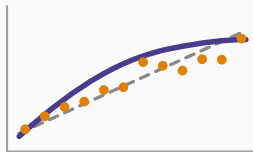


Économétrie avancée

Chapitre 18 : Sujets avancés en séries temporelles



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Avant de commencer

Modèles à retards échelonnés infinis

Tests de racine unitaire

Régression fallacieuse

Cointégration et modèles à correction d'erreur

Prévision

Synthèse

Avant de commencer

Pourquoi ce chapitre ?

Le chapitre 10 a posé les bases des séries temporelles avec des retards **finis**. Ce chapitre franchit trois étapes supplémentaires : autoriser des retards **infinis** mais paramétrés, tester formellement si une série est intégrée (racine unitaire), et exploiter des relations de long terme entre séries non stationnaires (cointégration) sans tomber dans le piège de la régression fallacieuse. On termine par les fondements de la prévision.

Modèles à retards échelonnés infinis

Retard distribué infini (IDL)

Un modèle IDL relie y_t à toute l'histoire de z : $y_t = \alpha + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + \dots + u_t$, avec $\delta_j \rightarrow 0$. Non estimable tel quel (nombre infini de paramètres), il exige une **paramétrisation**. Le **retard géométrique (Koyck)** pose $\delta_j = \gamma \rho^j$, $|\rho| < 1$: propension d'impact γ , propension de long terme $\gamma/(1 - \rho)$.

Estimation par transformation de Koyck

La différence pondérée $y_t - \rho y_{t-1}$ élimine la somme infinie et donne

$y_t = \alpha_0 + \gamma z_t + \rho y_{t-1} + u_t - \rho u_{t-1}$ – un modèle à variable dépendante retardée, mais dont l'erreur $u_t - \rho u_{t-1}$ est en général corrélée à y_{t-1} , rendant les MCO incohérents sauf hypothèse supplémentaire (ex. u_t suit elle-même un AR(1) de même paramètre ρ , auquel cas le modèle redevient dynamiquement complet et les MCO conviennent).

Investissement résidentiel et inflation des prix

Sur les données HSEINV, le retard géométrique est nettement rejeté (le retard $gprice_{t-1}$ ajouté est très significatif) au profit d'un **retard rationnel** plus flexible, qui ajoute un terme z_{t-1} à l'équation transformée. La propension de long terme diffère radicalement entre les deux spécifications (environ 4{,}7 pour le Koyck contre une valeur non significativement différente de zéro pour le rationnel) — illustration frappante du coût d'une dynamique mal spécifiée.

Tests de racine unitaire

Hypothèse nulle et statistique

Pour $y_t = \alpha + \rho y_{t-1} + e_t$, on teste $H_0 : \rho = 1$ (racine unitaire, $I(1)$) contre $H_1 : \rho < 1$ ($I(0)$) via la régression équivalente $\Delta y_t = \alpha + \theta y_{t-1} + e_t$, $\theta = \rho - 1$, et le test t usuel sur θ . Sous H_0 , y_{t-1} est $I(1)$: la statistique t **ne suit pas** une loi normale asymptotique, mais la distribution de **Dickey-Fuller**, tabulée séparément (valeurs critiques plus négatives que $-1,65$, ex. $-2,86$ au seuil de 5 %).

Test augmenté (ADF) et tendance

Le **test de Dickey-Fuller augmenté** ajoute des retards de Δy_{t-j} pour blanchir l'autocorrélation résiduelle : $\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + e_t$; le test sur θ reste inchangé. Pour une série manifestement tendancielle, ajouter un terme de tendance δt à la régression (test **avec tendance**) exige des valeurs critiques encore plus négatives (ex. $-3,41$ à 5 %), car détendre une série $I(1)$ la rend artificiellement plus proche d'un processus $I(0)$.

Racine unitaire dans le taux d'inflation américain

Sur données annuelles 1948–1996 (test ADF, un retard), la statistique t sur inf_{t-1} vaut $-3,01$, inférieure au seuil critique de 5 % ($-2,86$) : on **rejette** la racine unitaire pour l'inflation. À l'inverse, pour le log du PIB réel américain (test avec tendance), $t = -2,41$ ne permet pas de rejeter H_0 – conclusion classique mais fragile compte tenu de la faible puissance de ces tests en petit échantillon.

Régression fallacieuse

Le piège des séries $I(1)$ indépendantes

Le résultat de Granger et Newbold

Si $\{x_t\}$ et $\{y_t\}$ sont deux marches aléatoires **indépendantes**, régresser y_t sur x_t produit néanmoins une statistique t significative bien plus souvent que le seuil nominal (environ 66 % au lieu de 5 % pour $n = 50$, et cela empire avec n). Le R^2 ne converge pas vers zéro mais vers une **variable aléatoire**. C'est la **régression fallacieuse** : aucune relation économique n'existe, mais l'inférence usuelle suggère le contraire.

Pourquoi les MCO usuels échouent

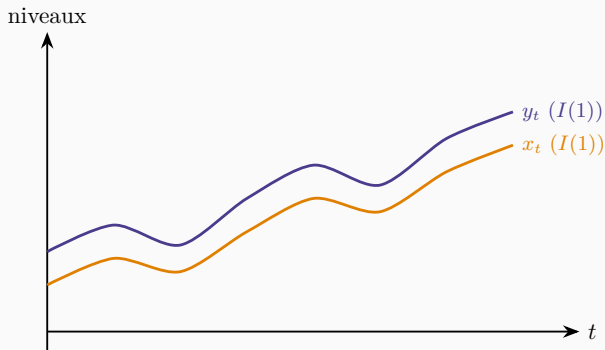
Sous $H_0 : \beta_1 = 0$, l'erreur du modèle en niveau doit être $u_t = y_t$ lui-même – c'est-à-dire une marche aléatoire. Les hypothèses de Gauss-Markov asymptotiques (chapitre 11), qui exigent un processus d'erreur faiblement dépendant, sont violées de façon flagrante : aucune des propriétés usuelles des MCO (normalité asymptotique du t , convergence du R^2) ne tient.

Cointégration et modèles à correction d'erreur

Une relation de long terme entre séries $I(1)$

Cointégration

Deux séries $I(1)$, y_t et x_t , sont **cointégrées** s'il existe $\beta \neq 0$ tel que $y_t - \beta x_t$ soit $I(0)$ – la combinaison linéaire ne dérive pas, malgré la non-stationnarité de chaque série prise isolément. β est le **paramètre de cointégration**. Test : régresser y_t sur x_t (MCO, consistant sous cointégration même si H_0 pose l'absence de cointégration), puis appliquer un test ADF aux résidus avec des valeurs critiques spécifiques (Engle-Granger), plus négatives que celles du test de racine unitaire standard.



Prévision

Perte quadratique et espérance conditionnelle

Pour prévoir y_{t+1} à partir de l'information I_t disponible en t , sous perte quadratique $E(e_{t+1}^2 \mid I_t)$, le prédicteur optimal est l'**espérance conditionnelle** $E(y_{t+1} \mid I_t)$. Pour un processus à différence de martingale, cette espérance est simplement la moyenne inconditionnelle — le passé n'aide pas à prévoir le futur (approximation usuelle pour les rendements financiers).

Modèles de prévision et VAR

Le modèle $y_t = \delta_0 + \alpha_1 y_{t-1} + \gamma_1 z_{t-1} + u_t$ évite d'avoir à prévoir z lui-même (contrairement au modèle statique en z_t contemporain). Généralisé à plusieurs séries, on obtient un **modèle vectoriel autorégressif (VAR)** : chaque variable est expliquée par ses propres retards et ceux des autres variables du système. Le test que z **cause au sens de Granger** y revient à tester la significativité jointe des retards de z dans l'équation de y , une fois les retards propres de y contrôlés.

Causalité de Granger : une mise en garde

« z cause y au sens de Granger » signifie seulement que le passé de z améliore la prévision de y au-delà du seul passé de y — cela ne dit **rien** sur une causalité contemporanée ni sur l'exogénéité de z dans une équation structurelle. Le terme est trompeur mais consacré par l'usage.

Critères hors échantillon : RMSE et MAE

Un ajustement en échantillon élevé (R^2 ajusté) ne garantit pas une bonne performance prédictive. On réserve les m dernières observations pour comparer les prévisions $\hat{f}_{n+h}$ aux valeurs réalisées :

$RMSE = \left(m^{-1} \sum_h \hat{e}_{n+h}^2\right)^{1/2}$ et $MAE = m^{-1} \sum_h |\hat{e}_{n+h}|$. On préfère le modèle qui minimise ces critères **hors échantillon**, pas nécessairement celui au meilleur R^2 en échantillon.

Prévoir le taux de chômage américain

Un modèle AR(1) et un modèle ajoutant l'inflation retardée sont comparés hors échantillon (1990–1996, sept observations réservées). Le second modèle domine sur les deux critères (RMSE = 0{,}550 contre 0{,}632; MAE = 0{,}362 contre 0{,}515) : l'inflation retardée améliore la prévision, confirmant *in fine* le résultat de Granger-causalité obtenu en échantillon.

Synthèse

L'essentiel du chapitre 17

- Les retards distribués infinis (Koyck, rationnels) modélisent une dynamique riche avec peu de paramètres, au prix d'hypothèses sur la structure des erreurs.
- Le test de Dickey-Fuller (et sa version augmentée) permet de tester formellement une racine unitaire, avec des valeurs critiques non standard.
- Régresser deux séries $I(1)$ indépendantes en niveaux produit typiquement une régression fallacieuse — statistiquement significative sans contenu économique.
- La cointégration légitime les régressions en niveaux entre séries $I(1)$ quand une combinaison linéaire est $I(0)$; le modèle à correction d'erreur en tire une dynamique de court terme cohérente avec l'équilibre de long terme.
- La prévision optimale sous perte quadratique est l'espérance conditionnelle; l'évaluation d'un modèle de prévision doit se faire hors échantillon (RMSE, MAE), pas seulement sur l'ajustement en échantillon.

Vers le chapitre 18

Après avoir couvert la théorie économétrique de bout en bout, le chapitre suivant change de nature : il s'agit de méthodologie de recherche appliquée — comment structurer, documenter et présenter un