Librairie RegArch

Modèles

Ollivier TARAMASCO

Table des matières

Objectifs de la librairie

La librairie RegArch est conçue pour simuler/estimer des modèles univariés avec erreurs de type ARCH. Elle est écrite en C++ et utilise les bibliothèques GNU GSL et NLOPT. Pour les interfaces Excel et R, se rajoutent des librairies écrites en C++/CLI, C# et R script. Il y a au total, plus de 25 000 lignes de code !

Modèle général

Le modèle de série chronologique simulé/estimé dans cette librairie est le suivant :

$$Y(t)|\,\underline{I_t} = m(t,\theta) + \sigma(t,\theta)\varepsilon(t,\beta) = m(t,\theta) + u(t,\theta)$$

où

- θ est le vecteur de l'ensemble des paramètres du modèle, β est l'éventuel vecteur des paramètres de la distribution de $\varepsilon(t,\beta)$,
- $\underline{I_t}$ est « l'information disponible » à la date t,
- $m(t,\theta)$ est l'espérance conditionnelle,
- $\sigma^2(t,\theta) = h(t,\theta)$ est la variance conditionnelle,
- $\varepsilon(t,\beta)$ est un bruit blanc, de variance égale à 1,
- $u(t,\theta) = \sigma(t,\theta)\varepsilon(t,\beta)$ est un processus de type *ARCH* « pur » (espérance nulle).

Les $\varepsilon(t,\beta)$ sont I.I.D. mais pas les $u(t,\theta)$ ou les $Y(t)$.

Pour que l'utilisateur ait le plus grand nombre possible de choix de modèles, on permet à l'espérance conditionnelle de s'écrire comme une somme « d'espérances conditionnelles élémentaires » :

$$m(t,\theta) = \sum_{i=1}^{N_m} m_i(t,\theta)$$

La librairie permet de choisir différentes « types » d'espérances conditionnelles « élémentaires », comme une constante, un modèle autorégressif d'ordre p, un modèle moyenne mobile d'ordre q, *etc*.

Pour la variance conditionnelle, le choix est donné entre divers modèles de type *ARCH*, comme *ARCH(p), GARCH(p, q), EGARCH(p, q), etc.*

De même, on peut choisir la distribution conditionnelle parmi plusieurs famille de loi, comme la loi normale, la loi de Student *etc.*

Les différents types d'espérances conditionnelles

Voici la liste des choix possibles pour les $m_i(t,\theta)$:

- ***Pas d'espérance conditionnelle***. Ce choix exclut tous les autres et on a donc $Y(t)|\underline{I_t} = u(t,\theta)$ = $ARCH$ « pur »
- ***Constante*** : $m_i(t,\theta) = CST$ et θ contient *CST.*
- ***AR(p)*** : $m_i(t,\theta) = \sum_{i=1}^{p} AR(i)Y(t-i)$ et θ contient le vecteur *AR(1), …, AR(p)*
- ***MA(q)*** : $m_i(t,\theta) = \sum_{j=1}^{q} MA(j)u(t-i)$ et θ contient le vecteur *MA(1), …, MA(q)*
- ***Régression linéaire*** : $m_i(t,\theta) = X(t)\beta$, où β est un vecteur de dimension *p > 0* et *X(t)* un vecteur ligne de régresseurs de dimension p. θ contient β.
- ***Variance in mean*** : $m_i(t,\theta) = \delta\sigma^2(t,\theta)$ et θ contient δ.
- ***Standard deviation in mean*** : : $m_i(t,\theta) = \delta\sigma(t,\theta)$ et θ contient δ.
- ***ARFIMA(p, d, q)*** : modèle *ARMA* fractal (voir par exemple cette page web). θ contient le vecteur *AR(1), …, AR(p),* le vecteur *MA(1), …, MA(q)* et le nombre (fractal) *d.*

Si vous voulez modéliser l'espérance conditionnelle par un modèle *ARMA(p, q)* avec une constante, il faudra ajouter le type *Constante,* le type *AR(p)* et le type *MA(q).*

Attention, la librairie ne vérifie pas si le modèle est bien « écrit » et estimable. Par exemple, éviter d'ajouter un modèle *AR(p)* à un modèle *ARFIM(p, d, q)* !

Les différents types de variances conditionnelles

Voici la liste des choix possibles pour les $\sigma^2(t,\theta) = h(t,\theta)$:

- ***Constante :*** $h(t,\theta) = CSTVAR$ et θ contient *CSTVAR,*
- ***ARCH(p)*** : $h(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH(i)u(t-i,\theta)^2$ et θ contient *CSTVAR* et le vecteur *ARCH(1), …, ARCH(p)*

- ***GARCH(p, q)*** : $h(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH(i)u(t-i,\theta)^2 + \sum_{j=1}^{q} GARCH(j)h(t-j,\theta)$ et θ contient *CSTVAR* , le vecteur *ARCH(1), ..., ARCH(p)* et le vecteur *GARCH(1), ..., GARCH(q),*
- ***TARCH(p) :*** $h(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH_{+}(i)u(t-i,\theta)^2\delta_{+}(u(t-i,\theta)) + \sum_{i=1}^{p} ARCH_{-}(i)u(t-i,\theta)^2\delta_{-}(u(t-i,\theta))$, où $\delta_{+}(x)$ (respectivement $\delta_{-}(x)$) est la fonction qui vaut 1 si $x > 0$ et 0 sinon (respectivement la fonction qui vaut 1 si $x < 0$ et 0 sinon). θ contient *CSTVAR* et les vecteurs *ARCH+(1), ..., ARCH+(p)* et *ARCH-(1), ..., ARCH- (p),*
- ***NGARCH(p, q)*** : $h(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH(i)[u(t-i,\theta)^2 - \varphi h(t-i,\theta)] + \sum_{j=1}^{q} GARCH(j)h(t-j,\theta)$ et θ contient *CSTVAR* , le nombre φ, le vecteur *ARCH(1), ..., ARCH(p)* et le vecteur *GARCH(1), ..., GARCH(q),*
- ***EGARCH(p, q)*** : $\operatorname{Log} h(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH(i)[\varphi\varepsilon(t-i,\beta) + \gamma\{|\varepsilon(t-i,\beta)| - E(|\varepsilon(t,\beta)|)\}] + \sum_{j=1}^{q} GARCH(j)\operatorname{Log} h(t-j,\theta)$, et θ contient *CSTVAR* , le nombre φ, le nombre γ, le vecteur *ARCH(1), ..., ARCH(p)* et le vecteur *GARCH(1), ..., GARCH(q)* et (éventuellement) le vecteur β*,*
- ***APARCH(p, q)*** : $\sigma^{\delta}(t,\theta) = CSTVAR + \sum_{i=1}^{p} ARCH(i)[|u(t-i,\theta)| - \gamma(i)u(t-i,\theta)]^{\delta} + \sum_{j=1}^{q} GARCH(j)\sigma^{\delta}(t-j,\theta)$ et θ contient *CSTVAR* , le nombre δ, le vecteur *ARCH(1), ..., ARCH(p),* le vecteur $\gamma(1),\ldots, \gamma(p)$ et le vecteur *GARCH(1), ..., GARCH(q),*
- ***UGARCH(p, q)*** : $h(t,\theta) = [CSTVAR] + X(t)\beta + \sum_{i=1}^{p} ARCH(i)u(t-i,\theta)^2 + \sum_{j=1}^{q} GARCH(j)h(t-j,\theta)$ et θ contient le vecteur β , le vecteur *ARCH(1), ..., ARCH(p)* et le vecteur *GARCH(1), ..., GARCH(q),* et éventuellement le nombre *CSTVAR.* Ici, *X(t)* un vecteur ligne de régresseurs de même dimension que β
- ***FIGARCH(p, d, q)*** : modèle *GARCH* fractal (voir par exemple cette page web). θ contient *CSTVAR* , le vecteur *ARCH(1), ..., ARCH(p)* et le vecteur *GARCH(1), ..., GARCH(q),* et le nombre fractal *d.*

Les différents types de distributions conditionnelles

Voici la liste des différents choix possibles pour la distribution de $\varepsilon(t,\beta)$

- ***Loi normale*** : pas de paramètre de distribution,
- ***Loi de Student*** : β est le nombre de degrés de liberté,
- ***GED*** : β est le paramètre : β de la *General Error Distribution.*
- **Mélange de deux lois normales** : β est le vecteur à trois paramètres contenant le paramètre p de la loi de Bernouilli, et les deux variances σ_X^2 et σ_Y^2