2017/12/26

新谷元嗣、藪友良

補論 3.2 多変量 GARCH モデル

対数尤度関数

3 章 7 節では、1 変量の対数尤度を求めた。ここでは多変量の場合、とくに 2 変量について対数尤度を求める。誤差項ε_{1t}、ε_{2t}は平均 0 で、2 次元の正規分布に従うとする。単純化のため、分散と共分散は時間を通じて一定としよう（この仮定は後で変更される）。したがって、h_{ijt}から添え字 t を除くことができる。このとき、ε_{1t}とε_{2t} の尤度関数は、2 次元正規分布の定義から、

$$L_t = \frac{1}{2\pi\sqrt{h_{11}h_{22}(1-\rho_{12}^2)}}\exp\left[-\frac{1}{2(1-\rho_{12}^2)}\left(\frac{\varepsilon_{1t}^2}{h_{11}}+\frac{\varepsilon_{2t}^2}{h_{22}}-\frac{2\rho_{12}\varepsilon_{1t}\varepsilon_{2t}}{(h_{11}h_{22})^{0.5}}\right)\right] \quad \text{(A3.1)}$$

となる。ただし、ρ_{12}はε_{1t}とε_{2t}の相関係数とする（$\rho_{12} = h_{12}/(h_{11}h_{22})^{0.5}$）。

ここで行列 H を分散共分散行列

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$$

と定義すれば、尤度関数は、

$$L_t = \frac{1}{2\pi|H|^{1/2}}\exp\left[-\frac{1}{2}\varepsilon_t' H^{-1}\varepsilon_t\right] \quad \text{(A3.2)}$$

と表現できる。ただし、$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ であり、$|H|$は Hの行列式である。ここで、「$'$」は行列の転置を意味する。

(A3.1)式と(A3.2)式が同じであることを示す。まず、行列式は$|H|=h_{11}h_{22}-(h_{12})^2$ であり、$h_{12} = \rho_{12}(h_{11}h_{22})^{0.5}$ から$|H|= (1-(\rho_{12})^2)h_{11}h_{22}$ と書ける。次に、H の逆行列

$$H^{-1} = \frac{1}{h_{11}h_{22}-h_{12}^2}\begin{bmatrix} h_{22} & -h_{12} \\ -h_{12} & h_{11} \end{bmatrix}$$

を用いて、

$$\varepsilon_t' H^{-1}\varepsilon_t = \frac{\varepsilon_{1t}^2 h_{22} - 2\varepsilon_{1t}\varepsilon_{2t}h_{12} + \varepsilon_{2t}^2 h_{11}}{h_{11}h_{22}-h_{12}^2}$$

となる。さらに、これは $h_{12} = \rho_{12}(h_{11}h_{22})^{0.5}$ を用いると、

$$\varepsilon_t' H^{-1}\varepsilon_t = \left[\frac{1}{(1-\rho_{12}^2)}\left(\frac{\varepsilon_{1t}^2}{h_{11}}+\frac{\varepsilon_{2t}^2}{h_{22}}-\frac{2\rho_{12}\varepsilon_{1t}\varepsilon_{2t}}{(h_{11}h_{22})^{0.5}}\right)\right]$$

となる。以上を用いれば、(A3.1)(A3.2)式が等しいことを示せる。

ここで$\{\varepsilon_t\}$は相互に独立としよう。したがって、$(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_T)$が同時に得られる尤度は、各尤度の積となる。分散が一定なら、尤度は、

$$L = \prod_{t=1}^{T}\frac{1}{2\pi|H|^{1/2}}\exp\left[-\frac{1}{2}\varepsilon_t' H^{-1}\varepsilon_t\right]$$

となる。対数をとると、対数尤度

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln|H| - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t' H^{-1}\varepsilon_t$$

を得る。最尤法では、実現値が得られる尤度を最大にするパラメータが選ばれる。具体的には、ε_t の実現値が与えられたもとで、対数尤度を最大にする h_{11}, h_{12}, h_{22} が選択される。

GARCH 効果を分析したいため、条件付き分散 h_{ij} は一定ではなく、時変的としよう。3 章 8 節を読み終えた方は、h_{11}, h_{22}, h_{12} が時変的なとき、いかに修正すればいいかは明らかだろう。このとき、尤度関数は、

$$L = \prod_{t=1}^{T}\frac{1}{2\pi|H_t|^{1/2}}\exp\left[-\frac{1}{2}\varepsilon_t' H_t^{-1}\varepsilon_t\right]$$

となり、H_t は

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix}$$

となる。尤度関数の対数をとると、対数尤度を得る。

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}(\ln|H_t| + \varepsilon_t' H_t^{-1}\varepsilon_t) \qquad \text{(A3.3)}$$

これまで 2 変量の場合を考えてきた。しかし、(A3.2)(A3.3)式は k 変量でも成立する。ただし、H は k x k の対称行列、ε_t は k x 1 の列ベクトルとなる。また、定数項 (2π)は k 乗される。

多変量 GARCH の定式化

(A3.3)を推定するためには、h_{ijt} の定式化を決めなくてはならない。3 章 10 節では、vech モデル、対角 vech モデル、CCC モデルの基本的な考え方を紹介した。ここでは、行列を用いて、より厳密に、vech モデル、対角 vech モデル、BEKK モデル、CCC モデル、DCC モデルを紹介していく。

1. vech モデル、対角 vech モデル

***vech* オペレータ**とは、対称行列の下（上）三角の要素を取り出して列ベクトルにする演算方法である。たとえば、 対称行列を

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix}$$

とすると、vech オペレータを用いれば

$$vech(H_t) = [\ h_{11t}, h_{12t}, h_{22t}\ ]'$$

という列ベクトルになる。

ベクトル$\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}]'$を考えよう。$\varepsilon_t\varepsilon_t' = [\varepsilon_{1t}, \varepsilon_{2t}]'[\varepsilon_{1t}, \varepsilon_{2t}]$を計算すると、2 x 2 の行列

$$\begin{bmatrix} \varepsilon_{1t}^2 & \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{1t}\varepsilon_{2t} & \varepsilon_{1t}^2 \end{bmatrix}$$

となる。さらに vech オペレータを用いると、$vech(\varepsilon_t\varepsilon_t') = \left[\varepsilon_{1t}^2, \varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2t}^2\right]'$ となる。ここで、$C = [\,c_1, c_2, c_3\,]'$、A =3 x 3 の行列（各要素はα_{ij}）、B =3 x 3 の行列（各要素はβ_{ij}）とする。このとき、**vech** モデルは、

$$vech(H_t) = C + A\ vech(\varepsilon_{t-1}\varepsilon_{t-1}') + B\ vech(H_{t-1})$$

と表せる。行列を展開して、(3.40) – (3.42)式のシステムと同じであることを確認してもらいたい。

$$h_{11t} = c_{10} + \alpha_{11}\varepsilon_{1t-1}^2 + \alpha_{12}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{13}\varepsilon_{2t-1}^2 + \beta_{11}h_{11t-1} + \beta_{12}h_{12t-1} + \beta_{13}h_{22t-1} \tag{3.40}$$

$$h_{12t} = c_{20} + \alpha_{21}\varepsilon_{1t-1}^2 + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{23}\varepsilon_{2t-1}^2 + \beta_{21}h_{11t-1} + \beta_{22}h_{12t-1} + \beta_{23}h_{22t-1} \tag{3.41}$$

$$h_{22t} = c_{30} + \alpha_{31}\varepsilon_{1t-1}^2 + \alpha_{32}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{33}\varepsilon_{2t-1}^2 + \beta_{31}h_{1t-1} + \beta_{32}h_{12t-1} + \beta_{33}h_{22t-1} \tag{3.42}$$

対角 ***vech*** モデルとは、A と B の対角要素だけを用いたものである。つまり、$i \neq j$ に対し $\alpha_{ij} = \beta_{ij} = 0$ とすればよい。このとき、モデルは、

$$h_{11t} = c_{10} + \alpha_{11}\varepsilon_{1t-1}^2 + \beta_{11}h_{11t-1}$$

$$h_{12t} = c_{20} + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \beta_{22}h_{12t-1}$$

$$h_{22t} = c_{30} + \alpha_{33}\varepsilon_{2t-1}^2 + \beta_{33}h_{22t-1}$$

となる。

2. BEKK モデル

エングルら(Engle and Kroner (1995))は、条件付き分散が正であることを保証するため、いわゆる **BEKK** モデルを考案した。当初、論文は Baba, Engle, Kraft, and Kroner (1991)として発表されたため、著者たちの頭文字をとって BEKK モデルと呼ばれている。

このモデルのアイデアは、全てのパラメータを 2 次形式(quadratic form)とし、分散が正であることを保証している点にある。変数が k 個あるとき、BEKK モデルの分散共分散行列は

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B$$

と定式化される。A と B は $k \times k$ の行列である。しかし、C は $k \times k$ の対称行列でなければならない（h_{ijt}の非対角要素の定数が同一になるため）。

たとえば、2 変量なら、各行列は以下のように定義される。

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix}; C = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}; A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}; B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}$$

また、ε_t はベクトルであり、$\varepsilon_t=(\varepsilon_{1t}, \varepsilon_{2t})'$とする。この行列を計算すると、

$$h_{11t} = (c_{11}^2 + c_{12}^2) + (\alpha_{11}^2\varepsilon_{1t-1}^2 + 2\alpha_{11}\alpha_{21}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}^2\varepsilon_{2t-1}^2) + (\beta_{11}^2 h_{11t-1} + 2\beta_{11}\beta_{21}h_{12t-1} + \beta_{21}^2 h_{22t-1})$$

$$\begin{aligned} h_{12t} = {} & c_{12}(c_{11} + c_{22}) + \alpha_{12}\alpha_{11}\varepsilon_{1t-1}^2 + (\alpha_{11}\alpha_{22} + \alpha_{12}\alpha_{21})\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}\alpha_{22}\varepsilon_{2t-1}^2 \\ & + \beta_{11}\beta_{12}h_{11t-1} + (\beta_{11}\beta_{22} + \beta_{12}\beta_{21})h_{12t-1} + \beta_{21}\beta_{22}h_{22t-1} \end{aligned}$$

$$h_{22t} = (c_{22}^2 + c_{12}^2) + (\alpha_{12}^2\varepsilon_{1t-1}^2 + 2\alpha_{12}\alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{22}^2\varepsilon_{2t-1}^2) + (\beta_{12}^2 h_{11t-1} + 2\beta_{12}\beta_{22}h_{12t-1} + \beta_{22}^2 h_{22t-1})$$

となる。一般的に、h_{ijt}は、誤差項の 2 乗、誤差項の積、全ての条件付き分散共分散に依存している。

3. CCC モデル

CCC モデルは、多変量 GARCH の特殊ケースである。2 変量の場合、H_tは

$$H_t = \begin{bmatrix} h_{11t} & \rho_{12}(h_{11t}h_{22t})^{0.5} \\ \rho_{12}(h_{11t}h_{22t})^{0.5} & h_{22t} \end{bmatrix}$$

となる。h_{11t}とh_{22t}はともに GARCH(1, 1)過程である。また、$h_{12t} = \rho_{12}(h_{11t}h_{22t})^{1/2}$となる。システムのパラメータは 7 個だけである（条件付き分散の式は 2 本あり、各式は 3 つパラメータがあり、さらにρ_{12}を加えると、パラメータ数は 7 個となる)。

4. DCC モデル

エングル（Engle (2002)）は、相関係数を時変的とし、CCC モデルを拡張した DCC(Dynamic Conditional Correlation)モデルを提案している。DCC モデルでは、すべてのパラメータを同時推定するのではなく、2 段階に分けて推定を行う。第 1 段階では、CCC モデルを推定し、標準化残差を求める。標準化残差$s_{it} = \hat{\varepsilon}_{it} / \hat{h}_{iit}^{0.5}$は、$v_{it}$の推定値である。第 2 段階では、条件付き共分散を推定するため、標準化残差を用いる。具体的には、標準化残差を平滑化することで、相関係数を求める。エングルは、いくつかの平滑化法を分析している。まず、シンプルな指数的平滑化がある。

$$q_{ijt} = (1 - \lambda)s_{it}s_{jt} + \lambda q_{ijt-1}$$

ただし、$\lambda < 1$とする。ここで、$\{q_{iit}\}$は、標準化残差の交差項$s_{it}s_{jt}$の加重平均となる。このとき、時変的相関係数は、q_{ijt} を用いて

$$\rho_{ijt} = q_{ijt} / (q_{iit}q_{jjt})^{0.5} \qquad \text{(A3.4)}$$

と推定される。エングルは、2 段階推定によって、時変的相関係数の一致推定量が得られることを示した。しかし、こうした 2 段階推定は同時推定の場合に比べると効率的ではない。

別の平滑化法として、

$$q_{ijt} = (1 - \alpha - \beta)\overline{s}_{ij} + \alpha s_{it}s_{jt} + \beta q_{ijt-1}$$

がある。ここで、$\overline{s}_{ij}$は、s_{it}とs_{jt}間の条件なし共分散である（第 1 段階 CCC からの推定値)。この平滑化では、$\overline{s}_{ij}$の係数を（1－α－β）と置くことで、q_{ijt}が条件なし共分散に収束するようにしている。第 1 段階で推定された係数を尤度関数に入れれば、パラメータはαとβ だけとなる。

なぜ 2 段階推定が可能なのだろうか。ここでは、その証明をしよう。まず、D_t=対角行列（その要素は$h_{iit}^{0.5}$)、R_tは時変的相関係数の行列（その要素は$r_{ijt} = (h_{ijt})/(h_{iit}h_{jjt})^{0.5}$）とする。このとき、行列$H_t$は、相関係数の定義から

$$H_t = D_tR_tD_t$$

と表せる。たとえば、2 変量であれば、

$$D_t = \begin{pmatrix} h_{11t}^{0.5} & 1 \\ 1 & h_{22t}^{0.5} \end{pmatrix}、\quad \mathrm{R_t} = \begin{pmatrix} 1 & h_{12t}/(h_{11t}h_{22t})^{0.5} \\ h_{12t}/(h_{11t}h_{22t})^{0.5} & 1 \end{pmatrix}$$

となり、$H_t = D_tR_tD_t$ と書ける。D_t の逆行列を用いて、$R_t = (D_t)^{-1}H_t(D_t)^{-1}$ となる。

$$R_t = \begin{pmatrix} h_{11t}^{0.5} & 0 \\ 0 & h_{22t}^{0.5} \end{pmatrix}^{-1} \begin{pmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{pmatrix} \begin{pmatrix} h_{11t}^{0.5} & 0 \\ 0 & h_{22t}^{0.5} \end{pmatrix}^{-1} = \begin{pmatrix} 1 & h_{12t}/(h_{11t}h_{22t})^{0.5} \\ h_{12t}/(h_{11t}h_{22t})^{0.5} & 1 \end{pmatrix}$$

また、標準化残差を $\mathrm{v_t}$ とすると、$\mathrm{v_t = D_t^{-1}\varepsilon_t}$ となる。

$$v_t = \begin{pmatrix} h_{11t}^{0.5} & 1 \\ 1 & h_{22t}^{0.5} \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

標準化残差の 2 乗和は、$v_t'v_t$ であり、これは $v_t'v_t = \varepsilon_t' D_t^{-1}D_t^{-1}\varepsilon_t$ とも書ける。

$$\varepsilon_t' D_t^{-1}D_t^{-1}\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}' \begin{pmatrix} h_{11t}^{0.5} & 0 \\ 0 & h_{11t}^{0.5} \end{pmatrix}^{-1} \begin{pmatrix} h_{11t}^{0.5} & 0 \\ 0 & h_{11t}^{0.5} \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \frac{\varepsilon_{1t}^2}{h_{11t}} + \frac{\varepsilon_{2t}^2}{h_{22t}}$$

対数尤度(A3.3)の H_t に $D_tR_tD_t$ を代入すると、

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}(\ln|D_tR_tD_t| + \varepsilon_t'(D_tR_tD_t)^{-1}\varepsilon_t) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}(2\ln|D_t| + \ln|R_t| + v_t'(R_t)^{-1}v_t)$$

$$= -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}(2\ln|D_t| + \ln|R_t| + v_t'(R_t)^{-1}v_t - v_t'v_t + \varepsilon_t'D_t^{-1}D_t^{-1}\varepsilon_t) \qquad \text{(A3.5)}$$

となる。最後の式展開では、標準化残差の 2 乗を引いて、さらに加えるという操作をした。

2 段階推定が適切であることを示そう。対数尤度において、D_t と R_t が別々に入っていることに気づいてもらいたい。したがって、D_t と R_t は別々に推定できる。まず、CCC モデルによって、D_t のパラメータを推定する。これは R_t の値を知ることなくできる。次に、D_t の値を用いて、標準化残差を作れる。これを尤度関数に代入し、R_t のパラメータを推定する。厳密には、(A.3.5)の最大化は、第 1 段階で、

$$-\frac{1}{2}\sum_{t=1}^{T}(2\ln|D_t| + \varepsilon_t'D_t^{-1}D_t^{-1}\varepsilon_t)$$

を最大化し、第 2 段階で次式を最大化すればよい。

$$-\frac{1}{2}\sum_{t=1}^{T}(\ln|R_t| + v_t'(R_t)^{-1}v_t - v_t'v_t)$$