Econométrie 2A
Période 4

Josephson Junior R.

April 28, 2024

Table des matières

Moindres Carrés Généralisés

Soit le modele suivant :

$$\begin{cases} Y = & X\beta + \varepsilon \\ o\grave{u} & E(\varepsilon'\varepsilon) = \Omega_\varepsilon \neq \sigma_\varepsilon^2 I \end{cases}$$

On estime le paramètre β par la MCG ou **estimateur d'Aitken** :

$$\hat{\beta}_{MCG} = \left(X'\Omega_\varepsilon^{-1}X\right)^{-1}\left(X'\Omega_\varepsilon^{-1}Y\right)$$

Sa matrice variance-covariance est donnée par :

$$\hat{\Omega}_{\hat{\beta}_{MCG}} = \left(X'\Omega_\varepsilon^{-1}X\right)^{-1}$$

NB : Il serait préferable d'adopter la notation suivante :

$$\varepsilon \sim \mathcal{N}(0, \Omega_\varepsilon)$$

Autocorrélation des erreurs

Causes

- Absence d'une variable explicative importante dont la valeur résiduelle permettrait de blanchir les erreurs
- Mauvaise spécififcation du modèle : la relation entre Y et X n'est pas linéaire
- Une lissage par moyenne mobile qui aurait créée une autocorrélation artificielle des erreurs.

Estimation du modèle

Soit le modèle suivant :

$$\begin{cases} Y = & X\beta + \varepsilon \quad & \varepsilon \sim AR(1) \\ \varepsilon_t = & \rho\varepsilon_{t-1} + \mu_t \quad & \mu_t \sim BB(0, \sigma_\mu^2) \end{cases}$$

Procédure d'estimation de ρ

1. Procédure 1 : par regression directe de e_t sur e_{t-1}

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t \; e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

2. Procédure 2 : par la statistique de DW :

$$\hat{\rho} = 1 - \frac{DW}{2}$$

Estimation des paramètres

Puisque les erreurs $u_t \sim AR(1)$ donc elles sont autocorellées d'ordre 1 $\Rightarrow$ la MCO n'est pas plus valide.
On applique la MCG au modèle initial ce qui revient à appliquer la MCO au modèle quasi-différence suivant : $Y_t - \hat{\rho} Y_{t-1} = \beta(X_t - \hat{\rho} X_{t-1}) + v_t$

$$\Longrightarrow y_t = \beta x_t + v_t$$

Test de Durbin et Watson

Le test est fait pour **identifier un problème d'autorrélation d'ordre 1** pour un modèle contenant **une constante** et un nombre d'observations $n \geq 15$.

$$\begin{cases} H0: & \rho = 0 \\ H1: & \rho \neq 0 \end{cases}$$

La statistique du test est définie par :

$$DW = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

La valeur du test est comprise entre [0,4] et possède sa propre table où l'on pourra déterminer d_1 et d_2. La décision du test se trouve sur l'intevalle ci-dessous.

0	$\rho < 0$	d_1	?	d_2	H0	2	H0	$4 - d_2$	?	$4 - d_1$	$\rho > 0$	4

Test de Breusch-Godfrey

- **Test de Fisher** : l'hypothèse est la même que celle du test DW. A la différence on applique le MCO pour l'équation intermédiaire suivante :

$$\hat{\varepsilon}_i = a + bx_i + \rho\hat{\varepsilon}_{i-1} + v_i$$

L'estimation des paramètres donne :

$$\hat{\beta} = \left(A'A\right)^{-1} \hat{A}'\hat{\varepsilon}$$

On admet la relation suivante :

$$SCR_{(2)} = \hat{\varepsilon}'\hat{\varepsilon} - \hat{\beta}'\hat{A}'\hat{\varepsilon}$$

La statistique de test est :

$$F = \frac{R^2_{(2)}}{1 - R^2_{(2)}} \frac{n - K - 2p}{p} \sim \mathcal{F}(p, n - K - 2p)$$

- **Test de LM** : LM pour dire Langrange Multiplier. La statistique de test est :

$$\chi_c^2 = (n-p)R_{(2)}^2 \sim \chi^2(p)$$

p est l'ordre de l'autorrélation à tester. La règle de décision est :

$$\begin{cases} \chi_c^2 \leq \chi_\alpha^2(p) : & \textbf{On accepte H0} \\ \chi_c^2 > \chi_\alpha^2(p) : & \textbf{On rejette H0} \end{cases}$$

Non-stationnarité et régression factice

Depuis le début du chapitre on a travaillé avec des modèles de régression simple ou multiple en supposant implicitement la stationnarité des series chronologiques. Mais dans la vie réelle certaines **variables macroéconomiques** et **financières** ne sont pas forcément de nature stationnaire.

C'est dans ce cadre que **Granger et Newfold (1974)** ont avancé le problème de régression factice. Ces régressions factices ont tendance à donner un coefficient **R^2 élevé** mais une statistique **DW faible** ce qui présage une forte autocorrélation des erreurs. De plus le **t-test** donne bien **la significativité des coefficients du modèle de régression**.

Comment résoudre le problème ?

Une première approche serait de **différencier** les séries non-stationnaires afin de les stationnariser et pouvoir appliquer sur eux les méthodes habituelles d'économétrie.
$\Rightarrow$ Cette opération de différenciation a pour limite essentielle de **masquer les propriétés de long terme** des séries étudiées.
Pour palier à ce problème nous allons introduire **le concept de cointégration** qui permet de **spécifier les relations à LT tout analysant la dynamique de CT entre les variables étudiées**.

Supposons y_t et x_t sont deux séries non-stationnaires et intégrées d'ordre d **I(d)** alors en général :

$$\mathbf{z_t = y_t - \alpha x_t \sim I(d)}$$

Mais des fois il se peut que z_t soit **I(d-b)** soit un ordre d'intégration plus bas que celui des deux series considérées.
$\Rightarrow$ On dit y_t et x_t sont **cointégrées** notées : $(\mathbf{y_t}, \mathbf{x_t}) \sim \mathbf{CI}(\mathbf{d}, \mathbf{b})$ avec :

$$\begin{cases} \alpha : & \textbf{le paramètre de cointégration} \\ (1, -\alpha) : & \textbf{vecteur de cointégration} \end{cases}$$

Le cas le plus fréquemment utilisé est $\mathbf{b} = \mathbf{d} = \mathbf{1}$.

Test de cointégration

L'algorihtme de **Engle et Granger** se déroule en deux étapes :

- **Etape 1** : Tester l'ordre d'intégration des variables. Deux séries ne peuvent être cointégrées que si **elles ont le même ordre d'intégration**. Pour rétrouver l'ordre d'intégration des séries on applique le test de **Dickey-Fuller** ou l'ADF aussi.
- **Etape 2** : Estimer par MCO la relation de LT définit par :

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

Pour que la relation de cointégration soit acceptée il faut que $\hat{\varepsilon}_t$ soit stationnaire c'est-à-dire :

$$\hat{\varepsilon}_t = y_t - \hat{\alpha} - \hat{\beta} x_t \sim \mathbf{I_0}$$

Pour tester la stationnarité des résidus on passe par le DF ou l'ADF test mais on utilise d'autres tables comme : **Engle et Yoo (1988)** ou bien **Mackinnon** qui dépendent du nombre d'observations T et le nombre de variables N (expliquer et explicatives).

Règle de décision:

- **Si les résidus sont non-stationnaire** alors la relation estimée est une régression factice.
- Sinon c'est **une relation de cointégration**

Théorème de représentation de Granger

Une des propriétés fondamentales des séries cointégrées est qu'elles peuvent être modéliser sous la forme d'un ECM. Ce résultat a été demontré dans le cadre du théorème de représentation de Granger valable pour les series cointégrés d'ordre 1. De tels modèles permettent **de modéliser les ajustements qui conduisent à un équilibre de LT**.
Pour $\mathbf{x_t}$ **et** $\mathbf{y_t} \sim \mathbf{I(1)}$ alors il existe une représentation à correction d'erreur de la forme :

$$\mathbf{\Delta y_t} = \gamma_0 + \gamma_1 \mathbf{\Delta x_t} + \mathbf{c}\hat{\varepsilon}_{\mathbf{t-1}} + \mathbf{u_t} \quad \textbf{où} \quad \hat{\varepsilon}_{\mathbf{t-1}} = \mathbf{y_{t-1}} - \hat{\alpha} - \hat{\beta}\mathbf{x_{t-1}}$$

- $\hat{\varepsilon}_{\mathbf{t-1}}$: résidu retardé d'une période d'estimation de la relation à LT.
- $\gamma_1 \mathbf{\Delta x_t}$: la composante dynamique de court terme du modèle.
- $\mathbf{c}$: vitesse d'ajustement de la variable endogène vers l'équilibre de LT (si $< \mathbf{0}$: **force de rappel** ; sinon **force de répulsion et le deséquilibre est toujours persistant**).

Le modèle ECM permet d'intégrer les fluctuations de CT autour de l'équilibre de LT donné par la relation de cointégration. Il décrit ainsi un processus d'ajustement et combine deux typees de variables :

- **Variables en différence prémière** : stationnaire ; représentatif de l'équilibre à court terme.
- **Variable en niveau** $(\hat{\varepsilon}_{\mathbf{t-1}})$: combinaison linéaire stationnaire de variables non stationnaire assurant la prise en compte de l'équilibre à LT.

Estimation et Validation de l'ECM : Méthode à deux étapes

Pour des séries **CI(1,1)** on estime l'ECM par la méthode à deux étapes de Engel et Granger.

- E1 : **Estimation par les MCO de la relation de LT**

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

- E2 : **Estimation par les MCO de l'ECM**

$$\Delta y_t = \gamma_0 + \gamma_1 \Delta x_t + c\hat{\varepsilon}_{t-1} + u_t$$

Validation du modèle : il est impératif que le coéfficient d'ajustement c soit **Significativement négatif** pour que le mécanisme à correction d'erreur existe. Dans le cas contraire, il n'existe pas de phénomène de retour à l'équilibre donc il convient de réjeter une spécification du modèle ECM.

Que se passerait-t-il si on prenais le cas de plusieurs variables explicatives qui sont toutes intégrées d'ordre 1.

$$y_t = \beta_0 + \beta_1 x_{1t} + ... + \beta_k x_{kt} + \varepsilon_t$$

On peut toujours revenir à l'application de la méthode à deux étapes d'Engel et Granger si **il s'agit d'une seule relation de cointégration (un seul vecteur de cointégration)**.

$$\Delta y_t = \gamma_0 + \gamma_1 \Delta x_{1t} + ... + \gamma_k \Delta x_{kt} + c\hat{\varepsilon}_{t-1} + u_t$$

Toutefois, dans le cas où l'on plus qu'un vecteur de cointégration il est impossible d'appliquer cette méthode. Dans ce cas, on a recours à **l'approche VECM** et aux tests de cointégration proposés par **Johansen** : **Test de trace et Test de la valeur propre maximale**.

Définition

La notion de retard est très importante dans un modèle de série temprelle car elle peut **exprimer une certaine forme d'inertie dans le comportement des agents économiques**.

Exemple : La fonction de consommation

$$\mathbf{C_t} = \beta_0 + \beta_1 \mathbf{R_t} + \beta_2 \mathbf{R_{t-1}} + \varepsilon_t$$

Dans ce modèle on peut déterminer **propensions marginales à consommer**, une à court terme qui est β_1 et une à long terme qui est $\beta_1 + \beta_2$.

On distingue deux types de modèle qui incorporent des retards :

- **Le modèle à retards échelonnés** :

$$\mathbf{y_t} = \alpha + \sum_{\mathbf{h=0}}^{\mathbf{q}} \beta_{\mathbf{h}} \mathbf{x_{t-h}} + \varepsilon_{\mathbf{t}} \Rightarrow \mathbf{y_t} \sim \mathbf{DL(q)} \quad (1)$$

- **Le modèle ARDL** :

$$\mathbf{y_t} = \mu + \sum_{\mathbf{h=0}}^{\mathbf{q}} \beta_{\mathbf{h}} \mathbf{x_{t-h}} + \sum_{\mathbf{k=1}}^{\mathbf{p}} \alpha_{\mathbf{k}} \mathbf{y_{t-k}} + \varepsilon_{\mathbf{t}} \Rightarrow \mathbf{y_t} \sim \mathbf{ARDL(p, q)} \quad (2)$$

Effets de court terme et long terme

Considérons le modèle DL(q), on peut distinguer deux effets sur la variables dépendante y suite **à une variation de x** :

- **Effet immédiat** : β_0 qui est le multiplicateur de court terme. Une variation de x augment la variable y de β_0 unité.
- **Effet cumulé** : $\sum \beta_h$ qui est appelé multiplicateur d'équilbre ou dynamique. Une augmatentation de x d'une unité se traduit par une hausse de y de $\sum \beta_h$

En fonction des polynomes retards on peut réécrire les deux modèles :

$$\mathbf{y_t} \sim \mathbf{DL(q)} \Rightarrow \mathbf{y_t} = \alpha + \mathbf{B(L)x_t} + \varepsilon_\mathbf{t} \quad où \quad B(L) = \sum_{h=0}^{q} \beta_h L^h$$

$$\mathbf{y_t} \sim \mathbf{ARDL(p,q)} \Rightarrow \mathbf{A(L)y_t} = \mu + \mathbf{B(L)x_t} + \varepsilon_\mathbf{t}$$

On définit pour chaque modèle le muliplicateur de long terme par :

$$(1)\ \mathbf{MLT} = \mathbf{B(1)} \quad ; \quad (2)\ \mathbf{MLT} = \frac{\mathbf{B(1)}}{\mathbf{A(1)}}$$

Pour que le modèle soit **stable** il faut que les racines du polynome A(L) soient supérieures à 1.

Retard Moyen

Le retard moyen exprime **le delai d'ajustement de y suite à un choc sur x**.

Pour le modèle (1) :

$$\mathbf{RM} = \frac{\mathbf{B}'(\mathbf{1})}{\mathbf{B}(\mathbf{1})}$$

Interpretation de RM : y retrouvve une nouvelle situation d'équilibre au bout de RM période après le choc sur x. C'est-à dire combien de temps en moyenne il faut pour que y soit à l'équilibre après un choc sur x.
Pour des modèles logarithmique les interprétations se font **en termes d'elasticité de court et de long terme**.

L'estimation des paramètres d'une modèle dynamique soulève deux de difficultés :

- **Difficulté de détermination du nombre de retard optimal** : on résout ce problème par les critères d'AIC ou BIC. On retient la valeur de q (nbre de retards) qui minimise ces critères.

$$\mathbf{AIC(q)} = \mathbf{ln}\left(\frac{\mathbf{SCR}}{\mathbf{T}}\right) + \frac{\mathbf{2q}}{\mathbf{T}} \; ; \; \mathbf{BIC(q)} = \mathbf{ln}\left(\frac{\mathbf{SCR}}{\mathbf{T}}\right) + \frac{\mathbf{q\,lnT}}{\mathbf{T}}$$

- **La multicolinéarité entre les variables explciatives** : on est amené à reduire le nombre de retards. Pour se faire des hypothèeses sont effectuées sur la forme du retard afin de réduire le nombre de paramètres à esimer. On distingue : **les modèles DL en nombre fini et les modèles DL en nombre infini**.

Les modèles DL en nombre fini : **les retards d'Almon**

La technique d'Almon permet d'éviter une estimation directe des coefficient β_h puisqu'elle consiste à supposer que la vraie distribution des retards peut être approchée oar un polynôme d'ordre h faible tel que $h < q$.

$$y_t \sim DL(q) \Rightarrow y_t = \mu + \beta_0 x_t + \ldots + \beta_q x_{t-q} + \varepsilon_t \quad (1)$$

$$\beta_{\mathbf{h}} = \sum_{\mathbf{i=0}}^{\mathbf{m}} \alpha_{\mathbf{i}} \mathbf{d}^{\mathbf{i}}$$

A titre d'exemple on va prendre i = 2.

$$\beta_h = \alpha_0 + \alpha_1 d + \alpha_2 d^2$$

On retrouve par cette approximation des nouvelles variables explicatives.

D'une manière générale on peut écrire les q retards et un polynôme de degré m comme suit :

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & q & q^2 & \dots & q^h \end{pmatrix} \begin{pmatrix} 0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \Rightarrow \beta = W\alpha$$

En guise de réecriture du modèle on a :

$$\mathbf{Y} = \mathbf{Z}\alpha + \varepsilon \ \ \textbf{où} \ \ \mathbf{Z} = \mathbf{X} \times \mathbf{W}$$

Il est à noter que la méthode d'Almon suppose que **l'on connaisse le degré m du polynôme utilisé pour l'approximation**.

Les modèles DL en nombre infini : **les retards de Koyck**

Pour ces modèles, l'effet de la variable explicative est illimité dans le temps. On suppose cependant que **le passé récent a plus d'influence que le passé lointain et que le poids des observationd passées tend à décroître au cours du temps**.

$$y_t \sim DL(\infty) \Rightarrow y_t = \mu + \sum_{h=0}^{\infty} \beta_h L^h x_t + \varepsilon_t \ (1)$$

Sous l'hypothèse que les β_h sont de même signe, Koyck suppose que **les retards decroissent de façon géométrique** telle que :

$$\beta_h = \lambda^h \beta_0 \quad (2) \qquad ; \qquad 0 < \lambda < 1$$

(2) dans (1) donne :

$$\mathbf{y_t} = \mu + \beta_\mathbf{0} \sum_{\mathbf{h=0}}^{\infty} \lambda^{\mathbf{h}} \mathbf{L}^{\mathbf{h}} \mathbf{x_t} + \varepsilon_\mathbf{t} \quad \mathbf{(3)}$$

Le polynôme retard s'écrit :

$$\mathbf{B(L)} = \frac{\beta_\mathbf{0}}{\mathbf{1} - \lambda \mathbf{L}} \quad \mathbf{(4)}$$

Portons (4) dans (3) :

$$\mathbf{y_t} = \mu(\mathbf{1} - \lambda) + \beta_\mathbf{0}\mathbf{x_t} + \lambda \mathbf{y_{t-1}} + \mathbf{v_t} \quad ; \quad \mathbf{v_t} = \varepsilon_\mathbf{t} - \lambda \varepsilon_\mathbf{t-1} \sim \mathbf{AR(1)}$$

Ainsi l'approche géométrique de Koyck a permi de passer d'un modèle $DL(\infty)$ à un modèle $\mathbf{ARDL(0,1)}$ avec des erreurs autocorrélées d'ordre 1. On a reduit les paramètres à estimer (réduction notable du problème de multicolinéarité).

Toutefois le modèle pose deux problème : **l(endogéneité et l'autocorrélation des erreurs**. Pour résoudre l'endogénéité on passe à la méthode VI en adoptant $\mathbf{x_{t-1}}$ **instrument de la variable endogène** $\mathbf{y_{t-1}}$

- **Multiplicateur de court terme**

$$MCT = \beta_0$$

- **Multiplicateur de long terme**

$$MLT = \frac{B(1)}{A(1)} = \frac{\beta_0}{1-\lambda}$$

- **Retard moyen**

$$RM = \frac{\lambda}{1-\lambda}$$

On montre que le cas particulier d'un ARDL(1,1) peut se ramener à un ECM.

$$\mathbf{y_t} \sim \mathbf{ARDL(1,1)} \Rightarrow \mathbf{y_t} = \mu + \beta_\mathbf{0}\mathbf{c_t} + \beta_\mathbf{1}\mathbf{c_{t-1}} + \alpha\mathbf{y_{t-1}} + \varepsilon_\mathbf{t}$$

$$\mathbf{ECM} \;:\; \mathbf{\Delta y_t} = \gamma_\mathbf{0} + \gamma_\mathbf{1}\mathbf{\Delta c_t} + \mathbf{c}\hat{\varepsilon}_\mathbf{t-1} + \mathbf{u_t}$$

Se souvenir que $\Delta y_t = y_t - y_{t-1}$; donc après réecriture on obtient :

$$\mathbf{\Delta y_t} = \mu + \beta_\mathbf{0}\mathbf{\Delta x_t} + (\alpha - \mathbf{1})\left[\mathbf{y_{t-1}} - \mathbf{x_{t-1}} \times \frac{\beta_\mathbf{1} + \beta_\mathbf{0}}{\mathbf{1} - \alpha}\right] + \varepsilon_\mathbf{t}$$

- $\beta_0 \Delta x_t$: composante dynamique de court terme du modèle.
- $(\alpha - 1)$: vitesse d'ajustement de la variable endogène vers l'équilibre de LT (si $< \mathbf{0}$: **force de rappel ; sinon force de répulsion)**