

Modélisation GARCH de la volatilité

Chapitre 14 : L'estimation par quasi-maximum
de vraisemblance



Jaouad Madkour

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

Avant de commencer

La vraisemblance

Le quasi-maximum de vraisemblance

Estimer avec une loi non gaussienne

Les difficultés numériques

Les propriétés asymptotiques

Synthèse

Avant de commencer

Le chapitre 9 a fermé cette voie

Le GARCH est un ARMA sur les carrés — on pourrait donc l'estimer par MCO. Mais l'innovation $v_t = \varepsilon_t^2 - h_t$ est fortement hétéroscédastique, très asymétrique, et rien ne garantirait la positivité des coefficients.

La méthode de référence est le **maximum de vraisemblance**.

La vraisemblance

Décomposition en produit de lois conditionnelles

$$L(\vartheta) = \prod_{t=1}^T f(r_t \mid \Omega_{t-1}; \vartheta).$$

Sous $z_t \sim \mathcal{N}(0, 1)$, la loi de r_t sachant le passé est $\mathcal{N}(\mu, h_t)$, d'où :

$$\ln L(\vartheta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \left[\ln h_t + \frac{(r_t - \mu)^2}{h_t} \right].$$

Comment lire cette expression

Deux termes qui s'opposent. $\ln h_t$ **pénalise** les variances élevées : le modèle ne peut pas gonfler h_t partout pour se mettre à l'abri. Et ε_t^2/h_t **pénalise** les variances trop faibles là où le choc a été grand.

L'optimum réalise l'équilibre : h_t doit être grande exactement les jours où ε_t^2 l'est.

Le point technique central

h_t n'est pas un paramètre : c'est une **fonction récursive** des paramètres et des données passées.

$$h_t(\vartheta) = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}(\vartheta).$$

La conséquence pratique

Chaque évaluation de la vraisemblance exige de parcourir toute la série, de $t = 1$ à $t = T$. Il n'existe **aucune** formule fermée : l'optimisation est nécessairement numérique.

C'est pourquoi l'estimation d'un GARCH est instantanée sur 2 500 points, et sensiblement plus lente sur un million.

Il faut une valeur pour h_1

Trois conventions courantes :

- $h_1 = \hat{\sigma}^2$, la variance empirique de l'échantillon — la plus employée;
- $h_1 = \omega / (1 - \alpha_1 - \beta_1)$, la variance de long terme du modèle;
- h_1 traitée comme un paramètre supplémentaire à estimer.

L'effet, et il est mesurable

L'influence de h_1 décroît en β_1^t . Avec $\beta_1 = 0,89$, elle est divisée par 100 au bout de 40 périodes.

Sur 2 500 observations, le choix est sans conséquence. Sur 200, il change la troisième décimale — et c'est pourquoi les logiciels ne donnent pas tous exactement les mêmes chiffres.

Le quasi-maximum de vraisemblance

Une incohérence apparente

Le chapitre 10 a établi que z_t n'est pas gaussienne. Or nous venons d'écrire la vraisemblance **sous hypothèse gaussienne**.

Est-ce que cela invalide l'estimation ?

Bollerslev et Wooldridge (1992)

Si l'équation de variance est **correctement spécifiée**, l'estimateur du maximum de vraisemblance **gaussien** est convergent et asymptotiquement normal, **même si la vraie loi de z_t n'est pas normale**.

On l'appelle estimateur du **quasi-maximum de vraisemblance** (QMV).

Pourquoi cela fonctionne

La vraisemblance gaussienne ne fait intervenir que les deux premiers moments conditionnels. Tant que $E(\varepsilon_t \mid \Omega_{t-1}) = 0$ et $V(\varepsilon_t \mid \Omega_{t-1}) = h_t$ sont exacts, les équations de score sont satisfaites en moyenne — quelle que soit la forme de la loi.

Écart-types robustes (sandwich)

$$\widehat{V}(\hat{\vartheta}) = \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1},$$

où $\mathbf{A}$ est le hessien et $\mathbf{B}$ la matrice des produits de scores. Sous normalité, $\mathbf{A} = \mathbf{B}$ et l'expression se simplifie; sinon, non.

Un réglage à ne jamais négliger

Les écart-types non robustes sont **sous-estimés** — souvent de 20 à 40 % sur données financières. On conclurait à tort à la significativité.

En R : `ugarchfit(..., solver.control = list(...))` fournit les deux jeux, sous les intitulés « Std.Error » et « Robust Std.Error ». **Toujours lire la seconde colonne.**

Estimer avec une loi non
gaussienne

Vraisemblance de Student

$$\ln L = \sum_{t=1}^T \left[\ln \Gamma\left(\frac{\nu+1}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2} \ln (\pi(\nu-2)) - \frac{1}{2} \ln h_t - \frac{\nu+1}{2} \ln \left(1 + \frac{\varepsilon_t^2}{h_t(\nu-2)}\right) \right]$$

Le paramètre ν est estimé **conjointement** avec les autres.

Le compromis

Si la loi est bien spécifiée, l'estimateur est **efficace** : il fait mieux que le QMV gaussien.

Si elle est mal spécifiée, la propriété de convergence est **perdue** — contrairement au QMV gaussien, qui reste robuste.

En pratique : la Student est presque toujours meilleure que la normale sur données financières, donc le risque de mauvaise spécification est faible, et on l'emploie.

Les difficultés numériques

Quatre symptômes et leurs causes

Symptôme	Cause probable	Remède
non-convergence	mauvaises valeurs initiales	partir d'un GARCH(1,1) simple
$\hat{\alpha}_1$ ou $\hat{\beta}_1$ collé à zéro	modèle sur-paramétré	réduire p et q
$\hat{\pi}$ estimé à 0,9999	rupture de régime	découper l'échantillon
écarts-types absurdes	quasi-non-identification	simplifier le modèle

Travailler en pourcentage

Si les rentabilités sont exprimées en décimal — 0,012 pour 1,2 % — alors ω est de l'ordre de 10^{-6} . L'optimiseur travaille alors sur des échelles très différentes selon les paramètres, et converge mal.

Multiplier les rentabilités par 100 avant estimation. Les logiciels le font parfois automatiquement, et affichent un avertissement quand ils ne le font pas — il faut le lire.

Vérifier la contrainte de stationnarité après coup

Un optimiseur non contraint peut renvoyer $\hat{\alpha}_1 + \hat{\beta}_1 > 1$. Ce n'est pas nécessairement une erreur de calcul : c'est le signe que le modèle n'est pas adapté, ou que l'échantillon contient une rupture.

Les propriétés asymptotiques

Convergence et normalité asymptotique

Sous stationnarité stricte, identification et existence de moments suffisants :

$$\sqrt{T}(\hat{\vartheta} - \vartheta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}).$$

Une réserve à connaître

Les conditions de régularité exigent l'existence de moments d'ordre élevé de ε_t — que le chapitre 8 a montrés souvent absents en finance.

La théorie asymptotique du GARCH repose donc sur des hypothèses plus fragiles qu'il n'y paraît. C'est une raison de plus d'employer les écarts-types robustes, et de se méfier de la troisième décimale.

Synthèse

L'essentiel du chapitre 14

- La vraisemblance équilibre deux pénalités : $\ln h_t$ contre ε_t^2/h_t .
- h_t est une **fonction réursive** des paramètres : pas de formule fermée, optimisation numérique obligatoire.
- **QMV** : le maximum de vraisemblance gaussien reste convergent même si z_t n'est pas gaussienne.
- Il faut alors des **écarts-types robustes** (sandwich) : les autres sont sous-estimés de 20 à 40 %.
- La vraisemblance de Student est plus efficace, et presque toujours préférable en finance.
- Précautions : rentabilités en pourcentage, bonnes valeurs initiales, vérification de $\hat{\pi} < 1$.

Vers le chapitre 15

Un modèle estimé n'est pas un modèle valide. Il faut interroger ses résidus standardisés — et le critère sera exactement celui du cours ARMA, transposé aux carrés.