GARCH 理論の習得の覚書

百合川　学籍番号：201311324

2016/09/30

本稿は GARCH 理論の習得の覚書である．論文には理論部分の証明が *Appendix* として載せられている．その証明を読まないとモデルが何を表しているのかわからないので，省略されている部分を補って，後学のために詳解を付けておくのである．本文中のわかりづらかった部分も補完したものを載せてある．

本稿の参照論文:

Tim Bollerslev(1986),

"Generalized Autoregressive Conditional Heteroskedasticity",

Journal of Econometrics 31 (1986) 307 − 327. *North − Holland.*

§1． Appendix の Appendix

GARCH モデル:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i},$$
$$\alpha_0 \geq 0, \qquad \alpha_i > 0, \qquad \beta_i \geq 0.$$

定理 1　GARCH($p,\ q$) 過程について，広義定常である為の必要十分条件は $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1$ である:

$$\mathrm{E}\left[\epsilon_t\right] = 0,\ \mathrm{V}\left[\epsilon_t\right] = \alpha_0 \left(1 - \sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{p}\beta_i\right)^{-1},\ \mathrm{Cov}\left[\epsilon_t,\ \epsilon_s\right] = 0 \qquad \Leftrightarrow \qquad \sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1.$$

証明 1　ϵ_t は次の様に表されると仮定する:

$$\epsilon_t \equiv \eta_t h_t^{\frac{1}{2}}, \qquad \eta_t \sim N(0,\ 1)\ i.i.d.$$

各時点の ϵ_* を繰り返し h_t に代入する:

$$\begin{aligned}
h_t &= \alpha_0 + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} \\
&= \alpha_0 + \sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 h_{t-i} + \sum_{i=1}^{p}\beta_i h_{t-i} \\
&= \alpha_0 + \sum_{i=1}^{q}\alpha_i\eta_{t-i}^2\left(\alpha_0 + \sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2 h_{t-i-j} + \sum_{j=1}^{p}\beta_j h_{t-i-j}\right) \\
&\quad + \sum_{i=1}^{p}\beta_i\left(\alpha_0 + \sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2 h_{t-i-j} + \sum_{j=1}^{p}\beta_j h_{t-i-j}\right) \\
&= \alpha_0 + \alpha_0\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 + \sum_{i=1}^{p}\beta_i\right) \\
&\quad + \left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 + \sum_{i=1}^{p}\beta_i\right)\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2 h_{t-i-j}
\end{aligned}$$

$$
\begin{aligned}
&+\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right)\sum_{j=1}^{p}\beta_j h_{t-i-j}\\
&=\alpha_0+\alpha_0\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right)\\
&\quad+\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right)\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2\left(\alpha_0+\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2h_{t-i-j-l}+\sum_{l=1}^{p}\beta_l h_{t-i-j-l}\right)\\
&\quad+\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right)\sum_{j=1}^{p}\beta_j\left(\alpha_0+\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2h_{t-i-j-l}+\sum_{l=1}^{p}\beta_l h_{t-i-j-l}\right)\\
&=\alpha_0+\alpha_0\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right)+\alpha_0\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\\
&\quad+\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2h_{t-i-j-l}\\
&\quad+\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\sum_{l=1}^{p}\beta_l h_{t-i-j-l}\\
&=\alpha_0M(t,\ 0)+\alpha_0M(t,\ 1)+\alpha_0M(t,\ 2)\\
&\quad+M(t,\ 2)\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2\left(\alpha_0+\sum_{r=1}^{q}\alpha_r\eta_{t-i-j-l-r}^2h_{t-i-j-l-r}+\sum_{r=1}^{p}\beta_r h_{t-i-j-l-r}\right)\\
&\quad+M(t,\ 2)\sum_{l=1}^{p}\beta_l\left(\alpha_0+\sum_{r=1}^{q}\alpha_r\eta_{t-i-j-l-r}^2h_{t-i-j-l-r}+\sum_{r=1}^{p}\beta_r h_{t-i-j-l-r}\right)\\
&\vdots\\
&=\alpha_0\sum_{k=0}^{\infty}M(t,\ k).
\end{aligned}
$$

上式で記述されている通り，$M(t,\ k)$ は以下の規則で表される:

$$
\begin{aligned}
M(t,\ 0)&=1,\\
M(t,\ 1)&=\left(\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\sum_{i=1}^{p}\beta_i\right),\\
M(t,\ 2)&=\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\\
&=M(t-i,\ 1)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+M(t-i,\ 1)\sum_{i=1}^{p}\beta_i\\
M(t,\ 3)&=M(t,\ 2)\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2+M(t,\ 2)\sum_{l=1}^{p}\beta_l\\
&=\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2\\
&\quad+\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2+\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\sum_{i=1}^{p}\beta_i\right)\sum_{l=1}^{p}\beta_l\\
&=\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\left(\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2+\sum_{l=1}^{p}\beta_l\right)\right)\\
&\quad+\sum_{i=1}^{p}\beta_i\left(\left(\sum_{j=1}^{q}\alpha_j\eta_{t-i-j}^2+\sum_{j=1}^{p}\beta_j\right)\left(\sum_{l=1}^{q}\alpha_l\eta_{t-i-j-l}^2+\sum_{l=1}^{p}\beta_l\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= M(t-i,\ 2)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 + M(t-i,\ 2)\sum_{i=1}^{p}\beta_i \\
&\ \vdots \\
M(t,\ k+1) &= M(t-i,\ k)\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 + M(t-i,\ k)\sum_{i=1}^{p}\beta_i \\
&\ \vdots
\end{aligned}
$$

η_t は独立に分布 $N(0,\ 1)$ に従うという仮定の下，$M(t,\ k)$ のモーメントは時点 t に依存しない:

$$
\mathrm{E}\left[M(t,\ k)^r\right] = \mathrm{E}\left[M(s,\ k)^r\right], \qquad t \neq s,\ r \in \mathbb{N}.
$$

従って 1 次モーメントの計算は次の結果を得る．$k \geq 1$ として，

$$
\begin{aligned}
\mathrm{E}\left[M(t,\ k)\right] &= \mathrm{E}\left[M(t-i,\ k-1)\textstyle\sum_{i=1}^{q}\alpha_i\eta_{t-i}^2 + M(t-i,\ k-1)\sum_{i=1}^{p}\beta_i\right] \\
&= \left(\sum_{i=1}^{q}\alpha_i \mathrm{E}\left[\eta_{t-i}^2\right] + \sum_{i=1}^{p}\beta_i\right)\mathrm{E}\left[M(t,\ k-1)\right] \\
&= \left(\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i\right)^2 \mathrm{E}\left[M(t,\ k-2)\right] \\
&\ \vdots \\
&= \left(\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i\right)^k \mathrm{E}\left[M(t,\ 0)\right] \\
&= \left(\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i\right)^k.
\end{aligned}
$$

最終的に ϵ_t^2 の 1 次モーメントは，

$$
\mathrm{E}\left[\epsilon_t^2\right] = \mathrm{E}\left[\mathrm{E}\left[\epsilon_t^2 \mid \psi_{t-1}\right]\right] = \mathrm{E}\left[h_t\right] = \mathrm{E}\left[\alpha_0\textstyle\sum_{k=0}^{\infty} M(t,\ k)\right] = \alpha_0 \mathrm{E}\left[\sum_{k=0}^{\infty} M(t,\ k)\right].
$$

ここで定理の広義定常性の仮定から $\mathrm{E}\left[h_t\right] < \infty$ でなければならず，従って $\sum_{k=0}^{\infty} M(t,\ k)\ (\geq 0)$ は可積分である必要がある．正項級数が可積分なら *Lebesgue* の収束定理により項別積分に持ち込むことができ，よって $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1$ である．逆に，$\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1$ が仮定されていれば無限級数が可積分である為，

$$
\alpha_0 \mathrm{E}\left[\textstyle\sum_{k=0}^{\infty} M(t,\ k)\right] = \alpha_0 \displaystyle\sum_{k=0}^{\infty}\mathrm{E}\left[M(t,\ k)\right] = \alpha_0\left(1 - \sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{p}\beta_i\right)^{-1} < \infty
$$

が正当化される．そしてこれは時点に依存しない分散である．また，時点 $t-1$ までの情報が与えられた下で h_t は決定されることと，η_t が時点に関して独立に標準正規分布に従うことに注意すれば，

$$
\begin{aligned}
\mathrm{E}\left[\epsilon_t \mid \psi_{t-1}\right] &= \mathrm{E}\left[\eta_t\right] h_t^{\frac{1}{2}} = 0 \\
&\Rightarrow\ \mathrm{E}\left[\epsilon_t\right] = 0, \\
\mathrm{Cov}\left[\epsilon_t,\ \epsilon_s \mid \psi_{t-1}\right] &= \mathrm{Cov}\left[\eta_t h_t^{\frac{1}{2}},\ \eta_s h_s^{\frac{1}{2}} \mid \psi_{t-1}\right] \\
&= \mathrm{E}\left[\eta_t h_t^{\frac{1}{2}}\ \eta_s h_s^{\frac{1}{2}} \mid \psi_{t-1}\right] \\
&= 0 \\
&\Rightarrow\ \mathrm{Cov}\left[\epsilon_t,\ \epsilon_s\right] = 0.
\end{aligned}
$$

であることも従うので，定理が示されるのである． □

定理 2　GARCH(1, 1) について，$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$過程の $2m$ 次のモーメント $\mathrm{E}\left[\epsilon_t^{2m}\right]$ が存在する為の必要十分条件は

$$\mu(\alpha_1,\ \beta_1,\ m) \equiv \sum_{j=0}^{m} \binom{m}{j} a_j \alpha_1^j \beta_1^{m-j} < 1, \qquad a_0 = 1,\ a_j = \prod_{i=1}^{j}(2i-1)$$

であり，このとき $2m$ 次のモーメントは次の様に表せる：

$$\mathrm{E}\left[\epsilon_t^{2m}\right] = a_m \left[\sum_{n=0}^{m-1} a_n^{-1} \mathrm{E}\left[\epsilon_t^{2n}\right] \alpha_0^{m-n} \binom{m}{m-n} \mu(\alpha_1, \beta_1, n)\right] \times \left[1 - \mu(\alpha_1, \beta_1, m)\right]^{-1}.$$

証明 2　標準正規分布に従う確率変数 X の $2m$ $(m \in \mathbb{N})$ 次のモーメントの計算を確認しておく．

$$\begin{aligned}
\mathrm{E}\left[x^{2m}\right] &= \int_{-\infty}^{\infty} x^{2m} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= 2\frac{1}{\sqrt{2\pi}} \int_0^{\infty} (2x)^m e^{-x} \frac{1}{\sqrt{2x}} dx \\
&= \frac{1}{\sqrt{2\pi}} 2^{m+\frac{1}{2}} \Gamma(m + \frac{1}{2}) \\
&= \frac{1}{\sqrt{2\pi}} 2^{m+\frac{1}{2}} \frac{(2m-1)(2m-3)\cdots 1}{2^m} \sqrt{\pi} \\
&= \prod_{i=1}^{m}(2i-1).
\end{aligned}$$

η_tの正規性と，情報構造ψ_{t-1}が与えられた下で h_tが定数となることに注意すると以下の式が成り立つ．

$$\begin{aligned}
a_m &\equiv \mathrm{E}\left[\eta_t^{2m}\right] = \prod_{i=1}^{m}(2i-1), \\
\mathrm{E}\left[\epsilon_t^{2m}\right] &= \mathrm{E}\left[\mathrm{E}\left[\epsilon_t^{2m} \mid \psi_{t-1}\right]\right] = \mathrm{E}\left[\mathrm{E}\left[\eta_t^{2m} h_t^m \mid \psi_{t-1}\right]\right] = \mathrm{E}\left[h_t^m\right] \mathrm{E}\left[\eta_t^{2m}\right] = a_m \mathrm{E}\left[h_t^m\right].
\end{aligned}$$

GARCH(1, 1) の下では $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$であるから，

$$\begin{aligned}
h_t^m &= \left(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}\right)^m \\
&= \sum_{n=0}^{m} \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \alpha_1^j \beta_1^{n-j} \epsilon_{t-1}^2 j h_{t-1}^{n-j}.
\end{aligned}$$

従って，時点 $t-2$ までの情報を与えると，

$$\begin{aligned}
\mathrm{E}\left[h_t^m \mid \psi_{t-2}\right] &= \sum_{n=0}^{m} \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} \alpha_1^j \beta_1^{n-j} \mathrm{E}\left[\epsilon_{t-1}^{2j} h_{t-1}^{n-j} \mid \psi_{t-2}\right] \\
&= \sum_{n=0}^{m} \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} \alpha_1^j \beta_1^{n-j} h_{t-1}^{n-j} \mathrm{E}\left[\epsilon_{t-1}^{2j} \mid \psi_{t-2}\right] \\
&= \sum_{n=0}^{m} \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} \alpha_1^j \beta_1^{n-j} h_{t-1}^{n-j} h_{t-1}^{j} a_j \\
&= \sum_{n=0}^{m} \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j} h_{t-1}^{n}.
\end{aligned}$$

ここで m 次元ベクトル $\boldsymbol{w}_t \equiv (h_t^m, h_t^{m-1}, \cdots, h_t)^T$ に対して，期待値を取ると，

$$
\begin{aligned}
\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-2}\right] &= \left(\ \mathrm{E}\left[h_t^m \mid \psi_{t-2}\right],\ \mathrm{E}\left[h_t^{m-1} \mid \psi_{t-2}\right],\ \mathrm{E}\left[h_t^{m-2} \mid \psi_{t-2}\right],\ \cdots\ ,\ \mathrm{E}\left[h_t \mid \psi_{t-2}\right]\ \right)^T \\
&= \begin{pmatrix}
\sum_{n=0}^{m} h_{t-1}^n \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j} \\
\sum_{n=0}^{m-1} h_{t-1}^n \binom{m-1}{n} \alpha_0^{m-1-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j} \\
\sum_{n=0}^{m-2} h_{t-1}^n \binom{m-2}{n} \alpha_0^{m-2-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j} \\
\vdots \\
\sum_{n=0}^{1} h_{t-1}^n \binom{1}{n} \alpha_0^{1-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j}
\end{pmatrix} \\
&= a_0 \begin{pmatrix} \alpha_0^m \\ \alpha_0^{m-1} \\ \alpha_0^{m-2} \\ \vdots \\ \alpha_0 \end{pmatrix} \\
&\quad + \begin{pmatrix}
\sum_{j=0}^{m} \binom{m}{j} a_j \alpha_1^j \beta_1^{m-j} & \binom{m}{m-1} \alpha_0 \sum_{j=0}^{m-1} \binom{m-1}{j} a_j \alpha_1^j \beta_1^{m-1-j} & \binom{m}{m-2} \alpha_0^2 \sum_{j=0}^{m-2} \binom{m-2}{j} a_j \alpha_1^j \beta_1^{m-2-j} & \cdots & \binom{m}{1} \alpha_0^{m-1} \sum_{j=0}^{1} \binom{1}{j} a_j \alpha_1^j \beta_1^{1-j} \\
0 & \sum_{j=0}^{m-1} \binom{m-1}{j} a_j \alpha_1^j \beta_1^{m-1-j} & \binom{m-1}{m-2} \alpha_0 \sum_{j=0}^{m-2} \binom{m-2}{j} a_j \alpha_1^j \beta_1^{m-2-j} & \cdots & \binom{m-1}{1} \alpha_0^{m-2} \sum_{j=0}^{1} \binom{1}{j} a_j \alpha_1^j \beta_1^{1-j} \\
0 & 0 & \sum_{j=0}^{m-2} \binom{m-2}{j} a_j \alpha_1^j \beta_1^{m-2-j} & \cdots & \binom{m-2}{1} \alpha_0^{m-3} \sum_{j=0}^{1} \binom{1}{j} a_j \alpha_1^j \beta_1^{1-j} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \sum_{j=0}^{1} \binom{1}{j} a_j \alpha_1^j \beta_1^{1-j}
\end{pmatrix}
\begin{pmatrix} h_{t-1}^m \\ h_{t-1}^{m-1} \\ h_{t-1}^{m-2} \\ \vdots \\ h_{t-1} \end{pmatrix} \\
&= \boldsymbol{d} + \boldsymbol{C}\boldsymbol{w}_{t-1}.
\end{aligned}
$$

ここで $m \times m$ 上三角行列 $\boldsymbol{C}$ の対角成分に現れる級数を以下の様に特別に表記し直す．上三角行列なので行列式は対角成分のみから計算されるのである．

$$
\mu(\alpha_1, \beta_1,\ n) \equiv \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j}
$$

また，与えられる情報を減少していくと，すなわちψ_*の添え字の時点を後退させると以下の式が成り立つ：

$$
\begin{aligned}
\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-3}\right] &= \mathrm{E}\left[\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-2}\right] \mid \psi_{t-3}\right] \\
&= \mathrm{E}\left[\boldsymbol{d} + \boldsymbol{C}\boldsymbol{w}_{t-1} \mid \psi_{t-3}\right] \\
&= \boldsymbol{d} + \boldsymbol{C}\mathrm{E}\left[\boldsymbol{w}_{t-1} \mid \psi_{t-3}\right] \\
&= \boldsymbol{d} + \boldsymbol{C}(\boldsymbol{d} + \boldsymbol{C}\boldsymbol{w}_{t-2}) \\
&= (\boldsymbol{I} + \boldsymbol{C})\boldsymbol{d} + \boldsymbol{C}^2\boldsymbol{w}_{t-2}, \\
\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-4}\right] &= \mathrm{E}\left[\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-3}\right] \mid \psi_{t-4}\right] \\
&= \mathrm{E}\left[(\boldsymbol{I} + \boldsymbol{C})\boldsymbol{d} + \boldsymbol{C}^2\boldsymbol{w}_{t-2} \mid \psi_{t-4}\right] \\
&= (\boldsymbol{I} + \boldsymbol{C})\boldsymbol{d} + \boldsymbol{C}^2\mathrm{E}\left[\boldsymbol{w}_{t-2} \mid \psi_{t-4}\right] \\
&= (\boldsymbol{I} + \boldsymbol{C} + \boldsymbol{C}^2)\boldsymbol{d} + \boldsymbol{C}^3\boldsymbol{w}_{t-3}, \\
&\ \ \vdots \\
\mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-k-1}\right] &= (\boldsymbol{I} + \boldsymbol{C} + \boldsymbol{C}^2 + \cdots + \boldsymbol{C}^{k-1})\boldsymbol{d} + \boldsymbol{C}^k\boldsymbol{w}_{t-k}, \\
&\ \ \vdots
\end{aligned}
$$

これは $k \to \infty$ まで考えると，情報が無い下での期待値，即ち $\mathrm{E}\left[\boldsymbol{w}_t\right]$ を表すことになり，$\mathrm{E}\left[h_t^m\right]$ が $\mathrm{E}\left[\boldsymbol{w}_t\right]$ の第一成分であることから，行列の冪級数の収束が求める $\mathrm{E}\left[\epsilon_t^{2m}\right]$ の存在の必要十分条件である：
ところで行列 $\boldsymbol{C}$ のノルムが 1 より小さいとき，冪級数は収束する．これを証明してみる．

行列 $\boldsymbol{A}$ のスペクトルノルムの定義は次の様に表される:

$$
\|\boldsymbol{A}\| = \sup_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}.
$$

定義に基づけば

$$\|\boldsymbol{AB}\| = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{ABx}\|}{\|\boldsymbol{x}\|} \leq \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\|\|\boldsymbol{Bx}\|}{\|\boldsymbol{x}\|} = \|\boldsymbol{A}\|\|\boldsymbol{B}\|,$$

が成り立ち，このことを利用すれば，$\|\boldsymbol{A}\| < \alpha$ であるとき，

$$\|\boldsymbol{A^n}\| = \left\|\boldsymbol{AA^{n-1}}\right\| \leq \|\boldsymbol{A}\|\left\|\boldsymbol{A^{n-1}}\right\| \leq \|\boldsymbol{A}\|^2\left\|\boldsymbol{A}^{n-2}\right\| \cdots \leq \|\boldsymbol{A}\|^n \ < \ \alpha^n$$

と表すことが出来る．ここで冪級数の話に戻る．冪級数の全体のノルムを考えると，

$$\left\|\boldsymbol{I} + \boldsymbol{C} + \boldsymbol{C}^2 + \cdots\right\| \leq \|\boldsymbol{I}\| + \|\boldsymbol{C}\| + \left\|\boldsymbol{C}^2\right\| + \cdots,$$

従って実数値の冪級数の収束条件が適用できて， $\|\boldsymbol{C}\| < 1$ であるなら上式左辺が項数に関して *Cauchy* 列をなす:

$$\forall\epsilon > 0,\ \exists N \in \mathbb{N},\ \forall n1, n2 \geq N,$$
$$\left\|\sum_{i=n1}^{n2} \boldsymbol{C}^i\right\| \leq \sum_{i=n1}^{n2} \|\boldsymbol{C}\|^i < \epsilon.$$

即ち或る有限な成分から構成される行列が存在し，冪級数はそれに収束する．収束先の計算は以下の様にする:

$$(\boldsymbol{I} - \boldsymbol{C})\sum_{i=0}^{\infty} \boldsymbol{C}^i = \sum_{i=0}^{\infty} \boldsymbol{C}^i - \sum_{i=1}^{\infty} \boldsymbol{C}^i = \boldsymbol{I},$$
$$\Rightarrow \sum_{i=0}^{\infty} \boldsymbol{C}^i = (\boldsymbol{I} - \boldsymbol{C})^{-1}.$$

$\|\boldsymbol{C}\| < 1$ が成り立つときには $\boldsymbol{C}$ の各成分 $c_{j,i}$ は $|c_{j,i}| < 1$ を満たす．これは次の様にして成り立つことである:

第 i 成分が 1 である m 次元ベクトル $\quad \boldsymbol{x} = (0, 0, \cdots, 0, 1, 0, \cdots, 0)^T,\ (i = 1, 2, \cdots, m)$ に対して，

$$|c_{j,i}| < \sqrt{\sum_{j=1}^{m} c_{j,i}^2} = \frac{\|\boldsymbol{Cx}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{C}\| < 1.$$

上三角行列の固有値は全て対角成分$\mu(\alpha_1, \beta_1,\ n)$ $(n = 1, 2, \cdots, m)$ に等しい．GARCH モデルの仮定からα_1, β_1は非負であり，a_jは標準正規分布の偶数次のモーメントであるから正数で，即ち，$\mu(\alpha_1, \beta_1,\ n)$ は n に関して非負単調増大である：

$$\begin{aligned}
\mu(\alpha_1, \beta_1, n+1) - \mu(\alpha_1, \beta_1, n) &= \sum_{j=0}^{n+1} \binom{n+1}{j} a_j \alpha_1^j \beta_1^{n+1-j} - \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j} \\
&= a_{n+1}\alpha_1^{n+1} + \sum_{j=0}^{n} \left(\binom{n+1}{j}\beta_1 - \binom{n}{j}\right) a_j \alpha_1^j \beta_1^{n-j} \\
&\geq a_{n+1}\alpha_1^{n+1} - \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j}
\end{aligned}$$

$\mu(\alpha_1, \beta_1,\ m) < 1$ であることが冪級数の収束の十分条件を為す．よって$\mu(\alpha_1, \beta_1,\ m) < 1$ であることは $\mathrm{E}\left[\epsilon_t^{2m}\right]$ の存在の十分条件となる．一方，$\mathrm{E}\left[\epsilon_t^{2m}\right]$ が存在すれば $2(m-1),\ 2(m-2),\ \cdots,\ 2$ 次モーメントも存在することになるので，冪級数は収束し，$\mu(\alpha_1, \beta_1,\ m) < 1$ が従う．

最後に

$$\mathrm{E}\left[\epsilon_t^{2m}\right] = a_m \left[\sum_{n=0}^{m-1} a_n^{-1} \mathrm{E}\left[\epsilon_t^{2n}\right] \alpha_0^{m-n} \binom{m}{m-n} \mu(\alpha_1, \beta_1, n)\right] \times [1 - \mu(\alpha_1, \beta_1, m)]^{-1}$$

を示す．

$$\mathrm{E}\left[\epsilon_t^{2m}\right] = a_m \mathrm{E}\left[h_t^m\right],$$
$$\mathrm{E}\left[h_t^m \mid \psi_{t-2}\right] = \sum_{n=0}^{m} \quad h_{t-1}^n \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j}$$

が成り立っていることにより，

$$
\begin{aligned}
\mathrm{E}\left[\epsilon_t^{2m}\right] &= a_m \mathrm{E}\left[h_t^m\right] \\
&= a_m \mathrm{E}\left[\mathrm{E}\left[h_t^m \mid \psi_{t-2}\right]\right] \\
&= a_m \mathrm{E}\left[\textstyle\sum_{n=0}^{m} \quad h_{t-1}^n \binom{m}{n} \alpha_0^{m-n} \sum_{j=0}^{n} \binom{n}{j} a_j \alpha_1^j \beta_1^{n-j}\right] \\
&= a_m \left[\sum_{n=0}^{m} \quad \mathrm{E}\left[h_{t-1}^n\right] \binom{m}{n} \alpha_0^{m-n} \mu(\alpha_1, \beta_1,\ n)\right] \\
&= a_m \left[\sum_{n=0}^{m-1} \quad a_n^{-1} \mathrm{E}\left[\epsilon_{t-1}^{2n}\right] \binom{m}{n} \alpha_0^{m-n} \mu(\alpha_1, \beta_1,\ n)\right] + a_m a_m^{-1} \mathrm{E}\left[\epsilon_{t-1}^{2m}\right] \mu(\alpha_1, \beta_1,\ m).
\end{aligned}
$$

$\mathrm{E}\left[\epsilon_t^{2m}\right]$ が時点に依らない値を持つことに注意しなければならない，というのも，先述の行列の冪級数の収束の箇所で，

$$
\begin{pmatrix}
a_m^{-1} \mathrm{E}\left[\epsilon_t^{2m}\right] \\
a_{m-1}^{-1} \mathrm{E}\left[\epsilon_t^{2(m-1)}\right] \\
a_{m-2}^{-1} \mathrm{E}\left[\epsilon_t^{2(m-2)}\right] \\
\vdots \\
a_1^{-1} \mathrm{E}\left[\epsilon_t^{2}\right]
\end{pmatrix}
= \mathrm{E}\left[\boldsymbol{w}_t\right] = \lim_{k\to\infty} \mathrm{E}\left[\boldsymbol{w}_t \mid \psi_{t-k-1}\right] = (\boldsymbol{I} - \boldsymbol{C})^{-1} \boldsymbol{d}
$$

が成り立っているのである．従って，最右辺第二項を左辺に移せば，求める式

$$
\mathrm{E}\left[\epsilon_t^{2m}\right] = a_m \left[\sum_{n=0}^{m-1} \quad a_n^{-1} \mathrm{E}\left[\epsilon_t^{2n}\right] \binom{m}{n} \alpha_0^{m-n} \mu(\alpha_1, \beta_1,\ n)\right] \left[1 - \mu(\alpha_1, \beta_1,\ m)\right]^{-1}
$$

が導かれる． □

§2． 本文中の補完説明

GARCH モデル式の同値形 原文第 2 章

本稿の最初に示した GARCH モデルの式の同値形を表記する:

$$
\begin{aligned}
\epsilon_t \equiv \eta_t h_t^{\frac{1}{2}}, &\qquad \eta_t \sim N(0,\ 1)\ i.i.d.\text{を仮定している下で，} \\
h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \\
&= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i (h_{t-i} - \epsilon_{t-i}^2) \\
&= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} (1 - \eta_{t-i}^2).
\end{aligned}
$$

両辺に $\epsilon_t^2 - h_t$ を加えると，

$$
\begin{aligned}
\epsilon_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} (1 - \eta_{t-i}^2) + \epsilon_t^2 - h_t \\
&= \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} (1 - \eta_{t-i}^2) + h_t (\eta_t^2 - 1).
\end{aligned}
$$

確率過程の自己相関係数 原文第 4 章

GARCH$(p,\ q)$ 過程に 4 次モーメントの存在を仮定する．共分散を次の様に表記する．

$$
\gamma_n \ \equiv \ \gamma_{-n} \ \equiv \mathrm{Cov}\left[\epsilon_t^2,\ \epsilon_{t-n}^2\right].
$$

上のスクリーンで表されるモデル式を利用すれば，共分散の計算は以下の様に書き下せる:

$$
\gamma_n \ = \mathrm{Cov}\left[\epsilon_t^2,\ \epsilon_{t-n}^2\right]
$$

$$
\begin{aligned}
&= \mathrm{Cov}\left[\alpha_0 + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}(1-\eta_{t-i}^2) + h_t(\eta_t^2-1),\ \epsilon_{t-n}^2\right] \\
&= \mathrm{Cov}\left[\sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}(1-\eta_{t-i}^2) + h_t(\eta_t^2-1),\ \epsilon_{t-n}^2\right].
\end{aligned}
$$

ここで，モデルの定常性が仮定されていれば，時点に依らず $\mathrm{E}\left[\epsilon_t^2\right] = 0$,
$\mathrm{E}\left[h_t(\eta_t^2-1)\right] = \mathrm{E}\left[h_t\right]\mathrm{E}\left[\eta_t^2-1\right] = 0$, $\mathrm{Cov}\left[\epsilon_t^2,\ \epsilon_{t-n}^2\right] = \mathrm{E}\left[\epsilon_t^2\epsilon_{t-n}^2\right]$ が成り立つので，共分散は

$$
\begin{aligned}
&\mathrm{Cov}\left[\sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}(1-\eta_{t-i}^2) + h_t(\eta_t^2-1),\ \epsilon_{t-n}^2\right] \\
&= \mathrm{E}\left[\left\{\sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}(1-\eta_{t-i}^2) + h_t(\eta_t^2-1)\right\}\epsilon_{t-n}^2\right] \\
&= \sum_{i=1}^{q}\alpha_i\mathrm{E}\left[\epsilon_{t-i}^2\epsilon_{t-n}^2\right] + \sum_{i=1}^{p}\beta_i\mathrm{E}\left[\epsilon_{t-i}^2\epsilon_{t-n}^2\right] + \sum_{i=1}^{p}\beta_i\mathrm{E}\left[h_{t-i}(1-\eta_{t-i}^2)\epsilon_{t-n}^2\right] + \mathrm{E}\left[h_t(\eta_t^2-1)\epsilon_{t-n}^2\right] \\
&= \sum_{i=1}^{q}\alpha_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\mathrm{E}\left[h_{t-i}(1-\eta_{t-i}^2)\epsilon_{t-n}^2\right] + \mathrm{E}\left[h_t\right]\mathrm{E}\left[\eta_t^2-1\right]\mathrm{E}\left[\epsilon_{t-n}^2\right] \\
&= \sum_{i=1}^{q}\alpha_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\mathrm{E}\left[h_{t-i}(1-\eta_{t-i}^2)\epsilon_{t-n}^2\right].
\end{aligned}
$$

と表すことが出来る．最右辺第二項には注意が必要で，$n \geq p+1$ でないと $(1-\eta_{t-i}^2)$ と ϵ_{t-n}^2 との独立性が保証されない．ここで $n \geq p+1$ の下で再び共分散式に戻ろう．

$$
\begin{aligned}
&\sum_{i=1}^{q}\alpha_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\gamma_{n-i} + \sum_{i=1}^{p}\beta_i\mathrm{E}\left[h_{t-i}(1-\eta_{t-i}^2)\epsilon_{t-n}^2\right] = \sum_{i=1}^{max(p,q)}(\alpha_i+\beta_i)\gamma_{n-i} = \sum_{i=1}^{m}\phi_i\gamma_{n-i}, \\
&s.t.\quad n \geq p+1,\quad m \equiv max(p,q),\quad \phi_i \equiv \alpha_i+\beta_i,\quad \alpha_i \equiv 0\ (i \geq q+1),\quad \beta_i \equiv 0\ (i \geq p+1).
\end{aligned}
$$

$\gamma_0^{-1} = \mathrm{E}\left[\epsilon_t^4\right] < \infty$ が時点に依らない値であることに注意して，時差 n の自己相関係数

$$
\rho_n \equiv \frac{\mathrm{Cov}\left[\epsilon_t^2,\ \epsilon_{t-n}^2\right]}{\sqrt{\mathrm{V}\left[\epsilon_t^2\right]\mathrm{V}\left[\epsilon_{t-n}^2\right]}} = \frac{\mathrm{Cov}\left[\epsilon_t^2,\ \epsilon_{t-n}^2\right]}{\mathrm{V}\left[\epsilon_t^2\right]} = \gamma_n\gamma_0^{-1}
$$

に関して次の等式を得る．

$$
\rho_n = \gamma_n\gamma_0^{-1} = \sum_{i=1}^{m}\phi_i\gamma_{n-i}\gamma_0^{-1} = \sum_{i=1}^{m}\phi_i\rho_{n-i},\quad n \geq p+1.
$$

この式が表していることは，自己相関係数 $\rho_p,\ \rho_{p-1},\ \rho_{p-2},\ \cdots,\ \rho_{p+1-m}$ によって任意の時点 n での ρ_n が一意に定まるということである．

GARCH モデルのパラメータの最尤推定 原文第 5 章

GARCH モデルの式をベクトルの積で表現し直す:

$$
\begin{aligned}
h_t &= \alpha_0 + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} \\
&= \boldsymbol{z}_t^T\boldsymbol{\omega}_t, \\
s.t.\quad \boldsymbol{z}_t &\equiv (1,\ \epsilon_{t-1}^2,\ \cdots, \epsilon_{t-q}^2,\ h_{t-1},\ \cdots,\ h_{t-p})^T, \\
\boldsymbol{\omega} &\equiv (\alpha_0,\ \alpha_1,\ \cdots,\ \alpha_q,\ \beta_1,\ \cdots,\ \beta_p)^T.
\end{aligned}
$$

推定するパラメータは $1+q+p$ 個だけではない．ここで，時点 t における確率変数であるϵ_tが線形回帰式の誤差項であると考えて，

$$
\epsilon_t = y_t - \boldsymbol{x_t}^T\boldsymbol{b}
$$

と置くと，推定するパラメータの個数は $dim(\boldsymbol{b})+dim(\boldsymbol{\omega})$ となる．モデルの母数空間が $Euclidean$ 空間のコンパクトな部分空間と考える．そして確率過程 $\{\epsilon_t\}_t$の二次モーメントが存在する母数空間であるとする．この母数空間を

$$\Theta \equiv \{\boldsymbol{\theta} = (\boldsymbol{b}^T,\ \boldsymbol{\omega}^T)\}$$

とし，真のパラメータ $\boldsymbol{\theta}_0$をΘの内部に含むものとする．最尤推定とは，与えられた観測を基に尤度関数を最大にする$\hat{\boldsymbol{\theta}} \in \Theta$を求めるものである．$T$ この時点の観測に対する対数尤度関数は

$$L_T(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^{T} l_t(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^{T}\left(-\frac{1}{2}logh_t - \frac{1}{2}\epsilon_t^2 h_t^{-1}\right).$$

$\boldsymbol{\omega}$ に関して微分すると以下の様になる：

$$\begin{aligned}
\boldsymbol{\omega} &= (\omega_1,\ \omega_2,\ \cdots)^T \text{と表して,}\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial \omega_i} &= -\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_i} + \frac{1}{2}\frac{\epsilon_t^2}{h_t^2}\frac{\partial h_t}{\partial \omega_i},\\
&= \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_i}\left(\frac{\epsilon_t^2}{h_t} - 1\right),\\
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \omega_j \partial \omega_i} &= \frac{1}{2}\frac{-1}{h_t^2}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial \omega_i}\left(\frac{\epsilon_t^2}{h_t} - 1\right) + \frac{1}{2}h_t^{-1}\frac{\partial^2 h_t}{\partial \omega_j \partial \omega_i}\left(\frac{\epsilon_t^2}{h_t} - 1\right) + \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_i}\frac{-\epsilon_t^2}{h_t^2}\frac{\partial h_t}{\partial \omega_j}\\
&= \frac{1}{2}\left(\frac{\epsilon_t^2}{h_t} - 1\right)\left(h_t^{-1}\frac{\partial^2 h_t}{\partial \omega_j \partial \omega_i} - \frac{1}{h_t^2}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial \omega_i}\right) - \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial \omega_i}\frac{\epsilon_t^2}{h_t^2}\\
&= \left(\frac{\epsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \omega_j}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_i}\right] - \frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial \omega_i}\frac{\epsilon_t^2}{h_t} \qquad \text{と計算できるから,}\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}} &= \left(\frac{\partial l_t}{\partial \omega_1},\ \frac{\partial l_t}{\partial \omega_2},\ \cdots\right)^T\\
&= \frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\left(\frac{\epsilon_t^2}{h_t} - 1\right),\\
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}^T} &= \begin{pmatrix}
\frac{\partial^2 l_t}{\partial \omega_1^2} & \frac{\partial^2 l_t}{\partial \omega_1 \partial \omega_2} & \frac{\partial^2 l_t}{\partial \omega_1 \partial \omega_3} & \cdots\\
\frac{\partial^2 l_t}{\partial \omega_2 \partial \omega_1} & \frac{\partial^2 l_t}{\partial \omega_2^2} & \frac{\partial^2 l_t}{\partial \omega_2 \partial \omega_3} & \cdots\\
\frac{\partial^2 l_t}{\partial \omega_3 \partial \omega_1} & \frac{\partial^2 l_t}{\partial \omega_3 \partial \omega_2} & \frac{\partial^2 l_t}{\partial \omega_3^2} & \cdots\\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}\\
&= \left(\frac{\epsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \boldsymbol{\omega}}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \boldsymbol{\omega}^T}\right] - \frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\frac{\partial h_t}{\partial \boldsymbol{\omega}^T}\frac{\epsilon_t^2}{h_t}.
\end{aligned}$$

ここで，$\frac{\partial h_t}{\partial \boldsymbol{\omega}}$ は次のように示される．

$$\begin{aligned}
\boldsymbol{z}_t &= (z_1,\ z_2,\ \cdots) \text{ と表して,}\\
\frac{\partial h_t}{\partial \omega_i} &= \frac{\partial \boldsymbol{z}_t^T \boldsymbol{\omega}_t}{\partial \omega_i} = \begin{cases} z_i & (\omega \neq \beta)\\ z_i + \omega_i \frac{\partial z_i}{\partial \omega_i} & (\omega = \beta)\end{cases} = \begin{cases} z_i & (\omega \neq \beta)\\ z_i + \beta_i \frac{\partial h_i}{\partial \omega_i} & (\omega = \beta)\end{cases} \qquad \text{であるから,}\\
\frac{\partial h_t}{\partial \boldsymbol{\omega}} &= \boldsymbol{z}_t + \sum_{i=1}^{p}\beta_i \frac{\partial h_{t-i}}{\partial \boldsymbol{\omega}}.
\end{aligned}$$

$(0 \leq)\ \sum_{i=1}^{p}\beta_i < 1$ であるなら $\frac{\partial h_t}{\partial \boldsymbol{\omega}} = \boldsymbol{z}_t + \sum_{i=1}^{p}\beta_i \frac{\partial h_{t-i}}{\partial \boldsymbol{\omega}}$ は定常となる．
ここで $Ficher$ 情報行列

$$\boldsymbol{I}(\boldsymbol{\theta}) \equiv -\mathrm{E}\left[\frac{\partial^2 l_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^T}|\psi_{t-1}\right] = \mathrm{E}\left[\frac{\partial l_t}{\partial \boldsymbol{\theta}}\frac{\partial l_t}{\partial \boldsymbol{\theta}^T}|\psi_{t-1}\right]$$

を考える．$\frac{\partial^2 l_t}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}^T}$ の右辺第一項 $\left(\frac{\epsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \boldsymbol{\omega}^T}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\right]$ の条件付き期待 0 となる:

$$\mathrm{E}\left[\left(\frac{\epsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \boldsymbol{\omega}^T}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\right]|\psi_{t-1}\right] = \mathrm{E}\left[\frac{\epsilon_t^2}{h_t} - 1\right]\frac{\partial}{\partial \boldsymbol{\omega}^T}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \boldsymbol{\omega}}\right] = 0, \qquad \because \mathrm{E}\left[\epsilon_t^2|\psi_{t-1}\right] = \mathrm{V}\left[\epsilon_t|\psi_{t-1}\right] = h_t.$$

この為，$Ficher$ 情報行列 $\boldsymbol{I(\theta)}$ の $\boldsymbol{\omega}$ に対応する部分は $\frac{\partial^2 l_t}{\partial\boldsymbol{\omega}\partial\boldsymbol{\omega}^T}$ の右辺第二項のみ関係することになる．以上では対数尤度関数を $\boldsymbol{\omega}$ についてみていたが，次に $\boldsymbol{b}$ について微分する:

$$
\begin{aligned}
&\boldsymbol{b}=(b_1,\ b_2,\ \cdots)^T,\ \boldsymbol{x}_t=(x_{t,1},\ x_{t,2},\ \cdots)^T \text{と表して，}\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial b_i} &= -\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}-\frac{1}{2}\frac{2\epsilon_t\frac{\partial\epsilon_t}{\partial b_i}h_t-\epsilon_t^2\frac{\partial h_t}{\partial b_i}}{h_t^2},\\
&= \epsilon_t x_{t,i}h_t^{-1}+\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right),\\
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial b_j\partial b_i} &= \frac{\partial\epsilon_t}{\partial b_j}x_{t,i}h_t^{-1}+\epsilon_t x_{t,i}\frac{-1}{h_t^2}\frac{\partial h_t}{\partial b_j}\\
&\quad+\frac{1}{2}\frac{-1}{h_t^2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)+\frac{1}{2}h_t^{-1}\frac{\partial^2 h_t}{\partial b_j\partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)+\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}\frac{2\epsilon_t\frac{\partial\epsilon_t}{\partial b_j}h_t-\epsilon_t^2\frac{\partial h_t}{\partial b_j}}{h_t^2}\\
&= -h_t^{-1}x_{t,j}x_{t,i}-\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}\right)\\
&\quad-\epsilon_t x_{t,i}\frac{1}{h_t^2}\frac{\partial h_t}{\partial b_j}-\epsilon_t x_{t,j}\frac{1}{h_t^2}\frac{\partial h_t}{\partial b_i}+\left(\frac{\epsilon_t^2}{h_t}-1\right)\left[\frac{1}{2}h_t^{-1}\frac{\partial^2 h_t}{\partial b_j\partial b_i}-\frac{1}{2}\frac{1}{h_t^2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial b_i}\right]\\
&= -h_t^{-1}x_{t,j}x_{t,i}-\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}\right)\\
&\quad-\epsilon_t x_{t,i}h_t^{-2}\frac{\partial h_t}{\partial b_j}-\epsilon_t x_{t,j}h_t^{-2}\frac{\partial h_t}{\partial b_i}+\left(\frac{\epsilon_t^2}{h_t}-1\right)\frac{\partial}{\partial b_j}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}\right]\qquad\text{と計算できるから，}\\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial\boldsymbol{b}} &= \left(\frac{\partial l_t}{\partial b_1},\ \frac{\partial l_t}{\partial b_2},\ \cdots\right)^T\\
&= \epsilon_t\boldsymbol{x}_t h_t^{-1}+\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial\boldsymbol{b}}\left(\frac{\epsilon_t^2}{h_t}-1\right),\\
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{b}^T\partial\boldsymbol{b}} &= \begin{pmatrix}\frac{\partial^2 l_t}{\partial b_1^2} & \frac{\partial^2 l_t}{\partial b_1\partial b_2} & \frac{\partial^2 l_t}{\partial b_1\partial b_3} & \cdots\\ \frac{\partial^2 l_t}{\partial b_2\partial b_1} & \frac{\partial^2 l_t}{\partial b_2^2} & \frac{\partial^2 l_t}{\partial b_2\partial b_3} & \cdots\\ \frac{\partial^2 l_t}{\partial b_3\partial b_1} & \frac{\partial^2 l_t}{\partial b_3\partial b_2} & \frac{\partial^2 l_t}{\partial b_3^2} & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix}\\
&= -h_t^{-1}\boldsymbol{x}_t\boldsymbol{x}_t^T-\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial\boldsymbol{b}}\frac{\partial h_t}{\partial\boldsymbol{b}^T}\left(\frac{\epsilon_t^2}{h_t}\right)\\
&\quad-\epsilon_t h_t^{-2}\frac{\partial h_t}{\partial\boldsymbol{b}}\boldsymbol{x}_t^T-\epsilon_t h_t^{-2}\boldsymbol{x}_t\frac{\partial h_t}{\partial\boldsymbol{b}^T}+\left(\frac{\epsilon_t^2}{h_t}-1\right)\frac{\partial}{\partial\boldsymbol{b}}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial\boldsymbol{b}^T}\right].
\end{aligned}
$$

ただし，

$$
\frac{\partial h_t}{\partial\boldsymbol{b}}=-2\sum_{i=1}^{q}\alpha_i\epsilon_{t-i}\boldsymbol{x_{t-i}}+\sum_{i=1}^{p}\beta_i\frac{\partial h_{t-i}}{\partial\boldsymbol{b}}.
$$

$Ficher$ 情報行列 $\boldsymbol{I(\theta)}$ の $\boldsymbol{b}$ に対応する部分を考えると，$\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{b}^T\partial\boldsymbol{b}}$ の第三項と第四項の条件付き期待値が

$$
\begin{aligned}
\mathrm{E}\left[\epsilon_t h_t^{-2}\tfrac{\partial h_t}{\partial\boldsymbol{b}}\boldsymbol{x}_t^T|\psi_{t-1}\right] &= \mathrm{E}\left[\epsilon_t|\psi_{t-1}\right]\frac{\partial h_t}{\partial\boldsymbol{b}}\boldsymbol{x}_t^T=0,\\
\mathrm{E}\left[\left(\tfrac{\epsilon_t^2}{h_t}-1\right)\tfrac{\partial}{\partial\boldsymbol{b}}\left[\tfrac{1}{2}h_t^{-1}\tfrac{\partial h_t}{\partial\boldsymbol{b}^T}\right]|\psi_{t-1}\right] &= \mathrm{E}\left[\tfrac{\epsilon_t^2}{h_t}-1|\psi_{t-1}\right]\frac{\partial}{\partial\boldsymbol{b}}\left[\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial\boldsymbol{b}^T}\right]=0,
\end{aligned}
$$

となるから，右辺第一項と右辺第二項のみ関係することになる．
最後に，対角ブロックでない部分，即ち $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\omega}^T\partial\boldsymbol{b}}$ と $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{b}^T\partial\boldsymbol{\omega}}$ を計算する:

$$
\begin{aligned}
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial b_j\partial\omega_i} &= \frac{\partial}{\partial b_j}\left(\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial\omega_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)\right)\\
&= -\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial\omega_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}h_t^{-1}\frac{\partial^2 h_t}{\partial b_j \partial \omega_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)\\
&+\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial \omega_i}\frac{2\epsilon_t\frac{\partial \epsilon_t}{\partial b_j}h_t-\epsilon_t^2\frac{\partial h_t}{\partial b_j}}{h_t^2}, \qquad \Rightarrow \mathrm{E}\left[\tfrac{\partial^2 l_t(\boldsymbol{\theta})}{\partial b_j \partial \omega_i}|\psi_{t-1}\right]=-\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial b_j}\frac{\partial h_t}{\partial \omega_i}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \omega_j \partial b_i} &= \frac{\partial}{\partial \omega_j}\left(\epsilon_t x_{t,i}h_t^{-1}+\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)\right)\\
&= \epsilon_t x_{t,i}h_t^{-2}\frac{\partial h_t}{\partial \omega_j}\\
&\quad+\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)\\
&\quad+\frac{1}{2}h_t^{-1}\frac{\partial^2 h_t}{\partial \omega_j \partial b_i}\left(\frac{\epsilon_t^2}{h_t}-1\right)\\
&\quad-\frac{1}{2}h_t^{-1}\frac{\partial h_t}{\partial b_i}\epsilon_t^2 h_t^{-2}\frac{\partial h_t}{\partial \omega_j}, \qquad \Rightarrow \mathrm{E}\left[\tfrac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \omega_j \partial b_i}|\psi_{t-1}\right]=-\frac{1}{2}h_t^{-2}\frac{\partial h_t}{\partial \omega_j}\frac{\partial h_t}{\partial b_i}.
\end{aligned}
$$

ここで $Ficher$ 情報行列の計算は終わる．
続いて最尤推定量の計算に移る．原論文 317 頁の $Berndt$, $Hall$, $Hall\ and\ Hausman$(1974) のアルゴリズムは $Newton$ 法のような再帰計算をすることで $\frac{\partial}{\partial \boldsymbol{\theta}}L_T(\boldsymbol{\theta})=\frac{\partial}{\partial \boldsymbol{\theta}}T^{-1}\sum_{t=1}^{T}l_t(\boldsymbol{\theta})=0$ の解を求めている:

$$
\boldsymbol{\theta}^{(i+1)}=\boldsymbol{\theta}^{(i)}+\lambda_i\left(\sum_{t=1}^{T}\frac{\partial l_t(\boldsymbol{\theta}^{(i)})}{\partial \boldsymbol{\theta}}\frac{\partial l_t(\boldsymbol{\theta}^{(i)})}{\partial \boldsymbol{\theta}^T}\right)^{-1}\sum_{t=1}^{T}\frac{\partial l_t(\boldsymbol{\theta}^{(i)})}{\partial \boldsymbol{\theta}}.
$$