Régressions avec résidus ARCH

Jean-Baptiste DURAND   Ollivier TARAMASCO

18 avril 2013

Modèle
Notations

- $\theta \in \mathbb{R}^n$ est l'ensemble des paramètres du modèle.
- $\beta \in \mathbb{R}^q$ est l'ensemble des paramètres de la distribution conditionnelle.
- $\phi \in \mathbb{R}^p$ est l'ensemble des paramètres de l'espérance et/ou de la variance conditionnelle.

On a donc

$$\begin{aligned} \theta &= \begin{pmatrix} \phi \\ \beta \end{pmatrix} \\ n &= p + q \end{aligned}$$

Modèle
Processus

- $Y(t) \,|\, \underline{\mathcal{I}_{t-1}} = m(t\,;\,\theta) + \sigma(t\,;\,\theta)\,\varepsilon(t\,;\,\beta) = m(t\,;\,\theta) + U(t\,;\,\theta)$
- $m(t\,;\,\theta)$: espérance conditionnelle.
- $\sigma(t\,;\,\theta)^2 = h(t\,;\,\theta)$: variance conditionnelle.
- $U(t\,;\,\theta)$: résidu (processus de type ARCH “pur”).
- $\varepsilon(t\,;\,\beta)$: bruit blanc, variance 1.

Remarque : Les $\varepsilon(t\,;\,\beta)$ sont i.i.d. Pas les $U(t\,;\,\theta)$ ni les $Y(t)$.

Vraisemblance
Calcul

On observe une trajectoire $\{y(t), t = 1 \ldots T\}$ du processus Y.

La vraisemblance s'écrit :

$$\begin{aligned}
\ln L(\theta, y) &= \sum_{t=t_0}^{T} l(t\,;\,\theta, y) \\
l(t\,;\,\theta, y) &= -\ln\sigma(t\,;\,\theta) + \ln f_\varepsilon\left(\frac{y(t) - m(t\,;\,\theta)}{\sigma(t\,;\,\theta)}\,;\,\beta\right) \\
&= -\ln\sigma(t\,;\,\theta) + \ln f_\varepsilon\left(\frac{u(t\,;\,\theta)}{\sigma(t\,;\,\theta)}\,;\,\beta\right) \\
&= -\ln\sigma(t\,;\,\theta) + \ln f_\varepsilon\left(e(t\,;\,\theta)\,;\,\beta\right)
\end{aligned}$$

où $f_\varepsilon(x\,;\,\beta)$ est la densité de la variable aléatoire $\varepsilon(t\,;\,\beta)$ (qui dépend des paramètres β)

Vraisemblance
Densité du bruit

Si $\varepsilon\,(t\,;\,\beta)$ est la variable centrée réduite d'un variable $\nu(t)$ de loi connue de paramètre β :

$$\varepsilon\,(t\,;\,\beta) \quad = \quad \frac{\nu(t) - m_\nu(\beta)}{\sigma_\nu(\beta)}$$

avec $m_\nu(\beta) = \mathrm{E}\,[\nu(t)]$ et $\sigma^2_\nu(\beta) = \mathrm{var}\,[\nu(t)]$, on a :

$$\ln f_\varepsilon(x\,;\,\beta) \quad = \quad \ln\sigma_\nu(\beta) + \ln f_\nu(\sigma_\nu(\beta)\,x + \mu_\nu(\beta)\,;\,\beta)$$

par exemple, si $\nu(t) \rightsquigarrow \mathcal{S}t(n)$, alors $\beta = n$, $m_\nu(n) = 0$, $\sigma_\nu(n) = \sqrt{\frac{n}{n-2}}$ et donc :

$$\ln f_\varepsilon(x\,;\,n) \quad = \quad \frac{1}{2}\ln\frac{n}{n-2} + \ln f_{\mathcal{S}t(n)}\left(\sqrt{\frac{n}{n-2}}\,x\right)$$

Gradient

Calcul

$$
\begin{aligned}
\operatorname{grad}_\theta \ln L\,(t\,;\,\theta, y) &= \sum_{t=t_0}^{T} \operatorname{grad}_\theta l\,(t\,;\,\theta, y) \quad \text{(Eq 1)} \\
\operatorname{grad}_\phi l\,(t\,;\,\theta, y) &= -\frac{1}{\sigma\,(t\,;\,\theta)} \operatorname{grad}_\phi \sigma\,(t\,;\,\theta) \\
&\quad + \frac{\partial \ln f_\varepsilon}{\partial x}\,(e\,(t\,;\,\theta)\,,\beta)\, \operatorname{grad}_\phi e\,(t\,;\,\theta) \quad \text{(Eq 2)}
\end{aligned}
$$

Gradient

Calcul suite ...

$$\begin{aligned}
\operatorname{grad}_\beta l(t\,;\,\theta, y) &= -\frac{1}{\sigma(t\,;\,\theta)} \operatorname{grad}_\beta \sigma(t\,;\,\theta) \\
&\quad + \frac{\partial \ln f_\varepsilon}{\partial x}\left(e(t\,;\,\theta), \beta\right) \operatorname{grad}_\beta e(t\,;\,\theta) \\
&\quad + \operatorname{grad}_\beta \ln f_\varepsilon\left(e(t\,;\,\theta), \beta\right) \qquad (\text{Eq } 3) \\
\operatorname{grad}_\theta e(t\,;\,\theta) &= \frac{1}{\sigma(t\,;\,\theta)} \left[\operatorname{grad}_\theta u(t\,;\,\theta)\right. \\
&\quad \left. - e(t\,;\,\theta) \operatorname{grad}_\theta \sigma(t\,;\,\theta)\right] \qquad (\text{Eq } 4) \\
\operatorname{grad}_\theta u(t\,;\,\theta) &= -\operatorname{grad}_\theta m(t\,;\,\theta) \qquad (\text{Eq } 5)
\end{aligned}$$

Gradient

Simplifications

Parfois le calcul est plus simple si on dérive $h(t\,;\,\theta) = \sigma(t\,;\,\theta)^2$ (ex. GARCH) ou $\nu(t\,;\,\theta) = \ln\sigma(t\,;\,\theta)$ (ex. EGARCH) on a :

$$\begin{aligned}
\mathrm{grad}_\theta\,\sigma(t\,;\,\theta) &= \frac{1}{2\sqrt{h(t\,;\,\theta)}}\,\mathrm{grad}_\theta\,h(t\,;\,\theta)\\
&= \frac{1}{2\,\sigma(t\,;\,\theta)}\,\mathrm{grad}_\theta\,h(t\,;\,\theta)\\
\mathrm{grad}_\theta\,\sigma(t\,;\,\theta) &= e^{\nu(t\,;\,\theta)}\,\mathrm{grad}_\theta\,\nu(t\,;\,\theta)\\
&= \sigma(t\,;\,\theta)\,\mathrm{grad}_\theta\,\nu(t\,;\,\theta)
\end{aligned}$$

Gradient
Algorithme

- Écrire le gradient pour chaque type d'espérance, variance et log-densité conditionnelle.
- Utiliser l'équation (Eq 5) pour en déduire $\mathrm{grad}_\theta\, u\,(t\,;\,\theta)$.
- Puis l'équation (Eq 4) pour en déduire $\mathrm{grad}_\theta\, e\,(t\,;\,\theta)$
- Puis les équations (Eq 2) et (Eq 3) pour obtenir $\mathrm{grad}_\theta\, l\,(t\,;\,\theta, y)$
- Sommer toutes ces valeurs pour calculer $\mathrm{grad}_\theta \ln L\,(t\,;\,\theta, y)$.

Récurrence

- on observe les $y(t)$,
- on connaît la valeur du vecteur θ des paramètres du modèle,
- on calcule $m(t\,;\,\theta)$, $\sigma(t\,;\,\theta)$ puis $u(t\,;\,\theta)$ et $e(t\,;\,\theta)$ à chaque date t.

Ces valeurs dépendent non seulement de l'observation $y(t)$ à la date t **mais aussi** de ce qui a précédé (donc les $m(s\,;\,\theta)$, $\sigma(s\,;\,\theta)$, $u(s\,;\,\theta)$, $e(s\,;\,\theta)$, $s < t$).

Donc le gradient de chacune de ces quantités va aussi dépendre des “gradients du passé”

Récurrence
Exemple

Soit un modèle MA(1) “pur” gaussien. Il y a deux paramètres MA(1) et la variance σ^2 du bruit blanc gaussien :

$$\theta = \begin{pmatrix} \mathrm{MA}(1) \\ \sigma^2 \end{pmatrix}$$

Le modèle s'écrit :

$$\begin{aligned} Y(t) &= \mathrm{MA}(1)\, U(t-1\,;\,\theta) + U\,(t\,;\,\theta) \\ &= \mathrm{MA}(1)\, U(t-1,;\,\theta) + \sigma\,\varepsilon(t) \end{aligned}$$

et donc

$$m\,(t\,;\,\theta) = \mathrm{MA}(1)\, u(t-1\,;\,\theta)$$

Récurrence
Exemple suite

Par conséquent :

$$\begin{aligned}
\operatorname{grad}_\theta m(t\,;\,\theta) &= \begin{pmatrix} u(t-1\,;\,\theta) \\ 0 \end{pmatrix} + \mathrm{MA}(1)\operatorname{grad}_\theta u(t-1\,;\,\theta) \\
&= \begin{pmatrix} u(t-1\,;\,\theta) \\ 0 \end{pmatrix} - \mathrm{MA}(1)\operatorname{grad}_\theta m(t-1\,;\,\theta)
\end{aligned}$$

Programmation

Vecteur des paramètres

Dans la représentation informatique du modèle, les diverses valeurs des paramètres se retrouvent stockées dans divers classes C++. Pour pouvoir faire les calculs, il faut “transférer” ces diverses valeurs dans un seul “cDVector” θ.

Par exemple, considérons le modèle Cstmean + ARMA(1, 1) - GARCH(1, 1) avec résidus $\mathcal{S}t(n)$. Ce modèle comporte 7 paramètres :

- paramètres de l'espérance conditionnelle : $\mathrm{CstMean}$, $\mathrm{AR}(1)$, $\mathrm{MA}(1)$ dispersés dans les 3 classes cConst, cAr et cMa.
- paramètres de la variance conditionnelle : CstVar, $\mathrm{ARCH}(1)$, $\mathrm{GARCH}(1)$ dans la classe cGarch
- paramètre de la distribution conditionnelle : n dans la classe cStudentResiduals

Programmation
Récupération

Pour récupérer les paramètres, il faut :

1. Calculer le nombre total de paramètres pour initialiser la taille du vecteur θ.
2. Récupérer la valeur de chaque paramètre dans chaque classe et les placer au bon endroit dans le vecteur θ.

En conséquence de quoi, il faut que dans la classe cRegArchParam, il y ait :

1. une fonction *int GetNParam(void)* qui retourne le nombre total de paramètres du modèle
2. une fonction *void RegArchParamToVector(cDVector& theDestVect)* qui récupère les valeurs et les mette dans theDestVect

Pour le calcul de la vraisemblance, on a utilisé une classe cRegArchValue qui permet de stocker les calculs à chaque date t. On aurait pu se permettre de ne garder que les “nLagMax” précédentes valeurs, où “nLagmax” est le nombre maximal de retards dans modèle.

A cause de la récurrence, il va falloir faire de même lors du calcul du gradient. Il faut donc créer une classe cRegArchGradient qui contiennent (entre autre) les “nGradLagMax” gradients du passés.

Il faut donc que

- dans la classe cRegArchParam, il y ait une fonction *uint GetNLags(void)* qui retourne “nGradLagMax”
- que dans la classe cRegArchGradient, après le calcul à la date t des divers gradients, il y ait une mise à jour pour préparer le calcul à la date $t+1$

Hessien
Calcul

$$
\begin{aligned}
\mathrm{hess}_{\theta\theta}\ \mathrm{ln}L\,(t\,;\,\theta, y) &= \sum_{t=t_0}^{T} \mathrm{hess}_{\theta\theta}\ l\,(t\,;\,\theta, y) \quad \text{(Eq 4)} \\
\mathrm{hess}_{\phi\phi}\ l\,(t\,;\,\theta, y) &= \frac{1}{\sigma^2()}\mathrm{grad}_{\phi}\ \sigma()\ {}^t\mathrm{grad}_{\phi}\ \sigma() - \frac{1}{\sigma()}\mathrm{hess}_{\phi\phi}\ \sigma() \\
&+ \frac{\partial^2 \mathrm{ln} f_{\varepsilon}}{\partial x^2}\,(e(), \beta)\ \mathrm{grad}_{\phi}\ e()\,{}^t\mathrm{grad}_{\phi}\ e() \\
&+ \frac{\partial \mathrm{ln} f_{\varepsilon}}{\partial x}\,(e(), \beta)\ \mathrm{hess}_{\phi\phi}\ e() \quad \text{(Eq 5)}
\end{aligned}
$$

Hessien

Calcul suite 1

$$
\begin{aligned}
\mathrm{hess}_{\phi\beta}\, l\,(t\,;\,\theta, y) \;=\;& \frac{1}{\sigma^2()}\mathrm{grad}_\phi\,\sigma()\,{}^t\mathrm{grad}_\beta\,\sigma() - \frac{1}{\sigma()}\mathrm{hess}_{\phi\beta}\,\sigma() \\
&+\frac{\partial^2 \ln f_\varepsilon}{\partial x^2}\,(e(),\beta)\;\mathrm{grad}_\phi\,e()\,{}^t\mathrm{grad}_\beta\,e() \\
&+\frac{\partial \ln f_\varepsilon}{\partial x}\,(e(),\beta)\;\mathrm{hess}_{\phi\beta}\,e() \\
&+\mathrm{grad}_\beta\,\frac{\partial \ln f_\varepsilon}{\partial x}\,(e(),\beta)\;{}^t\mathrm{grad}_\phi\,e() \qquad \text{(Eq 6)}
\end{aligned}
$$

Hessien

Calcul suite 2

$$
\begin{aligned}
\mathrm{hess}_{\beta\beta}\, l(t\,;\,\theta, y) \quad = \quad & \frac{1}{\sigma^2()}\mathrm{grad}_\beta\, \sigma()\, {}^t\mathrm{grad}_\beta\, \sigma() - \frac{1}{\sigma()}\mathrm{hess}_{\beta\beta}\, \sigma() \\
& + \frac{\partial^2 \ln f_\varepsilon}{\partial x^2}\,(e(), \beta)\ \mathrm{grad}_\beta\, e()^t\mathrm{grad}_\beta\, e() \\
& + \frac{\partial \ln f_\varepsilon}{\partial x}\,(e(), \beta)\ \mathrm{hess}_{\beta\beta}\, e() \\
& + \mathrm{grad}_\beta\, \frac{\partial \ln f_\varepsilon}{\partial x}\,(e(), \beta)\ {}^t\mathrm{grad}_\beta\, e() \\
& + \mathrm{grad}_\beta\, e()\, {}^t\mathrm{grad}_\beta\, \frac{\partial \ln f_\varepsilon}{\partial x}\,(e(), \beta) \\
& + \mathrm{hess}_{\beta\beta}\, \ln f_\varepsilon\,(e(), \beta) \qquad \text{(Eq 7)}
\end{aligned}
$$

Hessien

Calcul suite 2

$$
\begin{aligned}
\mathrm{hess}_{\theta\theta}\, e\,(t\,;\,\theta) &= -\frac{1}{\sigma^2()}\mathrm{grad}_\theta\, \sigma()\, {}^t\mathrm{grad}_\theta\, u() + \frac{1}{\sigma()}\mathrm{hess}_{\theta\theta}\, u() \\
&\quad + \frac{2\,u()}{\sigma^3()}\mathrm{grad}_\theta\, \sigma()\, {}^t\mathrm{grad}_\theta\, \sigma() \\
&\quad - \frac{1}{\sigma^2()}\mathrm{grad}_\theta\, u()\, {}^t\mathrm{grad}_\theta\, \sigma() \\
&\quad - \frac{u()}{\sigma^2()}\mathrm{hess}_{\theta\theta}\, \sigma() \quad \text{(Eq 8.)} \\
\mathrm{hess}_{\theta\theta}\, u() &= -\mathrm{hess}_{\theta\theta}\, m() \quad \text{(Eq 9.)}
\end{aligned}
$$