4. Model Selection

4.0. Testing for dependence and ARCH effect

Recall

$$Y_t = \mu_t + \eta_t \sqrt{h_t}.$$

Given a data set $\{Y_t\}$, if $\mu_t = \mu$–a constant, we do not need to fit an ARMA, TAR or Bilinear models to data set.

Assume $h_t = \sigma^2 < \infty$. Then the ACF of $\{Y_t\}$ is

$$\rho_k = 0,$$

for all $k \neq 0$. The sample ACF is

$$\widehat{\rho}_k = \frac{\sum_{t=1}^{n-k}(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2},$$

where $\bar{Y} = \sum_{t=1}^{n} Y_t / n$.

By the SLL theorem,

$$\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t = \mu + o(1).$$

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n}(Y_t - \bar{Y})^2 &= \frac{1}{n}\sum_{t=1}^{n}(Y_t - \mu)^2 + o(1) \\ &= \sigma^2 + o(1). \end{aligned}$$

For each k,

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}(Y_t-\bar{Y})(Y_{t+k}-\bar{Y})^2 \\
&\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}(Y_t-\mu)(Y_{t+k}-\mu)+o_p(1) \\
&\quad = \frac{\sigma^2}{\sqrt{n}}\sum_{t=1}^{n-k}\eta_t\eta_{t+k}+o_p(1)
\end{aligned}$$

Thus,

$$\begin{aligned}
\sqrt{n}(\widehat{\rho}_k-0) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t\eta_{t+k}+o_p(1) \\
&\to_d N(0,1).
\end{aligned}$$

This can be used to test

$$H_0:\rho_k=0 \text{ v.s. } H_a:\rho_k\neq 0.$$

The 95% confidence interval is $[-2/\sqrt{n}, 2/\sqrt{n}]$.

Joint test for serial correlation or white noise:

$$H_0:\rho_1=\rho_2\cdots=\rho_M$$

$$H_a:\rho_i\neq 0 \text{ for some } 1\leq i\leq M$$

Test statistic:

$$\mathcal{Q}(M) = n(n+2)\sum_{k=1}^{M}\frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(M).$$

It is called Ljung and Box (1978) test.

When $h_t \neq \sigma^2$ is strictly stationary with $Eh_t < \infty$, we still have

$$\rho_k = 0$$

and

$$\hat{\rho}_k \to 0 \text{ as } n \to \infty.$$

Ljung and Box (1978) test is not correct (over reject).

Testing for ARCH effect

Assume

$$Y_t = \mu + \eta_t \sqrt{h_t}.$$

Given a data set $\{Y_t\}$, if $h_t = \sigma^2$–a constant, we do not need to fit a GARCH-type model to data set.

Let $\xi_t = (Y_t - \mu)^2$. Then

$$E\xi_t = Eh_t = \sigma^2.$$

$$\nu_k = E(\xi_t - \sigma^2)(\xi_{t+k} - \sigma^2) = E(\eta_t^2 h_t - \sigma^2)(\eta_{t+k}^2 h_{t+k} - \sigma^2)$$

$$\rho_k = \frac{\nu_k}{\nu_0} = 0$$

if $h_t = \sigma^2$, and it is not 0 if $h_t \neq \sigma^2$.

Denote $\hat{\xi}_t = (Y_t - \bar{Y})^2$. σ^2 is estimated by

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \hat{\xi}_t.$$

ρ_k is estimated by

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (\hat{\xi}_t - \hat{\sigma}^2)(\hat{\xi}_{t+k} - \hat{\sigma}^2)}{\sum_{t=1}^{n} (\hat{\xi}_t - \hat{\sigma}^2)^2}.$$

Note that $\widehat{\sigma}^2 = \sigma^2 + o_p(1)$. We can show that

$$\begin{aligned}
\widehat{\rho}_k &= \frac{\sum_{t=1}^{n-k}(\xi_t - \sigma^2)(\xi_{t+k} - \sigma^2)}{\sum_{t=1}^{n}(\xi_t - \sigma^2)^2} + o_p(\frac{1}{\sqrt{n}}) \\
&= \frac{\sum_{t=1}^{n-k}(\eta_t^2 - 1)(\eta_{t+k}^2 - 1)}{\sum_{t=1}^{n}(\eta_t^2 - 1)^2} + o_p(\frac{1}{\sqrt{n}})
\end{aligned}$$

Thus,

$$\sqrt{n}(\widehat{\rho}_k - 0) \to_d N(0, 1).$$

McLeod and Li (1983):

$$H_0 : \rho_1 = \rho_2 \cdots = \rho_m$$

$$H_a : \rho_i \neq 0 \text{ for some } 1 \leq i \leq m$$

Test statistic:

$$Q^2(m) = n(n-1) \sum_{k=1}^{m} \frac{\widehat{\rho}_k^2}{n-k} \sim \chi^2(m).$$

4.1. Model with a constant variance

- Null model:

  Given the time series $\{Y_1, Y_2, \cdots, Y_n\}$ and the initial value Y_0, we try to use the following model to fit the data:

  $$Y_t = m(\theta, X_{t-1}) + \varepsilon_t,$$

  where $X_t = (Y_t, Y_{t-1}, \cdots)$ and ε_t is i.i.d. with mean zero and variance σ^2.

  Usually, we use the LSE method, i.e., look for the minimizer of the following objective function:

  $$L_n(\theta) = \sum_{t=1}^{n} \left[Y_t - m(\theta, X_{t-1}) \right]^2.$$

  Denote

  $$\widehat{\theta}_n = \arg\min_{\Theta} L_n(\theta).$$

$$
\begin{aligned}
D_t(\theta) &= -2\left[Y_t - m(\theta, X_{t-1})\right]\frac{\partial m(\theta, X_{t-1})}{\partial \theta}, \\
P_t(\theta) &= 2\frac{\partial m(\theta, X_{t-1})}{\partial \theta}\frac{\partial m(\theta, X_{t-1})}{\partial \theta'} \\
&\qquad\qquad -2\left[Y_t - m(\theta, X_{t-1})\right]\frac{\partial^2 m(\theta, X_{t-1})}{\partial \theta \partial \theta'}, \\
\Omega &= E\left[D_t(\theta_0) D_t'(\theta_0)\right] \\
&= 4\sigma^2 E\left[\frac{\partial m(\theta, X_{t-1})}{\partial \theta}\frac{\partial m(\theta, X_{t-1})}{\partial \theta'}\right], \\
\Sigma &= \frac{1}{2} E P_t(\theta_0) = \frac{\Omega}{4\sigma^2}.
\end{aligned}
$$

Assume that the conditions of Theorem 3.1 are satisfied. We can have the following expansion:

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = -\frac{\Sigma^{-1}}{2\sqrt{n}}\sum_{t=1}^{n} D_t(\theta_0) + o_p(1).$$

After we estimate the model, we have the residuals:

$$\widehat{\varepsilon}_t = \varepsilon_t(\widehat{\theta}_n) = Y_t - m(\widehat{\theta}_n, X_{t-1}).$$

The basic idea is that if the model is correct, then $\widehat{\varepsilon}_t$ should be very close to ε_t. We know that the noise ε_t is i.i.d.. Then, $\widehat{\varepsilon}_t$ should be almost i.i.d. or at least it should be uncorrelated. Thus, a simple way is to plot out the sample ACF of $\widehat{\varepsilon}_t$ and see if they are close to zero. However, how close is close? We need to qualify this difference. First, we see the following expansion:

$$
\begin{aligned}
\widehat{\varepsilon}_t &= Y_t - m(\widehat{\theta}_n, X_{t-1}) \\
&= \varepsilon_t + m(\theta_0, X_{t-1}) - m(\widehat{\theta}_n, X_{t-1}) \\
&= \varepsilon_t - (\widehat{\theta}_n - \theta_0)' \frac{\partial m(\theta_0, X_{t-1})}{\partial \theta} \\
&\quad -(\widehat{\theta}_n - \theta_0)' \frac{\partial^2 m(\xi^*, X_{t-1})}{\partial \theta \partial \theta'} (\widehat{\theta}_n - \theta_0) \\
&= \varepsilon_t - (\widehat{\theta}_n - \theta_0)' \frac{\partial m(\theta_0, X_{t-1})}{\partial \theta} + O_p\left(\frac{1}{n}\right).
\end{aligned}
$$

The sample ACF of $\widehat{\varepsilon}_t$ with the lag k is

$$
\widehat{\rho}_k = \frac{\sum_{t=k}^{n} \widehat{\varepsilon}_{t-k} \widehat{\varepsilon}_t}{\sum_{t=1}^{n} \widehat{\varepsilon}_t^2}.
$$

Note that

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\widehat{\varepsilon}_t^2 &= \frac{1}{n}\sum_{t=1}^{n}\left[\varepsilon_t - (\widehat{\theta}_n-\theta_0)'\frac{\partial m(\theta_0, X_{t-1})}{\partial\theta}\right. \\
&\quad \left. + O_p\left(\frac{1}{n}\right)\right]^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2 - \frac{2}{n}(\widehat{\theta}_n-\theta_0)'\sum_{t=1}^{n}u_{t-1}\varepsilon_t \\
&\qquad + \frac{1}{n}(\widehat{\theta}_n-\theta_0)'\sum_{t=1}^{n}u_{t-1}u_{t-1}'(\widehat{\theta}_n-\theta_0) \\
&\quad + O_p\left(\frac{1}{n}\right) \\
&= \sigma^2 + o_p(1),
\end{aligned}$$

where $u_{t-1} = \partial m(\theta_0, X_{t-1})/\partial\theta$. Thus,

$$\sqrt{n}\widehat{\rho}_k = \frac{1}{\sigma^2 + o_p(1)}\frac{1}{\sqrt{n}}\sum_{t=k}^{n}\widehat{\varepsilon}_{t-k}\widehat{\varepsilon}_t,$$

and

$$
\begin{aligned}
\sum_{t=k}^{n} \widehat{\varepsilon}_{t-k}\widehat{\varepsilon}_t &= \sum_{t=k}^{n} \varepsilon_{t-k}\varepsilon_t - (\widehat{\theta}_n - \theta_0)' \sum_{t=k}^{n} u_{t-k-1}\varepsilon_t \\
&\qquad -(\widehat{\theta}_n - \theta_0)' \sum_{t=k}^{n} u_{t-1}\varepsilon_{t-k} \\
&\qquad +(\widehat{\theta}_n - \theta_0)' \sum_{t=k}^{n} u_{t-k-1}u'_{t-1}(\widehat{\theta}_n - \theta_0) \\
&\quad +O_p(1) \\
&= \sum_{t=k}^{n} \varepsilon_{t-k}\varepsilon_t - n(\widehat{\theta}_n - \theta_0)' A_k + O_p(1),
\end{aligned}
$$

where $A_k = E\left(u_{t-1}\varepsilon_{t-k}\right)$.

Thus,

$$
\sigma^2 \sqrt{n}\widehat{\rho}_k = \frac{1}{\sqrt{n}} \sum_{t=k}^{n} \varepsilon_{t-k}\varepsilon_t - \sqrt{n}(\widehat{\theta}_n - \theta_0)' A_k + o_p(1).
$$

Usually, we need to check the lag of ACF up to M. Thus, we consider the vector as follows:

$$\sigma^2\sqrt{n}\begin{pmatrix}\widehat{\rho}_1\\ \widehat{\rho}_2\\ \cdot\\ \cdot\\ \widehat{\rho}_M\end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=M}^{n}\begin{pmatrix}\varepsilon_{t-1}\\ \varepsilon_{t-2}\\ \cdot\\ \cdot\\ \varepsilon_{t-M}\end{pmatrix}\varepsilon_t - \begin{pmatrix}A_1'\\ A_2'\\ \cdot\\ \cdot\\ A_M'\end{pmatrix}\sqrt{n}(\widehat{\theta}_n-\theta_0)+o_p(1)$$

$$= \frac{1}{\sqrt{n}}\sum_{t=M}^{n}\left[\begin{pmatrix}\varepsilon_{t-1}\\ \varepsilon_{t-2}\\ \cdot\\ \cdot\\ \varepsilon_{t-M}\end{pmatrix} - \begin{pmatrix}A_1'\\ A_2'\\ \cdot\\ \cdot\\ A_M'\end{pmatrix}\Sigma^{-1}u_{t-1}\right]\varepsilon_t$$

$$+o_p(1)$$

$$\equiv \frac{1}{\sqrt{n}}\sum_{t=M}^{n}R_{t-1}\varepsilon_t+o_p(1)$$

$$\longrightarrow^d N\left[0,\sigma^2E\left(R_{t-1}R_{t-1}'\right)\right],$$

as $n\to\infty$, by the central limiting theorem, where

$$E\left(R_{t-1}R_{t-1}'\right)=\sigma^2I_M-A'\Sigma^{-1}A,$$

with $A=[A_1,A_2,\cdots,A_M]$.

It follows that

$$\sqrt{n}\begin{pmatrix} \widehat{\rho}_1 \\ \widehat{\rho}_2 \\ \cdot \\ \cdot \\ \widehat{\rho}_M \end{pmatrix} \to^d N\left[0, I_M - \frac{A'\Sigma^{-1}A}{\sigma^2}\right],$$

as $n \to \infty$. Thus,

$$\begin{aligned} \mathcal{Q} &= n(\widehat{\rho}_1, \cdots, \widehat{\rho}_M)\left[I_M - A'\Sigma^{-1}A/\sigma^2\right]^{-1} \\ &\quad (\widehat{\rho}_1, \cdots, \widehat{\rho}_M)' \to \chi^2(M), \end{aligned}$$

as $n \to \infty$.

$\mathcal{Q}$ is called the portmanteau test.

Example: AR(1) model,

$$Y_t = \phi Y_{t-1} + \varepsilon_t.$$

In this case,

$$\begin{aligned} u_{t-1} &= Y_{t-1} = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-1-i}, \\ A_k &= E\left(u_{t-1}\varepsilon_{t-k}\right) = \sigma^2 \phi^{k-1}, \\ \Sigma &= E u_{t-1}^2 = \sigma^2 \sum_{i=0}^{\infty} \phi^{2i} = \frac{\sigma^2}{1-\phi^2}. \end{aligned}$$

Thus,

$$\begin{aligned} A'\Sigma^{-1}A &= \sigma^2(1-\phi^2) \begin{pmatrix} 1 & \phi & \cdot & \cdot & \phi^{M-1} \\ \phi & \phi^2 & \cdot & \cdot & \phi^M \\ \cdot & & & & \\ \cdot & & & & \\ \phi^{M-1} & \phi^M & \cdot & \cdot & \phi^{2M-2} \end{pmatrix} \\ &= \sigma^2(1-\phi^2) C' \begin{pmatrix} 1 & 0 & \cdot & \cdot & 0 \\ 0 & \phi^2 & \cdot & \cdot & 0 \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & \cdot & \cdot & \phi^{2M-2} \end{pmatrix} C \end{aligned}$$

where C is an orthogonal matrix.

It follows that

$$\begin{aligned} I_M - A'\Sigma^{-1}A/\sigma^2 &= Q' diag\{\phi^2, 1-\phi^2+\phi^4, \\ &\quad \cdots, 1-\phi^{2M-2}+\phi^{2M}\} Q. \end{aligned}$$

Note that ϕ^{2k} is very small as k is large. To make it easy, Box and Pierce (1970) considered the following modified test

$$\mathcal{Q}_1(M) = n \sum_{k=1}^{M} \widehat{\rho}_k^2 \sim \chi^2(M-1).$$

In general, we have

$$\mathcal{Q}_1(M) \to^d \chi^2_{M-p}, \text{ as } n \to \infty,$$

where p is the number of estimated parameters.

A further modification is Ljung and Box (1978) test as follows:

$$\mathcal{Q}_2(M) = n(n+2) \sum_{k=1}^{M} \frac{\widehat{\rho}_k^2}{n-k} \sim \chi^2(M-p).$$

4.2. Models with a zero mean

Given the time series $\varepsilon_1, \cdots, \varepsilon_n$, and the initial value ε_0, we try to fit the data by the model:

$$\begin{cases} \varepsilon_t = \eta_t \sqrt{h_t} \\ h_t = h(\theta, X_{t-1}), \end{cases} \tag{1}$$

where η_t is i.i.d. with zero mean and variance 1. We use the QMLE to estimate the model, i.e., find the minimizer of the quasi-likelihood function:

$$L_n(\theta) = -\frac{1}{2} \sum_{t=1}^{n} \left[\log h_t(\theta) + \frac{\varepsilon_t^2}{h_t(\theta)} \right],$$

where $h_t(\theta) = h(\theta, X_{t-1})$. Let $D_t(\theta)$ and $P_t(\theta)$ be defined as before with h_t being defined in the general form (1). Assume that the conditions in Theorem 3.1 are satisfied.

We have the expansion:

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = \frac{\Sigma^{-1}}{\sqrt{n}} \sum_{t=1}^{n} D_t(\theta_0) + o_p(1),$$

where

$$D_t(\theta_0) = \tfrac{1}{2h_t} \tfrac{\partial h_t(\theta_0)}{\partial \theta} (\eta_t^2 - 1),$$

$$\Sigma = E \left[\tfrac{1}{2h_t^2} \tfrac{\partial h_t(\theta_0)}{\partial \theta} \tfrac{\partial h_t(\theta_0)}{\partial \theta'} \right].$$

Now, the scaled residual, denoted by $\widehat{\eta}_t$, is defined as

$$\widehat{\eta}_t = \frac{\varepsilon_t}{\sqrt{h_t(\widehat{\theta}_n)}}.$$

Using Taylor's expansion, we have

$$\widehat{\eta}_t = \eta_t - \frac{\eta_t}{2h_t}\frac{\partial h_t(\theta_0)}{\partial \theta'}(\widehat{\theta}_n - \theta_0) + O_p\left(\frac{1}{n}\right) R_t(\xi_n^*),$$

where

$$R_t(\theta) = \frac{3\varepsilon_t}{4h_t^{5/2}(\theta)}\frac{\partial h_t(\theta)}{\partial \theta}\frac{\partial h_t(\theta)}{\partial \theta'} + \frac{\varepsilon_t}{2h_t^{3/2}(\theta)}\frac{\partial^2 h_t(\theta)}{\partial \theta \partial \theta'}.$$

Let

$$\widehat{\rho}_k = \frac{\sum_{t=k}^{n} \widehat{\eta}_{t-k}\widehat{\eta}_t}{\sum_{t=1}^{n} \widehat{\eta}_t^2}.$$

We can show that

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} \widehat{\eta}_t^2 &= \frac{1}{n}\sum_{t=1}^{n} \eta_t^2 + o_p(1) = 1 + o_p(1), \\
\frac{1}{\sqrt{n}}\sum_{t=k}^{n} \widehat{\eta}_{t-k}\widehat{\eta}_t &= \frac{1}{\sqrt{n}}\sum_{t=k}^{n} \eta_{t-k}\eta_t - A_k'\sqrt{n}(\widehat{\theta}_n - \theta_0) \\
&\quad + o_p(1),
\end{aligned}$$

where

$$A_k = E\left(\frac{1}{2h_t}\frac{\partial h_t(\theta_0)}{\partial \theta}\eta_{t-k}\eta_t\right).$$

Assume that $E\eta_t^3 = 0$. Then,

$$
\begin{aligned}
\sqrt{n}\begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \cdot \\ \cdot \\ \hat{\rho}_M \end{pmatrix} &= \frac{1}{\sqrt{n}} \sum_{t=M}^{n} \begin{pmatrix} \eta_{t-1} \\ \eta_{t-2} \\ \cdot \\ \cdot \\ \eta_{t-M} \end{pmatrix} \eta_t \\
&\quad - \begin{pmatrix} A'_1 \\ A'_2 \\ \cdot \\ \cdot \\ A'_M \end{pmatrix} \sqrt{n}(\hat{\theta}_n - \theta_0) + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{t=M}^{n} \left[\begin{pmatrix} \eta_{t-1} \\ \eta_{t-2} \\ \cdot \\ \cdot \\ \eta_{t-M} \end{pmatrix} \eta_t - \begin{pmatrix} A'_1 \\ A'_2 \\ \cdot \\ \cdot \\ A'_M \end{pmatrix} \right. \\
&\qquad \left. \frac{\Sigma^{-1}}{2h_t} \frac{\partial h_t(\theta_0)}{\partial \theta} (\eta_t^2 - 1) \right] + o_p(1) \\
&\to^d \; N\left[0, I_M - \frac{\kappa}{2} A' \Sigma^{-1} A\right],
\end{aligned}
$$

as $n \to \infty$, by the central limiting theorem, where

$$A = [A_1, A_2, \cdots, A_M] \text{ and } \kappa = E\eta_t^4 - 1.$$

Thus, we have

$$\begin{aligned}
\mathcal{Q}(M) &= n(\widehat{\rho}_1, \cdots, \widehat{\rho}_M) \\
&\quad \left[I_M - \frac{\kappa}{2} A' \Sigma^{-1} A \right]^{-1} (\widehat{\rho}_1, \cdots, \widehat{\rho}_M)' \\
&\quad \to^d \chi^2(M),
\end{aligned}$$

as $n \to \infty$. The simple modification as the Ljung-Box test is

$$\mathcal{Q}(M) = n(n+2) \sum_{k=1}^{M} \frac{\widehat{\rho}_k^2}{n-k}.$$

Li-Mak test:

$$\hat{r}_k = \frac{\sum_{t=k}^{n}(\hat{\eta}_{t-k}^2 - 1)(\hat{\eta}_t^2 - 1)}{\sum_{t=1}^{n}(\hat{\eta}_t^2 - 1)^2}.$$

where

$$\frac{1}{n}\sum_{t=1}^{n}(\hat{\eta}_t^2 - 1)^2 = E\eta_t^4 - 1 + o_p(1) \approx \kappa + o_p(1).$$

By Taylor's expansion, we have

$$\begin{aligned}
&(\hat{\eta}_{t-k}^2 - 1)(\hat{\eta}_t^2 - 1) \\
&\quad = (\eta_{t-k}^2 - 1)(\eta_t^2 - 1) \\
&\qquad -(\eta_{t-k}^2 - 1)\eta_t^2 \frac{1}{h_t}\frac{\partial h_t(\theta_0)}{\partial \theta}(\hat{\theta}_n - \theta_0)' \\
&\qquad -(\eta_t^2 - 1)\eta_{t-k}^2 \frac{1}{h_{t-k}}\frac{\partial h_{t-k}(\theta_0)}{\partial \theta}(\hat{\theta}_n - \theta_0)' \\
&\quad +O_p\left(\frac{1}{n}\right) R_t(\xi_n^*).
\end{aligned}$$

Moreover,

$$\frac{1}{\sqrt{n}}\sum_{t=k}^{n}(\hat{\eta}_{t-k}^2-1)(\hat{\eta}_t^2-1)$$
$$=\frac{1}{\sqrt{n}}\sum_{t=k}^{n}(\eta_{t-k}^2-1)(\eta_t^2-1)$$
$$-\frac{1}{\sqrt{n}}\sum_{t=k}^{n}(\eta_{t-k}^2-1)\eta_t^2\frac{1}{h_t}\frac{\partial h_t(\theta_0)}{\partial\theta}(\hat{\theta}_n-\theta_0)'$$
$$+o_p(1)$$
$$=\frac{1}{\sqrt{n}}\sum_{t=k}^{n}(\eta_{t-k}^2-1)(\eta_t^2-1)-B_k'\sqrt{n}(\hat{\theta}_n-\theta_0)$$
$$+o_p(1),$$

where

$$B_k=E\left[(\eta_{t-k}^2-1)\frac{1}{h_t}\frac{\partial h_t(\theta_0)}{\partial\theta}\right].$$

$$\frac{1}{\sqrt{n}}\sum_{t=k}^{n}(\hat{\eta}_{t-k}^2-1)(\hat{\eta}_t^2-1)$$
$$=\frac{1}{\sqrt{n}}\sum_{t=k}^{n}\Big[(\eta_{t-k}^2-1)$$
$$-B_k'\frac{\Sigma^{-1}}{2h_t}\frac{\partial h_t(\theta_0)}{\partial\theta}\Big](\eta_t^2-1)+o_p(1)$$
$$\equiv\frac{1}{\sqrt{n}}\sum_{t=k}^{n}\tilde{R}_{kt}'(\eta_t^2-1)+o_p(1).$$

Hence,

$$\sqrt{n}\begin{pmatrix} \hat{r}_1 \\ \hat{r}_2 \\ \cdot \\ \cdot \\ \hat{r}_M \end{pmatrix} = \frac{1}{\sqrt{n}}\sum_{t=M}^{n}\begin{pmatrix} \tilde{R}'_{1t} \\ \tilde{R}'_{2t} \\ \cdot \\ \cdot \\ \tilde{R}'_{Mt} \end{pmatrix}(\eta_t^2-1)+o_p(1)$$
$$\to^d N(0,\kappa EB_t)\,,\ \text{as } n\to\infty,$$

where

$$\begin{aligned} EB_t &= E\left[(\tilde{R}_{1t},\cdots,\tilde{R}_{Mt})'(\tilde{R}_{1t},\cdots,\tilde{R}_{Mt})\right] \\ &= \kappa I_M - \frac{1}{2}B'\Sigma^{-1}B, \end{aligned}$$

with $B=[B_1,B_2,\cdots,B_M]$. It follows that

$$\begin{aligned} \mathcal{Q}^2(M) &= n(\hat{r}_1,\cdots,\hat{r}_M) \\ &\quad \left[I_M-\frac{1}{2\kappa}B'\Sigma^{-1}B\right]^{-1}(\hat{r}_1,\cdots,\hat{r}_M)' \\ &\to \chi^2(M), \end{aligned}$$

as $n\to\infty$. Especially, when $\eta_t\sim N(0,1)$, $\kappa=3$.

Li and Make (1994) modified $\mathcal{Q}^2(M)$ as follows:

$$\mathcal{Q}^2(M) = n(n+2)\sum_{k=1}^{M}\frac{\hat{r}_k^2}{n-k}\sim\chi^2(M-p).$$

In Sas, it is called McLeod-Li test:

$$\mathcal{Q}^2(M) \; = \; n(n+2)\sum_{k=1}^{M}\frac{\widehat{r}_k^2}{n-k} \sim \chi^2(M).$$