

Francis Bismans
Olivier Damette

L3
Master
Doctorat

Économétrie dynamique

Modèles et applications



ellipses

Chapitre 1

Économétrie et séries temporelles

Sans entrer dans de longues discussions terminologiques, on se limitera à dire que l'économétrie est cette discipline qui combine théorie économique et analyse statistique. L'économétrie dynamique, elle, se concentre essentiellement sur la théorie économétrique fondée sur l'utilisation de données provenant de séries temporelles. En un mot, l'économétrie dynamique est donc tout simplement l'analyse économétrique de séries chronologiques.

1.1 Introduction

Le modèle linéaire de la régression multiple – avec la restriction qu'il ne comporte qu'une seule équation – peut se formuler comme suit :

$$y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i, \quad i = 1, \dots, n, \quad (1.1)$$

où y_i est la variable dépendante sur laquelle on dispose de n observations, les x_{ij} constituent les variables indépendantes ou explicatives et u_i est une variable aléatoire présentant les propriétés adéquates.

L'équation (1.1) se réécrit aisément sous la forme :

$$y_i = (1, x_{i2}, \dots, x_{ik}) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + u_i = \mathbf{x}'\boldsymbol{\beta} + u_i, \quad i = 1, \dots, n. \quad (1.2)$$

(Conventionnellement, un vecteur ligne se notera en gras et on lui associera le prime ('), tandis qu'un vecteur colonne sera simplement désigné par une lettre minuscule, elle aussi en gras.)

Cependant, le modèle formulé dans (1.1) ou dans (1.2) n'a rien de dynamique, la dimension chronologique n'y étant même pas présente. De ce point de vue, on peut dans un premier temps dater les variables du modèle et y introduire des variables explicatives décalées, ce qui donne

$$y_t = m + \beta_{10}x_{t1} + \beta_{11}x_{t-1,1} + \cdots + \beta_{1k_1}x_{t-k_1,1} + \beta_{20}x_{t1} + \beta_{22}x_{t-1,2} + \cdots + \beta_{2k_2}x_{t-k_2,2} \\ + \cdots + \beta_{K0}x_{tm} + \beta_{K1}x_{t-1,m} + \cdots + \beta_{Kk_K}x_{t-k_K,K} + u_t, \quad t = 1, \dots, T.$$

En utilisant une double sommation, l'équation précédente devient

$$y_t = m + \sum_{j=0}^{k_i} \sum_{i=1}^K \beta_{ij}x_{t-j,i} + u_t, \quad t = 1, \dots, T. \quad (1.3)$$

Un pas supplémentaire dans la dynamisation du modèle consistera à ajouter p variables dépendantes décalées à (1.3), soit

$$y_t = m + \sum_{r=1}^p \alpha_r y_{t-r} + \sum_{j=0}^{k_i} \sum_{i=1}^K \beta_{ij}x_{t-j,i} + u_t, \quad t = 1, \dots, T. \quad (1.4)$$

1.2 Modèles de séries temporelles stationnaires

Les modèles de séries chronologiques comportent plusieurs classes, dont celles dite ARMA ou encore celle des modèles proprement financiers du type GARCH.

1.2.1 Quelques résultats préliminaires

Considérons le **processus stochastique à temps discret** $(X_t, 1, \dots, T)$. De manière informelle, on pourrait définir un tel processus comme « une variable aléatoire plongée dans le temps ». Une réalisation de ce processus, notée $(x_t, t = 1, \dots, T)$, est appelée **série temporelle**.

L'**autocovariance d'ordre** k de la série $(x_t, t = 1, \dots, T)$, que l'on désignera par γ_k , est la covariance entre x_t et x_{t-k} :

$$\gamma_k = \text{Cov}(x_t, x_{t-k}) = E[(x_t - \mu_t)(x_{t-k} - \mu_{t-k})],$$

où μ_t est la moyenne de la série en t et $k = 0, \pm 1, \pm 2, \dots$. (Bien entendu, si $k = 0$, on obtient alors la variance.) La **fonction d'autocovariance** est une application de \mathbb{Z} dans \mathbb{R} . Comme elle est paire, c'est-à-dire que pour tout k , $\gamma_k = \gamma_{-k}$, son graphe est limité aux valeurs non négatives de k .

À partir des autocovariances, on introduit la notion d'autocorrélation, qui s'obtient en divisant chaque autocovariance de la série considérée par la variance correspondante. L'**autocorrélation d'ordre k** de $(x_t, t = 1, \dots, T)$ est ainsi donnée par

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = 0, \pm 1, \pm 2, \dots$$

On peut alors définir la fonction d'autocorrélation en parallèle à celle d'autocovariance :

Définition 1.1. On appelle **fonction d'autocorrélation** l'application ρ de \mathbb{Z} dans \mathbb{R} telle que

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k \in \mathbb{Z}.$$

Tout comme la fonction d'autocovariance, la fonction ρ_k est paire. Elle en diffère cependant du fait que $\rho_k = \gamma_k / \gamma_0 = 1$. Cette fonction est représentée graphiquement en se limitant aux valeurs non négatives de k , compte tenu de sa parité. Elle porte alors le nom de **corrélogramme**.

Introduisons à présent la notion de stationnarité d'un processus ou d'une série discrète.

Définition 1.2. Une série temporelle $(x_t, t \in \mathbb{Z})$ est dite **faiblement stationnaire** ou encore **stationnaire au second ordre** si pour tout $t \neq s$ et k arbitraire

- (i) $E(x_t) = E(x_s) = \mu < \infty$;
- (ii) $\text{Var}(x_t) = \text{Var}(x_s) = \sigma^2 < +\infty$;
- (iii) $\text{Cov}(x_t, x_{t-k}) = \text{Cov}(x_s, x_{s-k})$.

Exemple 1.1

Soit le processus **bruit blanc gaussien** défini par $(\varepsilon_t) \sim IIN(0, \sigma^2)$, où pour tout t , les variables aléatoires sont indépendamment et identiquement normalement distribuées. Ce processus est stationnaire au second ordre, puisque

- (i) $\forall t, E(\varepsilon_t) = 0$;
- (ii) $\forall t, \text{Var}(\varepsilon_t) = \sigma^2$;
- (iii) $\forall t \neq s \text{ et } k \neq 0, \text{Cov}(\varepsilon_t, \varepsilon_{t+k}) = \text{Cov}(\varepsilon_s, \varepsilon_{s+k})$.

Cette dernière égalité est vérifiée compte tenu que les ε_t sont indépendamment distribués.

Donnons à présent quelques précisions sur l'**opérateur retard** L défini par

$$L^j x_t = x_{t-j}, \quad j = 0, 1, \dots$$

Formellement, L peut être traité comme une variable, ce qui permet d'écrire un polynôme du p -ième degré en L , noté $\phi(L)$ et appelé **polynôme retard**, sous la forme

$$\phi(L) = \phi_0 + \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p. \quad (1.4)$$

En multipliant les deux membres de (1.4) par x_t , il vient alors

$$\begin{aligned} \phi(L)x_t &= (\phi_0 + \phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p)x_t \\ &= \phi_0 x_t + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} \\ &= \sum_{j=0}^p \phi_j x_{t-j}. \end{aligned}$$

Comme nous aurons l'occasion de le voir par la suite, l'opérateur retard permet de réaliser multiplications et divisions de séries temporelles. En particulier, dans la section suivante, nous utiliserons le résultat suivant :

Proposition 1.1 Si $|\phi| < 1$, alors

$$(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \dots)x_t = (1 - \phi)^{-1} x_t. \quad (1.5)$$

Démonstration

En multipliant les deux membres de (1.5) par $(1 - \phi L)$, il vient

$$(1 - \phi)(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \dots)x_t = x_t,$$

ce qui donne après réalisation du produit dans le premier membre

$$(1 - \phi L + \phi L - \phi^2 L^2 + \phi^2 L^2 - \phi^3 L^3 + \dots)x_t = \lim_{n \rightarrow +\infty} (1 - \phi^n L^n)x_t.$$

Compte tenu que $|\phi| < 1$, $\phi^n L^n$ tend vers 0 lorsque $n \rightarrow +\infty$ et en conséquence, (1.5) est vérifié. ♦

1.2.2 Processus autorégressifs

Commençons par les modèles autorégressifs, dits AR. (Cet acronyme est l'abréviation de « AutoRégressif ».)

Définition 1.3. Un **processus autorégressif d'ordre p** , noté $AR(p)$, s'écrit sous la forme

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t, \quad t = 1, \dots, T, \quad (1.6)$$

où $\varepsilon_t \sim bb(0, \sigma^2)$ est un bruit blanc de variance σ^2 . On note alors

$$y_t \sim AR(p).$$

Soit, à titre d'illustration, le modèle $AR(1)$

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \quad t = 1, \dots, T. \quad (1.7)$$

On vérifie sans peine que ce processus est stationnaire. En utilisant l'opérateur retard, il vient

$$\begin{aligned} y_t &= \phi L y_t + \varepsilon_t \\ (1 - \phi L) y_t &= \varepsilon_t. \end{aligned}$$

La proposition 1.1 permet alors d'écrire

$$\begin{aligned} y_t &= (1 - \phi L)^{-1} \varepsilon_t = (1 + \phi L + \phi^2 L^2 + \dots) \varepsilon_t \\ &= \varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \dots \end{aligned} \quad (1.8)$$

Si $|\phi| < 1$, la suite (ϕ, ϕ^2, \dots) converge. Il s'ensuit que la **condition de stationnarité** d'un $AR(1)$ est

$$|\phi| < 1.$$

Alternativement, cette condition est équivalente à ce que la solution de l'équation en l'opérateur retard

$$1 - \phi L = 0$$

soit supérieure à l'unité en **valeur absolue**, étant donné que si $|\phi| < 1$,

$$L = \frac{1}{\phi} > 1.$$

Toujours à titre d'illustration, voyons comment se comporte la fonction d'autocorrélation d'un AR(1). Multiplions d'abord (1.7) par y_{t-k} tout en prenant les espérances des deux membres :

$$E(y_t y_{t-k}) - \phi E(y_{t-1} y_{t-k}) = E(y_{t-k} \varepsilon_t). \quad (1.9)$$

Si $k > 0$, en servant de (1.8), on obtient

$$\begin{aligned} E(y_{t-k} \varepsilon_t) &= E[(\varepsilon_{t-k} + \phi \varepsilon_{t-k-1} + \phi^2 \varepsilon_{t-k-2} + \dots) \varepsilon_t] \\ &= E(\varepsilon_{t-k} \varepsilon_t) + \phi E(\varepsilon_{t-k-1} \varepsilon_t) + \phi^2 E(\varepsilon_{t-k-2} \varepsilon_t) + \dots \\ &= 0. \end{aligned}$$

Dès lors, toujours pour $k > 0$, en introduisant ce dernier résultat dans (1.9), il vient finalement

$$\gamma_k - \phi \gamma_{k-1} = 0. \quad (1.10)$$

Divisons (1.10) par γ_0 , ce qui conduit à l'équation de récurrence du premier ordre (sans terme constant) :

$$\rho_k = \phi \rho_{k-1}. \quad (1.11)$$

La solution de cette dernière équation s'obtient aisément par itération successive :

$$\begin{aligned} \rho_k &= \phi(\phi \rho_{k-2}) = \phi^2 \rho_{k-2} \\ &= \phi^2(\phi \rho_{k-3}) = \phi^3 \rho_{k-3} \\ &\vdots \\ &= \phi^{k-2}(\phi \rho_1) = \phi^{k-1} \rho_1 \\ &= \phi^k \rho_0 \\ &= \phi^k \quad (\rho_0 = \gamma_0 / \gamma_0 = 1). \end{aligned} \quad (1.12)$$

Au vu de (1.12), compte tenu que $|\phi| < 1$, on aura la proposition suivante :

La fonction d'autocorrélation (FAC) d'un processus AR(1) décroît exponentiellement vers 0 si $\phi > 0$. Elle décroît également vers zéro, mais en oscillant, si $\phi < 0$.

Cet ensemble de résultats se généralise au processus AR(p) tel que défini par (1.6). D'abord, ce modèle autorégressif peut s'écrire en utilisant l'opérateur retard :

$$(1 - \phi L - \phi^2 L^2 - \dots - \phi^p L^p) y_t = \varepsilon_t,$$

dont on peut extraire l'équation polynomiale

$$1 - \phi L - \phi^2 L^2 - \dots - \phi^p L^p = 0. \quad (1.13)$$

Dans le cas d'un modèle AR(1), la **condition de stationnarité** était que le paramètre ϕ , solution de l'équation du premier degré en L , soit supérieur à l'unité en valeur absolue. Eu égard à (1.13), cette condition implique que ses p racines – certaines d'entre elles pouvant être confondues ou complexes – soient **supérieures à l'unité en module**. En module et non en valeur absolue, parce que, répétons-le, des solutions peuvent se présenter sous la forme de complexes conjuguées. (On dit aussi, pour exprimer cette même condition, que les racines de (1.13) sont **extérieures au cercle unité**.)

La fonction d'autocorrélation, l'analogue de (1.11) pour un modèle AR(p), est donnée par l'équation de récurrence

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > 0. \quad (1.14)$$

Les racines de l'équation auxiliaire correspondant à (1.14)

$$m^p - \phi_1 m^{p-1} - \dots - \phi_{p-1} m - \phi_p = 0$$

sont soit réelles soit complexes. Dans le premier cas, si ces racines sont de plus inférieures à l'unité en valeur absolue, alors les solutions de (1.14) sont constituées d'exponentielles amorties. Dans le second cas, si les racines complexes conjuguées sont inférieures à l'unité en module, alors ces solutions sont des oscillations sinusoïdales amorties. En conséquence, il s'ensuit que la FAC est ici **une combinaison d'exponentielles amorties et d'oscillations également amorties**.

1.2.3 Les modèles ARMA

De tels modèles comportent deux parties : une partie autorégressive et une partie moyenne mobile désignée par l'acronyme MA, celui-ci étant une abréviation de « Moving Average », soit moyenne mobile en français. Nous commencerons par l'étude de cette dernière classe de modèles.

Définition 1.4. On appelle **processus moyenne mobile d'ordre q** un processus $(y_t, t \in \mathbb{Z})$ tel que

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \quad \forall t \in \mathbb{Z}, \quad (1.15)$$

où pour tout i , θ_i est un nombre réel et (ε_t) est un processus bruit blanc de moyenne nulle et de variance constante σ^2 .

On note alors

$$y_t \sim \text{MA}(q).$$

Le processus MA(1) s'écrit pour sa part

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1}, \quad \varepsilon_t \sim \text{bb}(0, \sigma^2). \quad (1.16)$$

Il est toujours stationnaire eu égard à ce que (ε_t) est un bruit blanc stationnaire.

En utilisant l'opérateur retard, (5.16) devient

$$\begin{aligned} y_t &= \varepsilon_t - \theta \varepsilon_{t-1} \\ &= (1 - \theta L) \varepsilon_t. \end{aligned}$$

Si $|\theta| < 1$, alors

$$(1 - \theta L)^{-1} y_t = (1 + \theta L + \theta^2 L^2 + \cdots) y_t = \varepsilon_t. \quad (1.17)$$

Autrement dit, pour obtenir une représentation autorégressive (infinie) du processus, la restriction $|\theta| < 1$ doit être imposée. Une telle restriction est connue sous de **condition d'inversibilité** du MA(1).

Voyons comment déterminer la fonction d'autocorrélation correspondante.

La variance, γ_0 en l'occurrence, est donnée par

$$\begin{aligned} \gamma_0 &= E[y_t^2] = E[(\varepsilon_t - \theta \varepsilon_{t-1})(\varepsilon_t - \theta \varepsilon_{t-1})] \\ &= E(\varepsilon_t^2) + \theta^2 E(\varepsilon_{t-1}^2) \\ &= (1 + \theta^2) \sigma^2. \end{aligned}$$

L'autocovariance d'ordre un est égale à

$$\begin{aligned} \gamma_1 &= E(y_t y_{t-1}) = E[(\varepsilon_t - \theta \varepsilon_{t-1})(\varepsilon_{t-1} - \theta \varepsilon_{t-2})] \\ &= -\theta E(\varepsilon_{t-1}^2) \\ &= -\theta \sigma^2. \end{aligned}$$