$ et $\cos(-k\pi) = \cos(k\pi) = \pm 1$ selon que k pair ou impair.

Densité spectrale d'un bruit blanc

- Soit $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2 < \infty)$ de représentation spectrale de ε_t

$$\varepsilon_t = \int_{-\pi}^{\pi} e^{i\lambda t} dM_\varepsilon(\lambda)$$

- Puisque $\mathbb{E}(dM_\varepsilon(\lambda)dM_\varepsilon(\nu)^*) = 0$ sauf si $\lambda = \nu$ et via le Th. 1

$$\begin{aligned} \mathbb{E}(\varepsilon_t \varepsilon_s) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(t\lambda - s\nu)} \mathbb{E}(dM_\varepsilon(\lambda)dM_\varepsilon(\nu)^*) \\ &= \int_{-\pi}^{\pi} e^{it\lambda} dF_\varepsilon(\lambda) = \int_{-\pi}^{\pi} e^{it\lambda} f_\varepsilon(\lambda) d\lambda \\ &= \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \gamma_\varepsilon(s) \int_{-\pi}^{\pi} e^{i(t-s)\lambda} d\lambda \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \int_{-\pi}^{\pi} e^{i(t-s)\lambda} d\lambda \end{aligned}$$

- On voit donc que $f_\varepsilon(\lambda) = \sigma_\varepsilon^2 (2\pi)^{-1}$

Densité spectrale d'un processus $I(0)$

- Soit x_t un processus linéaire $I(0)$ de représentation de Wold

$$x_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$$

- La représentation spectrale de x_t est alors

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} dM_x(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda t} \left(\sum_{j=0}^{\infty} a_j e^{-i\lambda j} \right) dM_{\varepsilon}(\lambda)$$

- On souligne donc ici que $dM_x(\lambda)$ se décompose comme

$$dM_x(\lambda) = \left(\sum_{j=0}^{\infty} a_j e^{-i\lambda j} \right) dM_{\varepsilon}(\lambda) = A(e^{-i\lambda}) dM_{\varepsilon}(\lambda)$$

- Par application du Th. 1 on trouve

$$f_x(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_x(h) e^{-i\lambda h} = \frac{\sigma_{\varepsilon}^2}{2\pi} \left| \sum_{j=0}^{\infty} a_j e^{-i\lambda j} \right|^2 = \frac{\sigma_{\varepsilon}^2}{2\pi} |A(e^{-i\lambda})|^2$$

Densité spectrale d'un AR(1)

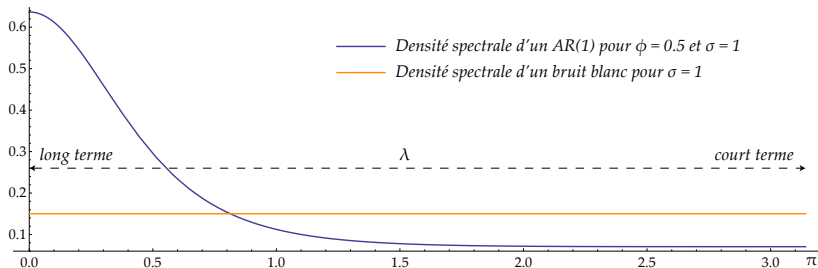
- Soit $x_t = \phi_1 x_{t-1} + \varepsilon_t$ avec $\varepsilon_t \sim i.i.d. WN(0, \sigma_\varepsilon^2 < \infty)$

$$\Rightarrow x_t = (1 - \phi_1 L)^{-1} \varepsilon_t = \sum_{j=1}^{\infty} a_j \varepsilon_{t-j}$$

- Comme $x_t \sim I(0)$, on peut écrire

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(e^{-i\lambda}) dM_\varepsilon(\lambda)$$

$$\text{et } f_x(\lambda) = \sigma_\varepsilon^2 (2\pi)^{-1} |A(e^{-i\lambda})|^2 = \sigma_\varepsilon^2 (2\pi)^{-1} |1 - \phi_1 e^{-i\lambda}|^{-2}$$



Densité spectrale d'un ARMA(p, q)

- Soit x_t un ARMA(p, q) dont on sait que

$$\begin{aligned} X_t &= \sum_{j=1}^p \phi_j X_{t-j} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. WN(0, \sigma_\varepsilon^2 < \infty) \\ &= \frac{1 - \theta_1 L + \dots + \theta_q L^q}{1 - \phi_1 L + \dots + \phi_p L^p} \varepsilon_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t = \psi(L) \varepsilon_t = \sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \end{aligned}$$

- Il s'agit d'un cas particulier de processus $I(0)$ et donc

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(e^{-i\lambda}) dM_\varepsilon(\lambda)$$

- Par application du Th. 1 on trouve

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |A(e^{-i\lambda})|^2 = \frac{\sigma_\varepsilon^2}{2\pi} \left| \frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})} \right|^2$$

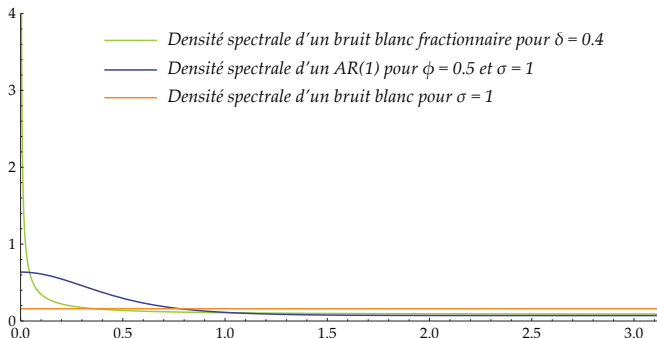
Densité spectrale d'un bruit blanc fractionnaire

■ Soit $(1-L)^\delta x_t = \varepsilon_t$ **bruit blanc fractionnaire** avec $\delta \in (-1/2, 1/2)$

■ x_t possède donc une représentation spectrale du type

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(e^{-i\lambda}) dM_\varepsilon(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda t} (1 - e^{-i\lambda})^{-\delta} dM_\varepsilon(\lambda)$$

\Rightarrow Le Th. 1 nous donne $f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |A(e^{-i\lambda})|^2 = \frac{\sigma_\varepsilon^2}{2\pi} |1 - e^{-i\lambda}|^{-2\delta}$



Densité spectrale d'un $ARFIMA(p, \delta, q)$ stationnaire

- Soit $x_t \sim ARFIMA(p, \delta, d)$ avec $\delta \in (-1/2, 1/2)$
- x_t possède donc une représentation spectrale du type

$$x_t = \int_{-\pi}^{\pi} e^{i\lambda t} A(e^{-i\lambda}) dM_{\varepsilon}(\lambda) = \int_{-\pi}^{\pi} e^{i\lambda t} (1 - e^{-i\lambda})^{-\delta} S(e^{-i\lambda}) dM_{\varepsilon}(\lambda)$$

avec $S(e^{-i\lambda})$ la composante $MA(\infty)$ absolument sommable

$$S(e^{-i\lambda}) = \left| \frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})} \right|$$

- Par application du Th. 1 on trouve

$$f_x(\lambda) = \frac{\sigma_{\varepsilon}^2}{2\pi} |A(e^{-i\lambda})|^2 = \frac{\sigma_{\varepsilon}^2}{2\pi} |1 - e^{-i\lambda}|^{-2\delta} \left| \frac{\theta(e^{-i\lambda})}{\phi(e^{-i\lambda})} \right|^2 = |1 - e^{-i\lambda}|^{-2\delta} f_S(\lambda)$$

Densité spectrale d'un $I(\delta)$ et approximation locale

- Soit $x_t \sim I(\delta)$ avec $\delta \in (-1/2, 1/2)$ de densité spectrale

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |A(e^{-i\lambda})|^2 = \frac{\sigma_\varepsilon^2}{2\pi} |1 - e^{-i\lambda}|^{-2\delta} |S(e^{-i\lambda})|^2$$

- Concentrons-nous sur les **basses fréquences** (long terme)

- Notons que $|1 - e^{i\lambda}| = |2 \sin(\lambda/2)|$ et donc

$$\lim_{\lambda \rightarrow 0} |1 - e^{i\lambda}|^{-2\delta} = \lim_{\lambda \rightarrow 0} |2 \sin(\lambda/2)|^{-2\delta} \approx \lim_{\lambda \rightarrow 0} |\lambda|^{-2\delta} = \infty$$

- f_x diverge en 0 car ici x_t possède une **dynamique de long terme**

Approximation locale et fonctions à variation lente

- Beran et al. (2013) rappellent que f_x peut s'écrire comme

$$f_x \sim L_f(\lambda)|\lambda|^{-2\delta}, \quad \lambda \rightarrow 0$$

- $L(\omega)$ est une fonction à **variation lente** (VL_∞) en l'infini si

$$\lim_{\omega \rightarrow \infty} \frac{L(c\omega)}{L(\omega)} = 1 \text{ noté } L(\omega) \sim L(c\omega) \quad \forall c > 0$$

et à variation lente en 0 (VL_0) si $L(1/\omega)$ est VL_∞

- $f(\omega) \sim \omega^\theta L(\omega)$ est à **variation régulière** en 0 (VR_0) si $L(1/\omega)$ est VL_∞

Theorem (2 : Équivalences de Beran)

Soit $x_t \sim I(\delta)$ avec $\delta \in (-1/2, 0) \cup (0, 1/2)$. Alors pour $\lambda \rightarrow 0$

$$\gamma_x(h) = L_\gamma(h)|h|^{2\delta-1} \text{ et } f_x \sim L_f(\lambda)|\lambda|^{-2\delta}$$

avec $L_f(\lambda) = L_\gamma(\lambda^{-1})\pi^{-1}\Gamma(2\delta)\sin(\pi/2 - \pi\delta)$

Approximation locale d'un bruit blanc fractionnaire

- Soit x_t un bruit blanc $I(\delta)$ dont on sait depuis le Chapitre 1 que

$$\gamma(h) \sim \underbrace{\sigma_\varepsilon^2 \pi^{-1} \Gamma(1 - 2\delta) \sin(\pi\delta)}_{L_\gamma} h^{2\delta-1} + O(h^{2\delta-3}), \quad h \rightarrow \infty$$

pour $\delta \in (-1/2, 0) \cup (0, 1/2)$

- Dans le Théorème 2, on a alors $f_x \sim \sigma_\varepsilon^2 (2\pi)^{-1} |\lambda|^{-2\delta}$ car

$$\begin{aligned} L_f(\lambda) &= L_\gamma(\lambda^{-1}) \pi^{-1} \Gamma(2\delta) \sin(\pi/2 - \pi\delta) \\ &= \sigma_\varepsilon^2 \pi^{-1} \Gamma(1 - 2\delta) \sin(\pi\delta) \pi^{-1} \Gamma(2\delta) \sin(\pi/2 - \pi\delta) \\ &= \frac{\sigma_\varepsilon^2}{\pi} \frac{\sin(\pi\delta) \sin(\pi/2 - \pi\delta)}{\sin(2\pi\delta)} = \frac{\sigma_\varepsilon^2}{\pi} \frac{\sin(\pi\delta) \cos(\pi\delta)}{\sin(2\pi\delta)} \\ &= \frac{\sigma_\varepsilon^2}{\pi} \frac{\sin(\pi\delta) \cos(\pi\delta)}{2 \sin(\pi\delta) \cos(\pi\delta)} = \frac{\sigma_\varepsilon^2}{2\pi} \end{aligned}$$

Note : d'après la formule de réflexion d'Euler : $\Gamma(2\delta)\Gamma(1-2\delta) = \pi / \sin(2\pi\delta)$

Approximation locale d'un processus $I(\delta)$

- L'approximation de Lieberman et Phillips (2008)

$$\gamma(h) \sim \underbrace{\sigma_\varepsilon^2 \pi^{-1} \Gamma(1 - 2\delta) \sin(\pi\delta)}_{L_\gamma} h^{2\delta-1} + O(h^{2\delta-3}), \quad h \rightarrow \infty$$

tient pour une large classe de processus $I(\delta)$ stationnaires et inversibles

- Le Théorème 2 de Beran et al. (2013) tient aussi pour une large classe de processus $I(\delta)$ stationnaires et inversibles
- Si x_t suit un ARFIMA(p, δ, q) on retrouve donc l'approximation

$$f_x \sim \frac{\sigma_{\varepsilon'}^2}{2\pi} |\lambda|^{-2\delta} = g |\lambda|^{-2\delta}, \quad \lambda \rightarrow 0$$

pour $\delta \in (-1/2, 0) \cup (0, 1/2)$, l'erreur d'approximation étant $O(h^{2\delta-3})$

... et $\sigma_{\varepsilon'}^2 = 2\pi f_S(0)$, la variance d'un ARFIMA(0, δ , 0)

L'erreur d'approximation étant $O(h^{2\delta-3})$ elle s'élimine plus vite que la constante de normalisation L_γ et $h^{2\delta-1}$

Comme en $\lambda = 0$ la composante de court terme de X a une densité spectrale supposée bornée, on peut opérer un changement de variable et définir un ARFIMA(0, δ , 0) de variance $\sigma_{\varepsilon'}^2 = 2\pi f_S(0)$ ayant pour densité spectral $f_x \sim g|\lambda|^{-2\delta}$

Estimation non-paramétrique de $f_x(\lambda)$

- Pour x_t un processus stationnaire, $\hat{\gamma}_x(h)$ est donnée par

$$\hat{\gamma}_x(h) = n^{-1} \sum_{t=1}^{n-k} x_t x_{t+k}, \text{ pour } k > 0 \text{ et } \mathbb{E}(x_t) = 0$$

- D'après le Théorème 1,

$$\hat{f}_x(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \hat{\gamma}_x(h) e^{-ih\lambda} = \frac{1}{2\pi} \sum_{h=-(n-1)}^{n-1} \hat{\gamma}_x(h) e^{-ih\lambda}$$

- D'après la formule de $\hat{\gamma}_x(h)$ on obtient un **périodogramme**

$$\hat{f}_x(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n x_t e^{-it\lambda} \right|^2 = \hat{I}_x(\lambda) \quad (1)$$

- En termes de **DFT** (symétrisée), le **périodogramme** s'écrit

$$\hat{I}_x(\lambda) = |\omega_x(\lambda)|^2 \text{ avec } \omega_x(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n x_t e^{-it\lambda_j} \text{ et } \lambda_j = \frac{2\pi j}{n}$$

Théorie limite du périodogramme dans le cas $I(0)$

- Si $x_t \sim I(0)$ est un processus linéaire on peut montrer que

- $\hat{I}_x(\lambda)$ est un estimateur **non-biaisé** de $f_x(\lambda)$

$$\mathbb{E}(\hat{I}_x(\lambda) - f_x(\lambda)) = O(n^{-1})$$

- Les masses de $\hat{I}_x(\lambda)$ sont **asymptotiquement non-corrélées**.

$$\left| \text{cov}(\hat{I}_x(\lambda_j), \hat{I}_x(\lambda_l)) \right| = cn^{-1}, \forall c > 0 \text{ et } j \neq l$$

- Raisonnons avec le périodogramme standardisé, $\hat{I}_x(\lambda)/f_x(\lambda)$

\Rightarrow La convergence en loi est assurée pour chaque fréquence

$$\left(\frac{\hat{I}_x(\lambda_1)}{f_x(\lambda_1)}, \dots, \frac{\hat{I}_x(\lambda_j)}{f_x(\lambda_j)} \right) \xrightarrow{d} \frac{1}{2}(Z_1^2, \dots, Z_j^2)$$

avec Z_1, \dots, Z_j des variables aléatoires Gaussiennes standardisées de sorte que $Z_j^2/2 \sim$ distribution exponentielle d'espérance 1

Théorie limite du périodogramme dans le cas $I(\delta)$

■ Si $x_t \sim I(\delta)$ est un processus linéaire on peut montrer que

■ $\hat{I}_x(\lambda)$ est un estimateur **positivement biaisé** de $f_x(\lambda)$ car

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{\hat{I}_x(\lambda_j)}{f_x(\lambda_j)} \right) = \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2} \left| \frac{\lambda}{2\pi j} \right|^{-2\delta} d\lambda \geq 1$$

\Rightarrow on voit une dépendance à la fréquence j ce qui préfigure une **dépendance asymptotique** dont la forme est

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{\omega_x(\lambda_j) \omega_x(\lambda_l)^*}{\sqrt{f_x(\lambda_j) f_x(\lambda_l)}} \right) = \gamma_\omega(j, l, \delta)$$

avec

$$\gamma_\omega(j, l, \delta) = (-1)^{j+l+1} |2\pi j|^\delta |2\pi l|^\delta \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)(2\pi l - \lambda)} |\lambda|^{-2\delta} d\lambda$$

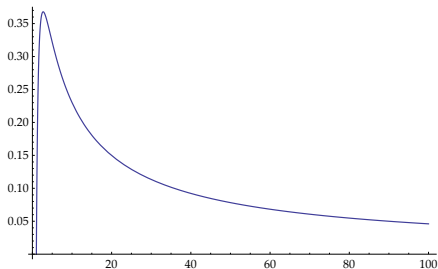
Pour $\delta = 0$, le processus est à mémoire courte et on a

$$\frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2} \left| \frac{\lambda}{2\pi j} \right|^{-2\delta} d\lambda = 1$$

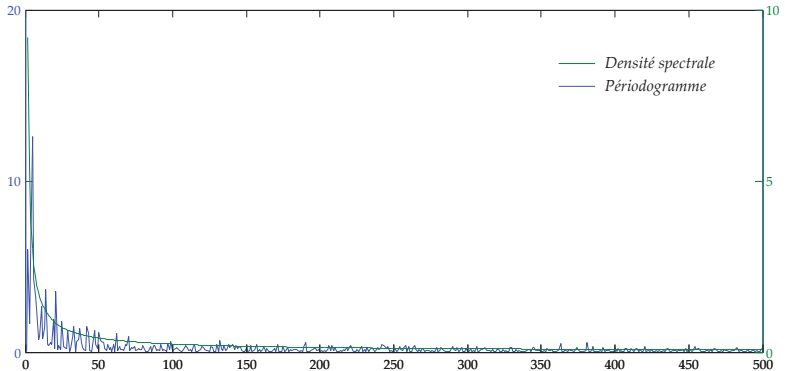
nous ramenant ainsi bien au résultat pour $x_t \sim I(0)$

Hurvich et al. (1998) montrent que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{\hat{I}_x(\lambda_j)}{f_x(\lambda_j)} \right) = O(\log j/j)$$



Périodogramme et densité spectrale d'un $I(\delta = 0.4)$



Plan

- | | |
|--|------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

Estimation de $\delta < 1/2$: approche temporelle

- Si $x_t \sim I(\delta < 1/2)$, le processus est stationnaire au second ordre

\Rightarrow la fonction d'autocovariance suffit pour décrire la dynamique de x_t

- Supposons que $x_t \sim ARFIMA(p, \delta, q)$

- Sous une hypothèse Gaussienne, la log-vraisemblance est

$$\begin{aligned}\mathcal{L}_E(\vartheta; \mathbf{x}) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma(\vartheta)| - \frac{1}{2} \mathbf{x}' \Sigma(\vartheta)^{-1} \mathbf{x} \\ &\approx \log(2\pi) + \frac{1}{n} \log |\Sigma(\vartheta)| + \frac{1}{n} \mathbf{x}' \Sigma(\vartheta)^{-1} \mathbf{x}\end{aligned}\quad (2)$$

avec $\mathbf{x} = (x_1, \dots, x_n)'$, $\vartheta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_p)$ et

$$\Sigma(\vartheta) = \gamma_x(r-s; \vartheta), \text{ pour } r, s = 1, \dots, n$$

et sachant qu'au Ch. 1 nous avons obtenu pour un ARFIMA(0, δ , 0)

$$\gamma(h) = \frac{(\sigma_\varepsilon^2/2\pi) \sin(\pi\delta) \Gamma(h+\delta) \Gamma(1-2\delta)}{\Gamma(h+1-\delta)}$$

Beran et al. (2013) p. 417 : on multiplie par $-2/n$ et du coup on minimise la fonction objectif

Estimation de $\delta < 1/2$: MLE exact

- L'estimateur du **maximum de vraisemblance exact** est alors

$$\hat{\vartheta} = \arg \min_{\vartheta \in \Theta} \mathcal{L}_E(\vartheta; \mathbf{x})$$

- Dahlhaus (1989) montre que $\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(0, \Xi)$
 - L'expression complexe de Ξ est détaillée dans Beran et al. (2013)
 - La théorie limite repose sur des conditions de régularité particulières
- Hosoya (1997) étend le MLE Gaussien au cas bivarié
- Sowell (1992) étudie les **propriétés numériques** de $\hat{\theta}$
 - Requiert l'inversion de $\Sigma(\vartheta)$, une matrice $n \times n$ à chaque itération
 - ⇒ pour n grand, cela devient très lourd informatiquement
 - ⇒ $\Sigma(\vartheta)$ approche la singularité pour $\delta \rightarrow 0.5$ dans l'optimisation

Approximation du MLE exact

- Soit $x_t \sim ARFIMA(p, \delta, d)$ avec $\delta \in (-1/2, 1/2)$ et

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |A(e^{-i\lambda})|^2 = \frac{\sigma_\varepsilon^2}{2\pi} h_x(\lambda)$$

- D'après le Théorème 1 on sait que

$$\gamma_x(h) = \frac{\sigma_\varepsilon^2}{2\pi} \int_{-\pi}^{\pi} e^{ih\lambda} h_x(\lambda) d\lambda$$

et donc que $\Sigma(\vartheta) = \sigma_\varepsilon^2 / (2\pi) \Sigma_h(\vartheta)$ avec

$$\Sigma_h(\vartheta) = \left(\int_{-\pi}^{\pi} e^{i(r-s)\lambda} h_x(\lambda; \vartheta) d\lambda \right)_{r,s=1,\dots,n} \quad (3)$$

- En utilisant cette représentation dans la vraisemblance (2) on a

$$\mathcal{L}_A(\vartheta; \mathbf{x}) = \log \sigma_\varepsilon^2 + \frac{1}{n} \log |\Sigma_h(\vartheta)| + \left(\frac{2\pi}{\sigma_\varepsilon^2} \right) \frac{1}{n} \mathbf{x}' \Sigma_h(\vartheta)^{-1} \mathbf{x}$$

Le terme $\log(2\pi)$ se simplifie avec le terme $-\log(2\pi)$ qui vient de $\Sigma(\vartheta) = \sigma_\varepsilon^2 / (2\pi) \Sigma_h(\vartheta)$

Théorème de Szegö

Theorem (3 : Szegö)

Pour $\Sigma_h(\vartheta)$ défini par l'Équation (3)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log |\Sigma_h(\vartheta)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log h_x(\lambda; \vartheta) d\lambda = 0$$

et il en découle que

$$\sigma_\varepsilon^2 = 2\pi \exp \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda \right)$$

- Le Théorème 3 simplifie $\mathcal{L}_A(\vartheta; \mathbf{x})$ qui devient

$$\mathcal{L}_A(\vartheta; \mathbf{x}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda + \left(\frac{2\pi}{\sigma_\varepsilon^2} \right) \frac{1}{n} \mathbf{x}' \Sigma_h(\vartheta)^{-1} \mathbf{x} \quad (4)$$

si on ignore un terme constant $\log 2\pi$

Esquisse de preuve pour le théorème de Szegö

- Puisque pour $n \rightarrow \infty$ on a

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log h_x(\lambda; \vartheta) d\lambda = 0$$

et on peut écrire

$$\begin{aligned} \sigma_\varepsilon^2 &= \sigma_\varepsilon^2 \exp \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log h_x(\lambda; \vartheta) d\lambda \right) \equiv \sigma_\varepsilon^2 \exp(0) \\ &= 2\pi \exp \left(\log \frac{\sigma_\varepsilon^2}{2\pi} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log h_x(\lambda; \vartheta) d\lambda \right) \\ &= 2\pi \exp \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda \right) \text{ car } f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} h_x(\lambda) \end{aligned}$$

- On en déduit aussi

$$\log \sigma_\varepsilon^2 = \log 2\pi + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda \quad (5)$$

où $\log 2\pi$ est souvent ignoré car constant

Approximation de Whittle

- Whittle (1953) propose d'approximer $\Sigma_h(\vartheta)^{-1}$ par

$$W(\vartheta) = \left(\frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} e^{i(r-s)\lambda} \frac{1}{h_x(\lambda; \vartheta)} d\lambda \right) = \left(\omega_{r-s}(\vartheta) \right), \quad r, s = 1, \dots, n$$

- Utilisons cette approximation dans l'Équation (4) de $\mathcal{L}_A(\vartheta; \mathbf{x})$

$$\begin{aligned} \mathcal{L}_W(\vartheta; \mathbf{x}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda + \left(\frac{2\pi}{\sigma_\varepsilon^2} \right) \frac{1}{n} \mathbf{x}' W(\vartheta) \mathbf{x} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda + \left(\frac{\sigma_\varepsilon^2}{2\pi} \right)^{-1} \frac{1}{n} \sum_{t,s=1}^n x_t x_s \omega_{t-s}(\vartheta) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{n} \sum_{t,s=1}^n \frac{e^{i(t-s)\lambda} x_t x_s}{f_x(\lambda; \vartheta)} d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_x(\lambda; \vartheta) d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_x(\lambda)}{f_x(\lambda; \vartheta)} d\lambda \end{aligned}$$

où $I_x(\lambda)$ est le périodogramme décrit par l'Équation (1)

Estimateur de type Whittle

- L'estimateur **Whittle** est alors donnée par $\arg \min_{\vartheta \in \Theta} \mathcal{L}_W(\vartheta; \mathbf{x})$

$$\mathcal{L}_W(\vartheta; \mathbf{x}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\log f_x(\lambda; \vartheta) + \frac{I_x(\lambda)}{f_x(\lambda; \vartheta)} \right) d\lambda \quad (6)$$

- En pratique on remplacera l'intégrale par une \sum de Riemann

$$\mathcal{L}_W(\vartheta; \mathbf{x}) \approx \frac{2}{n} \sum_{j=1}^n \left(\log f_x(\lambda_j; \vartheta) + \frac{I_x(\lambda_j)}{f_x(\lambda_j; \vartheta)} \right), \quad \lambda_j = 2\pi j/n$$

- L'approximation concerne au départ les processus $I(0)$ et Fox et Taquu (1986) étendent l'approche aux processus $I(\delta < 1/2)$
- Comme pour $\mathcal{L}_E(\vartheta; \mathbf{x})$, Dahlhaus (1989) montre que,

$$\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(0, \Xi)$$

avec Ξ identique au cas du MLE exact

Estimateur de type Whittle local

- Supposons que seules les **basses fréquences** nous intéressent
- On peut alors évaluer $\mathcal{L}_W(\vartheta; \mathbf{x})$ pour m fréquences autours de 0

$$\mathcal{L}_W(\vartheta; \mathbf{x}) \approx \frac{2}{m} \sum_{j=1}^m \left(\log f_x(\lambda_j; \vartheta) + \frac{I_x(\lambda_j)}{f_x(\lambda_j; \vartheta)} \right), \quad \lambda_j = 2\pi j/n$$

- Si $m/n \rightarrow 0$ on sait que $\lambda \rightarrow 0$ et donc $f_x(\lambda_j; \delta) \sim g\lambda_j^{-2\delta}$ et

$$\mathcal{Q}_W(\vartheta; \mathbf{x}) \approx \frac{1}{m} \sum_{j=1}^m \left(\log (g\lambda_j^{-2\delta}) + \frac{I_x(\lambda_j)}{g\lambda_j^{-2\delta}} \right), \quad \lambda_j = 2\pi j/n$$

- L'estimateur **local Whittle** est alors donné par $\arg \min_{\vartheta \in \Theta} \mathcal{Q}_W(\delta; \mathbf{x})$

\Rightarrow L'estimateur est **semi-paramétrique** car il ne modélise que δ

\Rightarrow **Robuste** à une mauvaise spécification du court terme (ARMA)

Estimateur de type Whittle local concentré

- Un raffinement de $\mathcal{Q}_W(\vartheta; \mathbf{x})$ permet une ultime simplification
- En dérivant $\mathcal{Q}_W(\vartheta; \mathbf{x})$ par rapport à g on obtient

$$\frac{\partial \mathcal{Q}_W(\vartheta; \mathbf{x})}{\partial g} = \frac{1}{m} \left(\sum_{j=1}^m g^{-1} - \frac{I_x(\lambda_j)}{g^2 \lambda_j^{-2\delta}} \right)$$

- Évaluée en 0 on obtient $\hat{g} = g_m(\delta) = m^{-1} \sum_{j=1}^m I_x(\lambda_j) \lambda_j^{2\delta}$
- En remplaçant $g_m(\delta)$ dans $\mathcal{Q}_W(\vartheta; \mathbf{x})$ on obtient

$$\mathcal{K}_W(\delta; \mathbf{x}) = \frac{1}{m} \sum_{j=1}^m \log(\hat{g} \lambda_j^{-2\delta}) + \frac{\hat{g}}{\hat{g}} = \log g_m(\delta) - \delta \frac{2}{m} \sum_{j=1}^m \log \lambda_j$$

\Rightarrow L'estimateur dévient alors $\hat{\delta} = \arg \min_{\delta \in \Theta} \mathcal{K}_W(\delta; \mathbf{x})$

- Robinson (1995) montre alors que $\sqrt{m}(\hat{\delta} - \delta) \xrightarrow{d} \mathcal{N}(0, 1/4)$

Notons que dans le calcul de $\mathcal{K}_W(\vartheta; \mathbf{x})$ on néglige le terme constant $\frac{\hat{g}}{g} = 1$ et

$$\frac{1}{m} \sum_{j=1}^m \log(\hat{g} \lambda_j^{-2\delta}) = \frac{m}{m} \log \hat{g} + \frac{1}{m} \sum_{j=1}^m \log \lambda_j^{-2\delta} = \log g_m(\delta) - \delta \frac{2}{m} \sum_{j=1}^m \log \lambda_j$$

Conditions d'application du TCL pour l'estimateur local Whittle

- $\delta \in (-1/2, 1/2)$ et $\mathbb{E}(\varepsilon_t^4) < \infty$
- $f_x(\lambda_j) = \lambda_j^{-2\delta} f_*(\lambda) = \lambda_j^{-2\delta} (f_*(0) + O(\lambda^\rho))$ pour $\rho \in (0, 2]$
- $m \rightarrow \infty$ et

$$\frac{(\log m)^2 m^{1+2\rho}}{n^{2\rho}} \rightarrow 0$$

$\hat{I}_x(\lambda_j)$ étant biaisé, la théorie limite de $\hat{\delta}$ n'est pas simple... Robinson (1995) montre qu'une borne sur $\mathbb{E}\hat{I}_x(\lambda_j)$ est suffisante pour assurer la consistance

Estimateur de Geweke et Porter-Hudak (1983)

- Geweke et Porter-Hudak (1983) adoptent une autre approche pour estimer δ **semi-paramétriquement**
- Partons du fait que pour $\lambda \rightarrow 0$ on a $f_x(\lambda_j; \vartheta) \sim g\lambda_j^{-2\delta}$
- On suppose que $\log I_x(\lambda_j) \approx \log f_x(\lambda_j; \vartheta)$ autour de 0 et donc

$$\log I_x(\lambda_j) = \log g + 2\delta \log \lambda_j + \varepsilon(\lambda_j), \quad j = 1, \dots, m, \quad (7)$$

$\varepsilon(\lambda_j) = \log(I_x(\lambda_j)/f_x(\lambda_j; \vartheta))$ un terme d'erreur et $m = o(n^{4/5})$

- On peut alors estimer l'Équation (7) par Least Squares (LSE)
- La théorie limite des LS nous donne $\sqrt{m}(\hat{\delta} - \delta) \xrightarrow{d} \mathcal{N}(0, \frac{\pi^2}{24})$
- $\hat{I}_x(\lambda_j)$ étant biaisé, la théorie limite de $\hat{\delta}$ est non-triviale comme le montrent Hurvich et al. (1998)

Conditions d'application du TCL pour GPH

- $\delta \in (-1/2, 1/2)$ et x_t est Gaussien
- $f_x(\lambda_j) = \lambda_j^{-2\delta} f_*(\lambda) = \lambda_j^{-2\delta} (f_*(0) + O(\lambda^\rho))$ pour $\rho \in (0, 2]$
- $f'_x(\lambda_j) = O(\lambda^{-2\delta-1})$
- $m \rightarrow \infty$ et $m = o(n^{2\rho/(2\rho+1)})$

Hurvich et al. (1998) montrent que pour m bien sélectionné, plusieurs sources de biais s'annihilent et

$$\mathbb{E}(\hat{\delta} - \delta)^2 = \frac{4\pi^4}{81} \frac{f^{*''}(0)}{f^*(0)} \frac{m^4}{n^4} + \frac{\pi^2}{24m} + O\left(\frac{m(\log(m)^3)}{n^2}\right) + o\left(\frac{m^4}{n^4}\right) + o\left(\frac{1}{m}\right) \rightarrow 0$$

Domaine des fréquences et non-stationnarité

- Les estimateurs de δ de type Whittle et LS éliminent les problèmes numériques
- Les estimateurs **semi-paramétriques** de δ éliminent le risque de mauvaise spécification des fréquences $\lambda \gg 0$
- Un autre avantage de ces estimateurs réside dans une certaine accommodation de la non-stationnarité
 - Si $x_t \sim I(\delta > 1/2)$, une première approche serait d'estimer $\check{\delta} \equiv \delta - 1$ pour Δx_t puis de considérer que $\hat{\delta} = \check{\delta} + 1 \Rightarrow$ peu efficace
 - Mais nous allons voir que pour $\delta \in [1/2, 1)$, ces estimateurs préservent en partie de bonnes propriétés asymptotiques
- Nous ne discuterons pas des processus non-inversibles car le cas $\delta < -1/2 < 0$ représente une situation d'intérêt moindre en économie

Pseudo-densité spectrale

- Soit $x_t \sim I(\delta)$ un processus stationnaire et $f_x(\lambda) \sim g|\lambda|^{-2\delta}$
- Si $\delta > 1/2$ rappelons que $f(\lambda)$ n'existe pas $\gamma_x(h) = \infty$
- Cependant, si $z_t^{(s)} = \Delta^s x_t$ pour $s = \lfloor \delta + 1/2 \rfloor$ est stationnaire et

$$f_{z^{(s)}}(\lambda) \sim \lambda^{-2(\delta-s)}$$

- Hurvich et Ray (1995) montrent alors que x_t admet une **pseudo-densité spectrale**

$$f_x(\lambda) := |1 - e^{i\lambda}|^{-2s} f_{z^{(s)}}(\lambda) \sim g|\lambda|^{-2\delta}$$

au sens où $f_x(\lambda)$ apparaît comme limite de $\mathbb{E}I_x(\lambda_j)$ pour $\delta \in [1/2, 3/2)$

$\Rightarrow f_x(\lambda)$ ne représente pas la décomposition de la variance (infinie) de x_t mais peut-être utilisé pour construire des estimateurs de δ

Estimateur de δ et non-stationnarité

■ Grâce au concept de **pseudo-densité spectrale**

■ Velasco (1999) montre que l'estimateur **LSE** est

- consistant pour $\delta \in [1/2, 1)$
- asymptotiquement normalement distribué pour $\delta \in [1/2, 3/4)$

■ Velasco et Robinson (2000) montrent que le **Whittle** est

- consistant pour $\delta \in [1/2, 1)$
- asymptotiquement normalement distribué pour $\delta \in [1/2, 3/4)$

■ Phillips et Shimotsu (2004) montrent que le **local Whittle** est

- consistant pour $\delta \in [1/2, 1]$
- asymptotiquement normalement distribué pour $\delta \in [1/2, 3/4)$

■ Shao (2010) montre que pour le **Whittle**

- $\hat{\delta} \xrightarrow{P} 1$ pour $\delta \in (1, 3/2)$ et conjecture la consistance pour $\delta = 1$

Représentation spectrale multivariée et VARFIMA

- Soit un système k -varié linéaire de la forme

$$\begin{pmatrix} \Delta^{\delta_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta^{\delta_k} \end{pmatrix} \begin{pmatrix} x_{1t} \\ \vdots \\ x_{kt} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{kt} \end{pmatrix}$$

où $x_t = (x_{1t}, \dots, x_{kt})'$ est un ARFIMA si u_t est un ARMA

- Si $\max(\delta_1, \dots, \delta_k)' < 1/2$, x_t est stationnaire et

$$\mathbb{E}\left((x_t - \mathbb{E}(x_t))(x'_{t+j} - \mathbb{E}(x'_t))\right) = \int_{-\pi}^{\pi} e^{i\lambda j} f_x(\lambda) d\lambda$$

avec $f_x(\lambda)$ une matrice de densité spectrale décrite en $\lambda \rightarrow 0$ par

$$f_x(\lambda) \sim \Lambda(\delta_a) G \Lambda(\delta_a)^*, \quad \Lambda(\delta_a) = \text{diag}(\lambda^{-\delta_a} e^{i(\pi-\lambda)\delta_a/2}) \quad (8)$$

où $a = 1, \dots, k$ et G est la matrice de VCV de $u_t = (u_{1t}, \dots, u_{kt})'$

Estimation multivariée et VARFIMA

- La densité (8) permet une approche **local Whittle multivariée**

$$\mathcal{Q}_W(\vartheta; \mathbf{x}) \approx \frac{1}{m} \sum_{j=1}^m \left(\log \det \Lambda(\delta_a) G \Lambda^*(\delta_a) + \text{tr} (\Lambda(\delta_a) G \Lambda(\delta_a)^*)^{-1} \hat{I}(\lambda) \right)$$

ou encore **local Whittle multivariée concentrée**

$$\mathcal{K}_W(\delta; \mathbf{x}) = \log \det \hat{G}_m(\delta) - 2 \sum_{a=1}^k \delta_a \frac{1}{m} \sum_{j=1}^m \log \lambda_j$$

où $\hat{G}_m(\delta) = \frac{1}{m} \sum_{j=1}^m \Lambda(\delta_a)^{-1} \hat{I}(\lambda) \Lambda^*(\delta_a)^{-1}$ et

$$\hat{I}_{k \times k}(\lambda) = \omega_x(\lambda) \omega_x^*(\lambda) \text{ avec } \omega_x(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n x_t e^{-it\lambda_j} \text{ et } \lambda_j = \frac{2\pi j}{n}$$

- Shimotsu (2007) montre que

$$m^{1/2}(\hat{\delta} - \delta) \xrightarrow{d} \mathcal{N}(0, \Xi^{-1}) \text{ où } \Xi = 2 \left(G \odot G^{-1} + I_k + \frac{\pi^2}{4} (G \odot G^{-1}) - I_k \right)$$

avec \odot le produit d'Hadamard et $\hat{G}_m(\delta) \xrightarrow{p} G$

Pour le cas Whittle paramétrique, c'est plus compliqué même si la théorie limite existe (voir Hosoya (1996) ...)

Plan

- | | |
|--|------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

Définition de la cointégration fractionnaire

- Repartons de la définition d'Engle et Granger (1987)

Definition (Cointégration selon Engle et Granger, 1987)

Les éléments du vecteur de variables aléatoires $X_t = (x_{1,t}, \dots, x_{k,t})$ sont dits cointégrés d'ordres (δ, b) , noté $CI \sim (\delta, b)$ si

- a) *tous les éléments du vecteur X_t sont $I(\delta)$*
- b) *il existe un vecteur $\beta \neq 0$ tel que $\varepsilon_t = \beta' X_t \sim I(\delta - b)$ avec $b > 0$*

Le vecteur β est de dimension k et se nomme vecteur cointégrant.

- Pour $\delta, b \in \mathbb{R}$ on parle de **cointégration fractionnaire**

Cointégration, forte, faible et stationnaire

■ Selon les sous-espaces de δ et $\gamma = \delta - b$ on distingue

■ la **cointégration forte**

■ si $\delta > 1/2$ et $\delta - \gamma > 1/2$ (i.e. $b > 1/2$)

$\Rightarrow \varepsilon_t$ est stationnaire et les déviations à l'équilibre de long terme sont peu persistantes relativement à X_t

■ la **cointégration faible**

■ si $\delta \lesseqgtr 1/2$ et $\delta - \gamma < 1/2$ (i.e. $b < 1/2$)

$\Rightarrow \varepsilon_t$ est très persistant relativement à X_t impliquant de larges et durables déviations à l'équilibre de long terme

■ la **cointégration stationnaire**

■ si $\delta < 1/2$ et $\delta - \gamma < 1/2$ (i.e. $b < 1/2$)

$\Rightarrow \varepsilon_t$ et X_t sont **stationnaires** donc les déviations à l'équilibre de long terme sont persistantes relativement à X_t

\Rightarrow il s'agit d'un sous-cas très particulier de **cointégration faible**

Le cadre fractionnaire de l'approche en 2 étapes

- Soit $x_t \sim I(\delta)$ un processus stochastique d'innovations v_t et

$$y_t = \beta x_t + \varepsilon_t \text{ avec } \varepsilon_t \sim I(\gamma) \text{ et } \beta \neq 0$$

- Cheung et Lai (1993) proposent d'étendre l'approche en 2 étapes

1a on estime δ_x et δ_y par un estimateur $\hat{\delta}$ consistant

1b on test l'égalité entre δ_x et δ_y

1c on estime β par OLS

... on récupère les résidus estimés $\hat{\varepsilon}_t = y_t - \hat{\beta}x_t$

2a on estime γ par un estimateur $\hat{\gamma}$ consistant

2b on teste si $\hat{\varepsilon}_t \sim I(\gamma < \delta)$

Étape 1 dans le cadre fractionnaire

- 1a L'estimation des δ_x et δ_y doit se faire à l'aide d'un estimateur consistant
 - 1b La théorie de la cointégration requiert $\delta_x = \delta_y = \delta$
 - Hualde (2013) propose un test d'égalité sous la nulle que nous allons détailler
 - 1c Sous l'hypothèse que $(y_t, x_t) \sim I(\delta)$ et $\beta \neq 0$ comment se comportent les OLS ?
 - Cheung et Lai (1993) révèlent que $n^b(\hat{\beta} - \beta) \xrightarrow{p} 0$ avec $b = \delta - \gamma$
 - ⇒ ce résultat généralise ceux des Chapitres 1 et 2 ($b = 0$ et $b = 1$)
- Note Si $\varepsilon_t \perp v_t$ n'est pas respecté, le problème d'endogénéité subsiste

Étape 1b : test d'égalité des ordres

- Hualde (2013) propose la statistique de test suivante

$$\hat{t} = \frac{(2\pi n)^{-1/2} \hat{a}' \sum_{t=1}^n z_t(\hat{\delta}_x, \hat{\delta}_y)}{(\hat{a}' \hat{f}_{z(\cdot)}(0) \hat{a})^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1)$$

sous $H_0 : \delta_y = \delta_x$ contre $H_1 : \delta_y \neq \delta_x$ où $\hat{t} \sim n^{|\delta_x - \delta_y|}$ et

- $z_t = (y_t, x_t)'$ et $z_t(\hat{\delta}_y, \hat{\delta}_x) = (\Delta^{\hat{\delta}_y} y_t, \Delta^{\hat{\delta}_x} x_t)'$

- $\hat{a} = \left(\mathbb{1}(n^\kappa(\hat{\delta}_y - \hat{\delta}_x) > h_n), \mathbb{1}(n^\kappa(\hat{\delta}_y - \hat{\delta}_x) \leq h_n) \right)'$

avec $\hat{a} \xrightarrow{p} a := (1, 0)' \mathbb{1}(\delta_y > \delta_x) + (0, 1)' \mathbb{1}(\delta_y \leq \delta_x)$ et $n \rightarrow \infty$

pour $h_n = \log n^\kappa$ ou $h_n = n^{\kappa/2}$ un paramètre de réglage

- $\kappa > 0$ est un autre paramètre de réglage (e.g. $\kappa = 1/2$)

- $\hat{\delta}_y$ et $\hat{\delta}_x$ sont obtenus par Whittle ou **local Whittle multivariés**

$\Rightarrow \hat{f}_{z(\cdot)}(0)$ est la densité spectral en $\lambda = 0$ de $z_t(\hat{\delta}_y, \hat{\delta}_x)$

Attention, en présence de cointégration, $\hat{f}_{z(.)}(\lambda \rightarrow 0)$ tend vers la VCV décrivant les relations de long terme entre les séries.

Tout comme dans le domaine temporel, la matrice des relations de long terme est singulière car de rang réduit, ce qui implique que l'on ne peut pas l'inverser.

Néanmoins, le test de [Hualde \(2013\)](#) ne nécessite pas une inversion car au dénominateur, les coefficients \hat{a}' et \hat{a} font ressortir uniquement les $f_{z(.)}^{(ii)}(0)$ pour $i = 1, 2$.

Le test est donc valide sous l'unique condition que $f_{z(.)}^{(ii)}(0) > 0$

Étape 1c : l'endogénéité dans le cadre fractionnaire

- Tout comme le cas $I(1)/I(0)$, l'endogénéité pose problème
 - En présence de **cointégration forte**
 - ⇒ les OLS sont **biaisés** mais ce biais est $O_p(n^{-1})$
 - En présence de **cointégration stationnaire**
 - ⇒ les OLS sont **inconsistant** car le biais ne disparaît pas asymptotiquement
- ⇒ Il existe des méthodes analogues au FMOLS mais dans le domaine des fréquences pour traiter l'**endogénéité**
- ⇒ voir Aloy et de Truchis (2015) pour une revue de la littérature

Étape 2 dans le cadre fractionnaire

2a Sur la base des résidus collectés $\hat{\varepsilon}_t$, on estime γ à l'aide d'un estimateur consistant

2b Test d'absence de cointégration

- Wang et al. (2015) proposent la statistique de test suivante

$$\hat{F} = \frac{n^{-1/2} \sum_{t=1}^n \Delta^{\hat{\gamma}} x_t}{(2\pi \hat{f}_v(0))^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1)$$

sous $H_0 : \gamma = \delta$ contre l'alternative $H_1 : \gamma < \delta$ où $\hat{F} = O_p(T^{\delta-\gamma})$

- Précisons les notations

- $\hat{f}_v(0)$ est la densité spectral en $\lambda = 0$ de $\hat{v}_t = \Delta^{\hat{\delta}} x_t$
- $\hat{\delta}$ est obtenu par Whittle ou local Whittle
- $\hat{\gamma}$, estimé à partir de $\hat{\varepsilon}_t$, est obtenu par Whittle ou local Whittle
- $\hat{\varepsilon}_t$ est obtenu à partir d'un estimateur consistant de β

Un biais peut apparaître dans l'estimation de γ par local Whittle sous H_0 : $\gamma = \delta$, i.e. quand $\beta = 0$.

Cela semble lié à l'utilisation des OLS dans un cadre de régression factice, combiné à un estimateur qui converge $m^{1/2}$, mais à confirmer (voir Hualde et Velasco 2008 et Velasco 2003)

Dans ce cas le test Wang et al. (2015) est oversized mais on peut retrouver une bonne taille en estimant γ à l'aide d'un estimateur convergent en $n^{1/2}$, e.g. l'estimateur Whittle paramétrique.

Cointégration fractionnaire et systèmes triangulaires

- L'approche en 2 étapes de Cheung et Lai (1993) repose sur un système triangulaire sous-jacent

$$\begin{aligned} x_{1t} &= \beta x_{2t} + \Delta^{-\delta_1} \varepsilon_{1t}, \\ x_{2t} &= \Delta^{-\delta_2} \varepsilon_{2t}, \end{aligned} \quad (9)$$

- (9) est un système de cointégration si $\delta_2 > \delta_1 > 0$ et $\beta \neq 0$ et admet la représentation matricielle suivante,

$$\begin{pmatrix} 1 & -\beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} (1-L)^{-\delta_1} & 0 \\ 0 & (1-L)^{-\delta_2} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \quad (10)$$

où $\varepsilon = (\varepsilon_{1t}, \varepsilon_{2t})' \sim I(0)$

- De plus, si $1/2 > \delta_2 > \delta_1 > 0$, $z_t = (\varepsilon_{1t}, x_{2t})'$ est stationnaire et

$$\mathbb{E}\left((z_t - \mathbb{E}(z_t))(z'_{t+j} - \mathbb{E}(z'_t))\right) = \int_{-\pi}^{\pi} e^{i\lambda j} f_z(\lambda) d\lambda$$

avec $f_z(\lambda)$ une matrice de densité spectrale

Systèmes triangulaires stationnaires et densité spectrale

- Dans le cas stationnaire du système cointégré (9) on a donc

$$f_z(\lambda) = \Lambda(\delta) f_\varepsilon(\lambda) \Lambda^*(\delta)$$

avec $\Lambda(\delta) = \text{diag}((1 - e^{i\lambda})^{-\delta_1}, (1 - e^{i\lambda})^{-\delta_2})$ et si $\lambda \rightarrow 0$

$$f_z(\lambda) \sim \Lambda(\delta) G \Lambda^*(\delta), \quad \Lambda(\delta) \sim \text{diag}(\lambda^{-\delta_1}, \lambda^{-\delta_2})$$

- Le périodogramme matriciel du système (9) est particulier

$$I_z(\lambda) = \omega_z(\lambda) \omega_z^*(\lambda) = \begin{pmatrix} \omega_1(\lambda) - \beta \omega_2(\lambda) \\ \omega_2(\lambda) \end{pmatrix} \begin{pmatrix} \omega_1(\lambda) - \beta \omega_2(\lambda) \\ \omega_2(\lambda) \end{pmatrix}^* \quad (11)$$

car il tient compte du paramètre β dans le développement de

$$\omega_z(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n z_t e^{-it\lambda_j}, \quad z_t = (\varepsilon_{1t}, x_{2t})'$$

- Les 3 paramètres du système sont donc bien $\vartheta = (\delta_1, \delta_2, \beta)'$

Systèmes triangulaires stationnaires et estimation Whittle

- Nielsen (2007) propose un estimateur **local Whittle** joint de θ

$$\mathcal{Q}_W(\vartheta; \mathbf{x}) \approx \frac{1}{m} \sum_{j=1}^m \left(\log \det \Lambda(\delta) G \Lambda^*(\delta) + \text{tr} (\Lambda(\delta) G \Lambda(\delta)^*)^{-1} \hat{I}_z(\lambda) \right)$$

ou encore **local Whittle multivariée concentrée**

$$\mathcal{K}_W(\vartheta; \mathbf{x}) = \log \det \hat{G}_m(\delta) - 2(\delta_1 + \delta_2) \frac{1}{m} \sum_{j=1}^m \log \lambda_j$$

$$\text{où } \hat{G}_m(\delta) = \frac{1}{m} \sum_{j=1}^m \Lambda(\delta)^{-1} \hat{I}_z(\lambda) \Lambda^*(\delta)^{-1}$$

- Nielsen (2007) montre que pour $\vartheta = (\delta_1, \delta_2, \beta)'$

$$\sqrt{m} \text{diag}(I_2, \lambda_m^{\delta_1 - \delta_2}) (\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Xi^{-1}), \quad \Xi = \begin{pmatrix} \Xi^{(1)} & 0 \\ 0 & \Xi^{(2)} \end{pmatrix}$$

$$\text{où } \Xi^{(1)} = 2(G \odot G^{-1} + I_2) \text{ et } \Xi^{(2)} = 2G_{22}/(G_{11}(1 - 2\delta_2 + 2\delta_1))$$

- de Truchis (2013) propose la version **Whittle** paramétrique

Systèmes triangulaires non-stationnaires et taper

- Nous avons vu que pour $\delta \in [1/2, 3/2)$, $x_t \sim I(\delta)$ admet une **pseudo-densité spectrale** avec $z_t^{(s)} = \Delta^s x_t$

$$f_x(\lambda) := |1 - e^{i\lambda}|^{-2s} f_{z^{(s)}}(\lambda) \sim g|\lambda|^{-2\delta}, \quad s = \lfloor \delta + 1/2 \rfloor$$

- Néanmoins, le biais du périodogramme, déjà présent pour $|\delta| \in (0, 1/2)$, augmente considérablement si $|\delta| > 1/2$
- Velasco (1999) propose de réduire ce biais en **“biseautant”** le spectre aux extrémités, i.e. quand $\lambda \rightarrow 0$ et $\lambda \rightarrow 2\pi$
 - l'idée est de multiplier x_t par des poids $h_t \in [0, 1]$, valant 1 au centre et diminuant doucement vers 0 aux extrémités
 - La **DFT** modifiée est alors

$$\omega_x(\lambda_j) = \left(2\pi \sum_{t=1}^n h_t^2 \right)^{-1/2} \sum_{t=1}^n h_t x_t e^{i\lambda_j t}, \quad \lambda_j = \frac{2\pi j}{n}$$

- Un choix typique pour h_t , appelé **taper**, est $h_t = 1 - |(2t - n)n^{-1}|$

Cointégration non-stationnaires et estimation Whittle

- Shimotsu (2012) utilise le concept de **taper** pour étendre l'approche de Nielsen (2007) au cas **non-stationnaire**
- Soit un système cointégré décrit par (10) et $0 < \delta_1 < \delta_2 < \infty$
- Soit $I(\lambda)_z = \omega_z(\lambda)\omega_z^*(\lambda)$ où $\omega_z(\lambda)$ est obtenu à l'aide d'un **taper**
- Une approche **local Whittle multivariée concentrée** donne

$$\mathcal{K}_W(\vartheta; \mathbf{x}) = \log \det \hat{G}_m(\delta) - 2(\delta_1 + \delta_2) \frac{2}{(1 - \kappa)m} \sum_{j=2}^m \log \lambda_j$$

$$\text{où } \hat{G}_m(\delta) = \frac{2}{(1 - \kappa)m} \sum_{j=2}^m \Lambda(\delta)^{-1} \hat{I}_z(\lambda) \Lambda^*(\delta)^{-1} \text{ et } \kappa = 0.05$$

- Shimotsu (2012) montre que

$$\sqrt{m} \text{diag}(I_2, \lambda_m^{\delta_1 - \delta_2})(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Xi^{-1})$$

la VCV asymptotique Ξ est détaillée dans Shimotsu (2012), p. 269

Plan

- | | |
|--|------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

VAR co-fractionnaire

- Faute de temps, cette section ne sera pas traitée mais voici quelques références pour répondre à une éventuelle curiosité
 - Johansen et Nielsen (2012)
 - Johansen et Nielsen (2010)
 - Johansen (2008)
- Ces papiers sont très théoriques et dépassent largement le niveau M2 donc pour une approche plus appliquée voir Jones et al. (2014)

Plan

- | | |
|--|------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

VAR cointégré non-linéaire

- Faute de temps, cette section ne sera pas traitée mais voici quelques références pour répondre à une éventuelle curiosité
 - Hansen et Seo (2002)
 - Kristensen et Rahbek (2010)
 - Kristensen et Rahbek (2013)
 - Seo (2007)

Plan

- | | |
|--|------------------------------|
| 1 Le domaine des fréquences | 4 CVAR fractionnaire |
| 2 Estimation des processus $I(\delta)$ | 5 Cointégration non-linéaire |
| 3 Cointégration fractionnaire | 6 Conclusions |

Ce qu'il faut retenir

- La mémoire longue est importante pour
 - décrire précisément l'autocovariance des processus stationnaires
 - définir rigoureusement la frontière de la stationnarité
 - définir précisément où débute le risque de **régression fallacieuse**

⇒ L'estimation de δ est cruciale mais

- complexe et coûteuse dans le domaine temporel
 - facilitée dans le **domaine des fréquences**
 - gérable semi-paramétriquement dans le domaine des fréquences
- La prise en compte de la mémoire longue ouvre la porte à
 - la spécification de modèles de **cointégration fractionnaire**
 - l'estimation des systèmes triangulaires fractionnaires en 2 étapes
 - l'estimation des systèmes triangulaires fractionnaires en 1 étape

- Ailliot, P., Pene, F., 2015. Consistency of the maximum likelihood estimate for non-homogeneous Markov-switching models. *ESAIM : Probability and Statistics*, 19, 268-292.
- Aloy, M., de Truchis, G., 2015. Optimal Estimation Strategies for Bivariate Fractional Cointegration Systems and the Co-persistence Analysis of Stock Market Realized Volatilities. *Computational Economics* 48, 83-104.
- Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2016. Long-Memory Processes. Springer-Verlag.
- Cheung, Y., Lai, K., 1993. A fractional cointegration analysis of purchasing power parity. *Journal of Business & Economic Statistics* 11, 103-112.
- Dahlhaus, R., 1989. Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 17(4), 1749-1766.
- de Truchis, G., 2013. Approximate Whittle analysis of fractional cointegration and the stock market synchronization issue. *Economic Modelling* 34, 98-105.
- Engle, R. F., Granger, C. W. J. (1987). Co-integration and error correction : representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- Fox, R., Taqqu, M., 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics* 14, 517-532.
- Geweke, J., Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of time series analysis* 4, 221-238.
- Hansen, B.E., Seo, B., 2002. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics* 110, 293-318.

- Hosoya, Y., 1996. The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence. *Journal of econometrics* 73, 217-236.
- Hosoya, Y., 1997. A limit theory for long-range dependence and statistical inference on related models. *The Annals of Statistics* 25, 105-137.
- Hualde, J., 2013. A simple test for the equality of integration orders. *Economics Letters* 119, 233-237.
- Hurvich, C., Ray, B., 1995. Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. *Journal of Time Series Analysis* 16.
- Hurvich, C.M., Deo, R., Brodsky, J., 1998. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis* 19, 19-46.
- Johansen, S., 2008. A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory* 24, 651-676.
- Johansen, S., Nielsen, M.O., 2010. Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics* 158, 51-66.
- Johansen, S., Nielsen, M.O., 2012. Likelihood Inference for a Fractionally Cointegrated Vector Autoregressive Model. *Econometrica* 80, 2667-2732.
- Jones, M., Nielsen, M.O., Popiel, M.K., 2014. A fractionally cointegrated VAR analysis of economic voting and political support. *Canadian Journal of Economics* 14, 1-53.
- Kristensen, D., Rahbek, A., 2010. Likelihood-based inference for cointegration with nonlinear error-correction. *Journal of Econometrics* 158, 78-94.

- Kristensen, D., Rahbek, A., 2013. Testing and Inference in Nonlinear Cointegrating Vector Error Correction Models, *Econometric Theory*.
- Lieberman, O., Phillips, P.C.B. (2008). A complete asymptotic series for the autocovariance function of a long memory process. *Journal of Econometrics* 147, 99-103.
- Nielsen, M.O., 2007. Local Whittle Analysis of Stationary Fractional Cointegration and the Implied–Realized Volatility Relation. *Journal of Business & Economic Statistics* 25, 427-446.
- Phillips, P., Shimotsu, K., 2004. Local Whittle estimation in nonstationary and unit root cases. *The Annals of Statistics* 32, 656-692.
- Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence. *The Annals of statistics* 23, 1630-1661.
- Seo, B., 2007. Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity. *Journal of Econometrics* 137, 68-111.
- Shao, X., 2010. Nonstationarity-Extended Whittle Estimation. *Econometric Theory* 26, 1060-1087.
- Shimotsu, K., 2007. Gaussian semiparametric estimation of multivariate fractionally integrated processes. *Journal of Econometrics* 137, 277-310.
- Shimotsu, K., 2012. Exact local Whittle estimation of fractionally cointegrated systems. *Journal of Econometrics* 169, 266-278.
- Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53, 165-188.

- Velasco, C., 1999. Non-stationary log-periodogram regression. Journal of Econometrics 91, 325-371.
- Velasco, C., Robinson, P.M., 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229-1243.
- Wang, B., Wang, M., Chan, N.H., 2015. Residual-based test for fractional cointegration. Economics Letters 126, 43-46.
- Whittle, P. (1953). Estimation and information in stationary time series. Arkiv för Matematik, 2, 423-434.