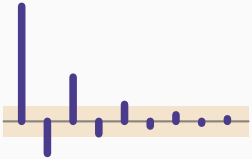
Modélisation ARMA des rentabilités financières

Chapitre 14 : L'estimation — moindres carrés et maximum de vraisemblance



Jaouad Madkour

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

Plan du chapitre

Avant de commencer

Les ordres p et q sont fixés. Il reste à chiffrer les coefficients $\alpha_1, \dots, \alpha_p$, $\beta_1, \dots, \beta_q$, la constante et σ_ε^2.

Une difficulté propre aux modèles MA

Un AR se régresse sur des variables **observées**. Un MA se régresse sur des chocs **inobservables**. Cette différence commande tout le chapitre.

Les moindres carrés

MCO sur un AR(p)

$$y_t = \theta + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon_t$$

est une régression linéaire ordinaire : les régresseurs sont des valeurs passées de la série, toutes observées.

Propriétés

L'estimateur des MCO est **biaisé** à distance finie — les régresseurs ne sont pas exogènes au sens strict — mais **convergent** et asymptotiquement normal, si le processus est stationnaire.

L'ordre de grandeur du biais

Pour un AR(1), $\mathrm{E}(\hat{\alpha}) \approx \alpha - (1 + 3\alpha)/T$. Avec $T = 1000$, le biais est de l'ordre de $0{,}003$: négligeable sur données financières, sensible sur données annuelles.

Pourquoi cela échoue pour un MA

L'obstacle

$$y_t = \theta + \varepsilon_t + \beta\varepsilon_{t-1}$$

Le régresseur ε_{t-1} n'est pas observé. Et il ne peut pas l'être : il dépend de β, que l'on cherche précisément à estimer.

La solution : la récursion

On peut **reconstruire** les chocs si l'on se donne une valeur de β et une valeur initiale :

$$\varepsilon_0 = 0, \qquad \varepsilon_t = y_t - \theta - \beta\,\varepsilon_{t-1}.$$

On obtient alors la somme des carrés $S(\beta) = \sum_t \varepsilon_t^2$, que l'on minimise **numériquement** sur β. Le problème n'est plus linéaire : il n'y a pas de formule fermée.

Moindres carrés conditionnels

La méthode

Fixer les chocs initiaux à zéro, reconstruire les ε_t par récursion, minimiser $S(\alpha, \beta)$ par un algorithme d'optimisation.

Son défaut

Le résultat dépend de l'initialisation arbitraire $\varepsilon_0 = 0$. L'effet s'estompe quand T grandit — d'autant plus vite que le modèle est inversible — mais il reste sensible sur petits échantillons.

Le maximum de vraisemblance

Vraisemblance

On se donne une loi pour les innovations, en général $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, et l'on cherche les paramètres qui rendent l'échantillon observé **le plus probable** :

$$\hat{\vartheta} = \arg\max_{\vartheta} \; L(\vartheta \mid x_1, \dots, x_T).$$

En décomposant la loi jointe en produit de lois conditionnelles :

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln v_t - \frac{1}{2}\sum_{t=1}^{T}\frac{(x_t - \widehat{x}_{t|t-1})^2}{v_t},$$

où $\widehat{x}_{t|t-1}$ est la prévision à un pas et v_t sa variance.

Comment la lire

La vraisemblance récompense les paramètres qui **prévoient bien à un pas**. Elle est calculée en pratique par le filtre de Kalman, qui produit les $\widehat{x}_{t|t-1}$ et les v_t de proche en proche.

Exacte ou conditionnelle

Deux variantes

- **Exacte** : traite correctement les premières observations, dont la loi diffère. Plus coûteuse, meilleure sur petits échantillons.
- **Conditionnelle** : conditionne aux premières valeurs, comme les moindres carrés conditionnels. Plus rapide.

Le choix, en pratique

Avec T de l'ordre de plusieurs milliers — le cas des données financières quotidiennes — les deux coïncident à la troisième décimale. La question ne se pose vraiment que sur données annuelles.

Ce que l'on obtient

Propriétés asymptotiques

Sous stationnarité, inversibilité et bonne spécification :

$$\sqrt{T}(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(0,\ \mathcal{I}(\vartheta)^{-1}),$$

où $\mathcal{I}$ est la matrice d'information de Fisher.

L'estimateur est convergent, asymptotiquement normal, et **efficace** : aucun estimateur sans biais ne fait mieux asymptotiquement.

Un point rassurant sur la normalité

La quasi-vraisemblance

On suppose des innovations gaussiennes pour écrire L. Si elles ne le sont pas — et les rentabilités ne le sont jamais — l'estimateur reste **convergent**. On parle de quasi-maximum de vraisemblance.

Seuls les écarts-types demandent une correction, dite « robuste ». C'est ce qui rend la méthode utilisable sur des données à queues épaisses.

Comparaison des méthodes

Quelle méthode pour quel modèle

Méthode	AR pur	MA / ARMA	Coût	Propriétés
MCO	direct	impossible	faible	convergent
Yule-Walker	direct	non	très faible	convergent, moins précis
MC conditionnels	oui	oui	moyen	dépend de l'initialisation
MV conditionnelle	oui	oui	moyen	très bonne
MV exacte	oui	oui	élevé	optimale

Ce que font les logiciels

EViews, R (`arima`) et Python (`statsmodels`) emploient par défaut le **maximum de vraisemblance exact** via le filtre de Kalman. Les autres méthodes servent surtout à fournir les valeurs initiales de l'optimisation.

Deux difficultés numériques

Ce qui peut mal se passer

- **Non-convergence.** Souvent le signe d'une racine commune (chapitre 7) ou d'un modèle trop grand. Remède : réduire les ordres.
- **Racine sur le cercle unité.** Un $\widehat{\beta}$ estimé à $0{,}999$ indique une série sur-différenciée. Remède : différencier une fois de moins.

Synthèse

Ce qu'il faut retenir

L'essentiel du chapitre 14

- Un AR s'estime par MCO : régression sur des valeurs observées, estimateur convergent.
- Un MA ne le peut pas : ses régresseurs sont des chocs inobservables.
- On reconstruit les chocs par récursion et l'on minimise **numériquement** — pas de formule fermée.
- Le **maximum de vraisemblance**, calculé par le filtre de Kalman, est la méthode de référence.
- L'estimateur est convergent même si les innovations ne sont pas gaussiennes : c'est la quasi-vraisemblance.
- Non-convergence ou racine proche de 1 : le modèle est mal spécifié, pas l'algorithme.

Vers le chapitre 15

Reste l'étape décisive. Un modèle estimé n'est pas un modèle valide : il faut interroger ses résidus. Et c'est là que les rentabilités financières vont révéler ce qu'un ARMA ne peut pas capturer.