Économétrie II

Ch. 4b Stationnarité & cointégration
L3 Économétrie – L3 MASS

Prof. Philippe Polomé, U. Lyon 2

Année 2015-2016

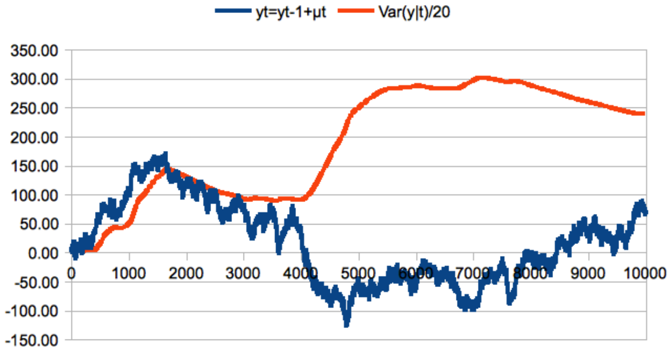
Table des matières

Table des matières

Marche aléatoire

- Dans le AR(1), l'hypothèse $|\rho| < 1$ est cruciale pour que la série soit faiblement dépendante (=I(0))
- Beaucoup de séries chronologiques économiques sont mieux caractérisées par un AR(1) avec $|\rho| = 1$:
  - $y_t = y_{t-1} + e_t$: une **marche aléatoire (random walk)**
- Variance d'une marche aléatoire $\nearrow$ linéairement avec le temps
  - Le processus ne peut donc être stationnaire
    - car sa distribution change dans le temps
  - Pas I(0) non plus
    - x_t et x_{t+h} ne deviennent pas presqu'indépendants lorsque $h \to \infty$
  - Donc pas utilisable en régression
- Puisque $E\left(e_{t+j}|y_t\right) = 0 \ \forall j \geq 1$, on a $E\left(y_{t+h}|y_t\right) = y_t \ \forall h \geq 1$
  - Donc, quel que soit l'éloignement temporel h, la meilleure prédiction pour y_{t+h} est y_t

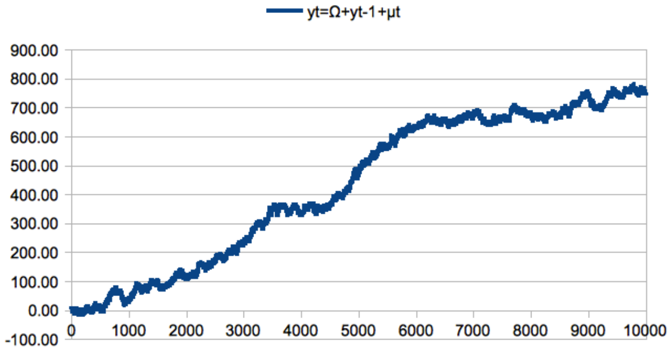
$y_t = y_{t-1} + \varepsilon_t$ avec $\varepsilon \sim n(0,4)$ et $y_0 = 0$ dans Trend....ods



I(1)

- Marche aléatoire = un cas de racine unitaire ou processus I(1)
- Un processus I(1) est "fortement persistant" ou "à mémoire longue"
  - "Tendance" $\neq$ "fortement persistent"
  - Des séries comme les taux d'intérêt, d'inflation ou de chômage sont souvent considérées "à mémoire longue"
    - mais n'ont pas de tendance claire
- Mais souvent, une série à mémoire longue a aussi une tendance claire
  - P.e. marche aléatoire avec dérive (random walk with drift) : $y_t = \Omega + y_{t-1} + e_t$
    - Ω est la dérive
    - Voir graphique

$$y_t = \Omega + y_{t-1} + \varepsilon_t \text{ avec } \varepsilon \sim n(0,4),\ y_0 = 0 \text{ et } \Omega = .05$$



Régression entre I(1), même sans trend

- Une régression simple entre 2 I(1) indépendantes va souvent résulter en une stat t significative
  - Même sans trend dans aucune variable
- Soit 2 marches aléatoires $y_t = y_{t-1} + e_t$ and $x_t = x_{t-1} + a_t$
  - On spécifie $y_t = \beta_0 + \beta_1 x_t + \xi_t$,
  - Alors $H_0 : \beta_1 = 0$ est vraie,
  - mais ξ_t contient y_{t-1} qui est une marche aléatoire,
- Alors l'inférence basée sur les résidus OLS mène à une stat t dont la distribution limite n'est pas normale
  - en fait cette stat $t \to \infty$ quand $T \to \infty$
  - Voir Trend....ods tab Spurious I(1)

Remarque : Autre régression spurieuse

- En coupe transversale
  - Régression spurieuse = cigognes
  - 2 variables sont reliées via une troisième
  - Si on régresse la 1º sur la 2º, on trouve une relation significative
    - mais si on intègre la 3º variable, la 2º perd sa significativité
  - Ce phénomène peut **aussi** arriver en séries temp
- Une relation spurieuse peut aussi être trouvée entre séries qui ont un trend
  - Si ces séries sont faiblement dépendante autour de leur trend, le problème se résoud en insérant un trend dans le modèle

Différence première

- La différence 1º d'une racine unitaire y_t
  - est faiblement dépendante (I(0))
    - x_t et x_{t+h} deviennent presqu'indépendants lorsque $h \to \infty$
  - et est souvent stationnaire
    - sa distribution ne change pas dans le temps
- Une série temporelle I(1) est souvent "différence-stationnaire"
  - Si y_t est I(1), alors $y_t - y_{t-1}$ est souvent I(0)
- Beaucoup de séries y_t qui sont > 0 sont t.q. $\ln(y_t)$ est I(1)
  - Alors on peut utiliser $\ln(y_t) - \ln(y_{t-1})$ en régression
  - Comme $\ln(y_t) - \ln(y_{t-1}) \approx \dfrac{y_t - y_{t-1}}{y_{t-1}}$ l'interprétation est en terme de taux de croissance
- Différencier une série temp avant de l'utiliser en régression retire aussi toute tendance linéaire

Table des matières

Corrélation

- $\rho_1 = Corr\left(y_t, y_{t-1}\right)$
  - Estimer ρ_1 à partir de la corrélation entre y_t et y_{t-1} dans l'échantillon
  - Se nomme **autocorrelation de 1º ordre** de $\{y_t\}$
- Les distributions d'échantillonnage de $\hat{\rho}_1$ sont très différentes lorsque ρ_1 proche de 1 et lorsque ρ_1 bien moins que 1
  - Quand ρ_1 proche de 1, $\hat{\rho}_1$ peut avoir un biais important à la baisse
  - On considère qu'il faut différencier si $\hat{\rho}_1 > .9$, voire $\hat{\rho}_1 > .8$
- Quand la série a un trend clair, on estime ρ_1 après avoir enlevé la tendance
  - Sinon, $\hat{\rho}_1$ tend à être surestimé

Test de racine unitaire

- Modèle AR(1) $y_t = \alpha + \rho y_{t-1} + e_t$
- Test **Dickey-Fuller** (DF) $H_0 : \rho = 1$ contre $H_1 : \rho < 1$
  - Soustraire y_{t-1} de chaque côté
    - $\Delta y_t = \alpha + \theta y_{t-1} + e_t$ avec $\theta = \rho - 1$
  - Sous H_0, y_{t-1} est I(1)
    - D'où la stat t associée à θ dans un MCO ne converge pas à une normale
    - mais bien à une distribution Dickey-Fuller,
  - On teste $\theta = 0$ (donc $\rho = 1$) par l'habituelle stat t
    - mais avec les valeurs tabulées de la Dickey-Fuller distribution
- Test de DF augmenté
  - On teste toujours pour $\rho = 1$ mais dans
  - $\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \cdots + \gamma_p \Delta y_{t-p} + e_t$

Application du test de DF

- $r3_t$ taux d'intérêt (annualisé) des bons du trésor à 3 mois
  - Bond equivalent yields, dans les pages financières
- Données dans INTQRT.gdt (Wooldridge)
  - Attention : changer structure du jeu de données :
  - mensuelles, date initiale inconnue
- Si on estime $\Delta y_t = \alpha + \theta y_{t-1} + e_t$
  - cr3 0 r3_1
  - Le coefficient de r3_1 est $-0{,}0907$, donc $\hat{\rho}$ =0.9093
  - t-stat sur r3_1 est -2.47, mais ne suit pas une distribution t
  - Sur la variable r3,
    - Prendre menu "variable" → "test de racine unitaire"
    - Pas de retard, avec constante, sans tendance
    - On retrouve les résultats de la régression, p-valeur .12 donc $\neg R\ H_0$: racine unitaire

Table des matières

Définition

- Prendre les différences 1º de séries I(1) avant de les utiliser en régression est "sûr"
  - mais limite l'analyse à des relations de court terme
  - et est peu efficient (estimateur peu précis)
  - La **cointégration** peut redonner un sens à des régressions entre séries I(1)
- Si $\{y_t\}$ et $\{x_t\}$ sont I(1), alors en général $y_t - \beta x_t$ est I(1) $\forall \beta$
  - Cependant, il est *possible* que pour un $\beta \neq 0$, $y_t - \beta x_t$ soit
    - I(0) : Asymptotiquement non corrélé
    - Stationnaire : Espérance & variance constantes
  - Si un tel β existe, on dit que $\{y_t\}$ et $\{x_t\}$ sont **cointégrées**
    - β est le paramètre de cointégration
- Interprétation :
  - $\{y_t\}$ et $\{x_t\}$ ne peuvent pas s'écarter trop l'un de l'autre dans le long terme

Exemple

- $r6_t$ taux d'intérêt (annualisé) des bons du trésor à 6 mois
  - $r3_t$ idem à 3 mois
- Données dans INTQRT.gdt (Wooldridge)
  - On a vu plus tôt que $r3_t$ avait une racine unitaire
  - Vrai aussi pour $r6_t$
- Soit $Ecart_t = r6_t - r3_t$
  - DF stat est -7.71 avec une p-valeur infime
  - donc RH_0 racine unitaire
  - donc $r6_t$ et $r3_t$ sont cointégrées avec paramètre 1
- Interprétation : si les taux s'écartaient, l'un des deux deviendrait un investissement plus attractif
  - son prix augmenterait
  - son rendement baisserait donc

Test de cointégration

- Si on a une valeur pour le coefficient de cointégration β
  - Alors on teste si $y_t - \beta x_t$ a une racine unitaire
- Si on ne connait pas β
  - Si y_t et x_t sont cointégrées
    - MCO est **consistant** pour β dans $y_t = \alpha + \beta x_t + u_t$
    - sinon, MCO spurieux et β faussement significatif
  - Test de Engle-Granger = Dickey-Fuller sur $\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$
    - Régresser $\Delta \hat{u}_t$ sur $\hat{u}_{t-1}$ avec une constante, sans retard $\Delta \hat{u}_t = \delta + \gamma \hat{u}_{t-1} + \xi_t$
    - Si $\hat{u}_{t-1}$ n'est pas significatif, alors $\hat{u}_t$ est I(0)
    - Donc y_t **et** x_t **sont cointégrées**
    - Le test est fait sur une distribution spéciale, pas sur une t
- Si y_t ou x_t a une dérive, c'est plus compliqué
  - Wooldridge 2012 p648
- Une tendance doit être modélisée

Exemple : cointégration entre fertilité et déduction fiscale

- Aux USA “personal exemption” est une déduction fiscale
  - Entre autres plus on a d’enfants (ou de dépendants), plus la réduction est grande
  - Le montant est assez faible, mais fluctue dans le temps
  - On peut donc imaginer lier la déduction et le nombre de naissances
- Fertil3.gdt (Wooldridge)
  - Modifier structure du jeu pour série temp, annuelle, début 1913
  - gfr naissances / 1000 femmes 15-44 ans
    - DF : p-valeur .80 donc $\neg R\ H_0$: racine unitaire
  - pe “personal exemption”, en $ réels
    - DF : p-valeur .45 donc $\neg R\ H_0$: racine unitaire
- Régressions
  - Niveau $gfr_t = \alpha + \beta pe_t + u_t$
  - Différences premières $\Delta gfr_t = \alpha + \beta \Delta pe_t + \Delta u_t$

gfr et pe

gfr_t	coef (p-val)			Δgfr_t	coef(p-val)		
Cst	**99.4 (0)**	**92.9 (0)**	**108.6 (0)**	Cst	-.08 (.92)	-.32 (.68)	**-3.45 (0)**
pe_t	.05 (.40)	-.06 (.36)	.03 (.66)	Δpe_t	-.05 (.27)	-.05 (.17)	-.05 (.19)
pe_{t-1}		-.02 (.83)	-.04 (.72)	Δpe_{t-1}		-.01 (.69)	-.009 (.75)
pe_{t-2}		.11 (.07)	.13 (.11)	Δpe_{t-2}		**.09 (0)**	**.09 (0)**
pe_{t-3}		-.005 (.93)	-.01 (.88)	Δpe_{t-3}		.04 (.17)	.04 (.15)
pe_{t-4}		.08 (.16)	**.02 (.04)**	Δpe_{t-4}		**-.04 (.04)**	**-.36 (.05)**
Pill (63)	**-27.8 (0)**	**-30.9 (0)**	.38 (.97)	Pill (63)	-2.23 (.07)	-1.78 (.14)	**-5.43 (.005)**
t			**-1.17 (0)**	t			**.11 (.01)**
DW	.12	.17	.25		1.44	1.34	1.57
T	72	68	68	T	71	67	67

La différence entre le modèle en niveaux et en différences 1º suggère de tester la cointégration car si les séries ne sont pas cointégrées, les régressions en niveaux sont spurieuses

gfr et pe

- Test de cointégration
  - Gretl : "modèle" → "Séries temp" → "Test de coint" → "Engle-Granger"
  - Variable gfr et pe, sans retard car on teste $\Delta \hat{u}_t = \alpha + \beta \hat{u}_{t-1} + \varepsilon_t$
  - Sortie complète
    - DF sur gfr et sur pe : I(1) chacune
    - MCO gfr sur pe
    - MCO résidus : $\neg R\ H_0 : \beta = 0$
    - Donc $\neg R\ H_0 : 1 - \beta = 1$: Résidus sont I(1)
    - Donc gfr et pe PAS cointégrés
- Contrôler pour une éventuelle tendance commune entre gfr et pe
  - Même procédure, mais sélectionner "constante et tendance temp"
  - Même conclusion
- Donc, la relation en niveau est spurieuse

Table des matières

Définition

- Si y_t et x_t sont I(1)
  - On ne peut estimer qu'un modèle en différences 1º
  - p.e. $\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + u_t$
- Mais, si de plus y_t et x_t sont cointégrées
  - On peut introduire des variables I(0) supplémentaires
  - Soit $s_t = y_t - \beta x_t$ qui est donc I(0)
    - Par simplicité on suppose $E(s_t) = 0$
  - Dans le cas le plus simple, on inclut un lag de s_t
    - $\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta s_{t-1} + u_t$
  - Le terme δs_{t-1} est appelé correction d'erreur
    - Ce modèle est dit à correction d'erreur

Discussion

- Un modèle à correction d'erreur permet d'étudier les dynamiques de court terme entre y_t et x_t
  - Il faut généralement estimer β
    - MCO est consistant
- Par simplicité, un modèle sans lags de Δy_t ni de Δx_t
  - $\Delta y_t = \alpha_0 + \gamma_0 \Delta x_t + \delta s_{t-1} + u_t$
  - $\Delta y_t = \alpha_0 + \gamma_0 \Delta x_t + \delta \left(y_{t-1} - \beta x_{t-1} \right) + u_t$
- Alors il faut que $\delta < 0$
  - Si $y_{t-1} - \beta x_{t-1} > 0$ alors y a dépassé l'équilibre en $t-1$
  - La cointégration impose qu'on y revienne
  - Comme $\delta < 0$ la correction d'erreur tend à réduire Δy_t
    - Ce qui ramene à l'équilibre
  - Même raisonnement lorsque $y_{t-1} - \beta x_{t-1} < 0$

Table des matières

Conclusions du Chapitre 4

- Autocorrélation des erreurs : Séries chronologiques
- On a vu test & correction pour AR(1)
  - Peuvent être développés pour tester et corriger autocorrélations plus complexes
- MCG et MCGF : corriger les biais induits par une forme d'autocorrélation **que l'on connait**
- Si autocorrélation mal évaluée : MCGF peut être pire que MCO
  - Par ex. résidus saisonniers et on suppose AR(1)
- On va plutôt essayer de mieux comprendre la série temporelle
  - MCO est consistant si les séries sont
    - Faiblement dépendantes I(0) : x_t et x_{t+h} presqu'indépendants lorsque $h \to \infty$
    - Stationnaires : leurs distributions ne changent pas dans le temps
  - Une série avec trend est régressée sur une tendance

Conclusions

- Quand les séries sont persistentes
  - Racines unitaire, I(1)
  - MCO produit des résultats spurieux
  - Les différences 1º sont souvent I(0)
- Deux séries sont cointégrées lorsque
  - Elles sont I(1)
  - Mais une de leurs combinaisons linéaires est I(0)
  - Alors, la régression de l'une sur l'autre indique une relation de long terme
    - Le court terme est étudié par un modèle à correction d'erreur

Devoir #4

Conception d'une feuille de tableur pour montrer l'effet de **l'autocorrélation** sur les coefficients estimés dans une régression linéaire MCO à une variable x et une constante.

1. Générer le terme d'erreur ε ; par exemple pour chaque observation i, on génère d'abord $\varepsilon_1 = n(0,1)$ puis $\varepsilon_i = \rho\varepsilon_{i-1} + \mu_i$ avec $|\rho| < 1$ et μ un bruit blanc
2. Illustrer que les MCO sont inconsistants ou non lorsqu'il y a autocorrélation
3. Créer deux séries I(1) sur le même principe
   3.1 Montrer que régresser l'une sur l'autre aboutit à des t-stat souvent très élevées, quelle que soit la taille de l'échantillon