

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:

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

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:

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

証明 1 ϵ_t は次の様に表されると仮定する:

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

各時点の ϵ_* を繰り返し h_t に代入する:

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

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

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：

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

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

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:

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

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,

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

最終的に ϵ_t^2 の 1 次モーメントは,

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

ここで定理の広義定常性の仮定から $\mathbb{E}[h_t] < \infty$ 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,

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

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,

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

であることも従うので, 定理が示されるのである. \square

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 $E[\epsilon_t^{2m}]$ が存在する為の必要十分条件は

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

であり, このとき $2m$ 次のモーメントは次の様に表せる:

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

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.

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

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.

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

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,

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

従って, 時点 $t-2$ までの情報を与えると,

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

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

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

ここで $m \times m$ 上三角行列 \mathbf{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_* の添え字の時点の後退させると以下の式が成り立つ：

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

これは $k \rightarrow \infty$ まで考えると、情報が無い下での期待値、即ち $\mathbb{E}[\mathbf{w}_t]$ を表すことになり、 $\mathbb{E}[h_t^m]$ が $\mathbb{E}[\mathbf{w}_t]$ 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 $\mathbb{E}[e_t^m]$ の存在の必要十分条件である：

ところで行列 \mathbf{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 \mathbf{A} のスペクトルノルムの定義は次の様に表される：

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

定義に基づけば

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

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

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

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、

$$\|I + C + C^2 + \cdots\| \leq \|I\| + \|C\| + \|C^2\| + \cdots,$$

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

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n_1, n_2 \geq N, \\ \left\| \sum_{i=n_1}^{n_2} C^i \right\| \leq \sum_{i=n_1}^{n_2} \|C\|^i < \epsilon.$$

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：

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

ところで、 $\|C\|$ の値は $C^T C$ の最大固有値に等しい．なぜならば、 C が上三角行列であることから $C^T C$ は実対称行列になる為である． $C^T C$ を対角化すると、或る直交行列 U によって

$$C^T C = U^T \Lambda U, \\ s.t. \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m), \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m,$$

と表すことが出来る．即ち、 $\|C\| = \sqrt{\lambda_1}$. C の対角要素の総乗の 2 乗が $C^T C$ の行列式の値に等しく、 $\|C\| < 1$ ならば $|C^T C| = |\Lambda|$

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り、 $\mu(\alpha_1, \beta_1, m) < 1$ 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 $E[\epsilon_t^{2m}]$ の存在の十分条件となる．一方、 $E[\epsilon_t^{2m}]$ 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

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

を示す．

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

が成り立っていることにより、

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

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

$E[\epsilon_t^{2m}]$ 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，

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

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

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

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：

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

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．

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

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：

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

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 $E[\epsilon_t^2] = 0$,
 $E[h_t(\eta_t^2 - 1)] = E[h_t]E[\eta_t^2 - 1] = 0$, $\text{Cov}[\epsilon_t^2, \epsilon_{t-n}^2] = E[\epsilon_t^2 \epsilon_{t-n}^2]$ が成り立つので、共分散は

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

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。

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

$\gamma_0^{-1} = E[\epsilon_t^4] < \infty$ 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

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

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 $\rho_p, \rho_{p-1}, \rho_{p-2}, \dots, \rho_{p+1-m}$ 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：

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

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、

$$\epsilon_t = y_t - \mathbf{x}_t^T \mathbf{b}$$

と置くと、推定するパラメータの個数は $\dim(\mathbf{b}) + \dim(\boldsymbol{\omega})$ 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

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

とし、真のパラメータ θ_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 $\hat{\theta} \in \Theta$ 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

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

ω に関して微分すると以下の様になる：

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

ここで、 $\frac{\partial h_t}{\partial \omega}$ は次のように示される。

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

$(0 \leq) \sum_{i=1}^p \beta_i < 1$ であるなら $\frac{\partial h_t}{\partial \omega} = z_t + \sum_{i=1}^p \beta_i \frac{\partial h_{t-i}}{\partial \omega}$ は定常となる。

ここで *Fisher* 情報行列

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

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

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

この為、 *Fisher* 情報行列 $I(\theta)$ の ω に対応する部分は $\frac{\partial^2 l_t}{\partial \omega \partial \omega^T}$ 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 ω についてみていたが、次に b について微分する：

$$b = (b_1, b_2, \dots)^T, \quad x_t = (x_{t,1}, x_{t,2}, \dots)^T \text{と表して,}$$

$$\begin{aligned}
\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} - 1\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} - 1\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] \quad \text{と計算できるから,} \\
\frac{\partial l_t(\boldsymbol{\theta})}{\partial \mathbf{b}} &= \left(\frac{\partial l_t}{\partial b_1}, \frac{\partial l_t}{\partial b_2}, \dots \right)^T \\
&= \epsilon_t \mathbf{x}_t h_t^{-1} + \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \mathbf{b}} \left(\frac{\epsilon_t^2}{h_t} - 1\right), \\
\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \mathbf{b}^T \partial \mathbf{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} & \dots \\ \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} & \dots \\ \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} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
&= -h_t^{-1} \mathbf{x}_t \mathbf{x}_t^T - \frac{1}{2} h_t^{-2} \frac{\partial h_t}{\partial \mathbf{b}} \frac{\partial h_t}{\partial \mathbf{b}^T} \left(\frac{\epsilon_t^2}{h_t} - 1\right) \\
&\quad - \epsilon_t h_t^{-2} \frac{\partial h_t}{\partial \mathbf{b}} \mathbf{x}_t^T - \epsilon_t h_t^{-2} \mathbf{x}_t \frac{\partial h_t}{\partial \mathbf{b}^T} + \left(\frac{\epsilon_t^2}{h_t} - 1\right) \frac{\partial}{\partial \mathbf{b}} \left[\frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \mathbf{b}^T} \right].
\end{aligned}$$

ただし,

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

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

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

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 $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}^T \partial \mathbf{b}}$ と $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \mathbf{b}^T \partial \boldsymbol{\omega}}$ を計算する:

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

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

ここで *Fisher* 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:

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