1 Define the GARCHX model

1.1 Mean Model

$$r_t = \mu + \sigma_t z_t \tag{1}$$
$$z_t \sim i.i.d\ N(0,1) \tag{2}$$

where $e_t = \sigma_t z_t$.

1.2 Volatility Model

We introduce an exogenous term x_t^2 to the traditional GARCH model. The squared time series ensures non-negativity of σ_t^2 at any time step.

$$\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma x_t^2 \tag{3}$$

2 Log Likelihood

For simplicity, assume $\mu = 0$. Let $\theta = (\omega, \alpha, \beta, \gamma)$. Then the log likelihood function is defined as a function of the conditional densities of r_t as such:

$$l(\theta) = \sum_{t=1}^{T} \frac{1}{2} \left(-\log 2\pi - \log \sigma_t^2 - \frac{e_t^2}{\sigma_t^2} \right) \tag{4}$$

The first and second partial derivatives of $l(\theta)$ are as follows:

$$\frac{\partial}{\partial \theta} l(\theta) = \sum_{t=1}^{T} \frac{1}{2} \left(\frac{e_t^2}{\sigma_t^4} - \frac{1}{\sigma_t^2} \right) \frac{\partial \sigma_t^2}{\partial \theta} \tag{5}$$

$$\frac{\partial^2}{\partial \theta_1 \partial \theta_2} l(\theta) = -\frac{1}{2} \sum_{t=1}^{T} \left(\frac{\partial^2 \sigma_t^2}{\partial \theta_1 \partial \theta_2} \left(\frac{1}{\sigma_t^2} - \frac{e_t^2}{\sigma_t^4} \right) + \frac{\partial \sigma_t^2}{\partial \theta_1} \frac{\partial \sigma_t^2}{\partial \theta_2} \left(\frac{2e_t^2}{\sigma_t^6} - \frac{1}{\sigma_t^4} \right) \right) \tag{6}$$

3 Partial Derivatives of σ^2

This section merely serves to supplement the previous section, as the first and second partial derivatives of σ^2 are used in the score funtion and information matrix.

3.1 First Partial Derivative

$$\frac{\partial\sigma_t^2}{\partial\omega} = 1 + \beta\frac{\partial\sigma_{t-1}^2}{\partial\omega} \tag{7}$$

$$\frac{\partial\sigma_t^2}{\partial\alpha} = e_{t-1}^2 + \beta\frac{\partial\sigma_{t-1}^2}{\partial\alpha} \tag{8}$$

$$\frac{\partial\sigma_t^2}{\partial\beta} = \sigma_{t-1}^2 + \beta\frac{\partial\sigma_{t-1}^2}{\partial\beta} \tag{9}$$

$$\frac{\partial\sigma_t^2}{\partial\gamma} = x_{t-1}^2 + \beta\frac{\partial\sigma_{t-1}^2}{\partial\gamma} \tag{10}$$

3.2 Unconditional Expectation of First Derivatives of σ^2

We first derive the unconditional expectation of σ^2 with respect to α. The same logic is then applied to the other parameters. We have shown that $\mathbb{E}(e_t^2) = \mathbb{E}(\sigma_t^2) = \frac{\omega+\gamma k}{1-(\alpha+\beta)}$ for all t, where $k = \mathbb{E}(x^2)$. Let $\frac{\omega+\gamma k}{1-(\alpha+\beta)} = C$ for brevity. Then when $\beta < 1$,

$$\begin{aligned}
\frac{\partial\sigma_t^2}{\partial\alpha} &= e_{t-1}^2 + \beta\frac{\partial\sigma_{t-1}^2}{\partial\alpha} \\
\mathbb{E}\left[\frac{\partial\sigma_t^2}{\partial\alpha}\right] &= C + \beta\,\mathbb{E}\left[\frac{\partial\sigma_{t-1}^2}{\partial\alpha}\right] \\
&= C + \beta\left(C + \beta\,\mathbb{E}\left[\frac{\partial\sigma_{t-2}^2}{\partial\alpha}\right]\right) \\
&= C + \beta C + \beta^2\left(C + \beta\,\mathbb{E}\left[\frac{\partial\sigma_{t-3}^2}{\partial\alpha}\right]\right) \\
&= \sum_{i=0}^{\infty} C\beta^i \\
&= \frac{C}{1-\beta} \\
&= \frac{\omega+\gamma k}{(1-\beta)(1-(\alpha+\beta))}
\end{aligned}$$

As long as $\beta < 1$, the unconditional expectation of first derivatives exist. Using the same logic, the first partial derivatives are as follows:

$$E\left[\frac{\partial\sigma_t^2}{\partial\omega}\right] = \frac{1}{1-\beta} \tag{11}$$
$$E\left[\frac{\partial\sigma_t^2}{\partial\alpha}\right] = \frac{\omega+\gamma k}{(1-\beta)(1-\alpha-\beta)} \tag{12}$$
$$E\left[\frac{\partial\sigma_t^2}{\partial\beta}\right] = \frac{\omega+\gamma k}{(1-\beta)(1-\alpha-\beta)} \tag{13}$$
$$E\left[\frac{\partial\sigma_t^2}{\partial\gamma}\right] = \frac{k}{1-\beta} \tag{14}$$

3.3 Second Derivative to σ^2

There exist two cases.

Case a: If both parameters θ_i, θ_j are not β

$$\frac{\partial^2\sigma_t^2}{\partial\theta_i\theta_j} = \beta\frac{\partial^2\sigma_{t-1}^2}{\partial\theta_i\theta_j} \tag{15}$$

with expectation:

$$E\left[\frac{\partial^2\sigma_t^2}{\partial\theta_i\partial\theta_j}\right] = 0 \tag{16}$$

Case b: If at least one of the parameters θ_i, θ_j is β

$$\frac{\partial^2\sigma_t^2}{\partial\beta\theta_j} = \frac{\partial\sigma_{t-1}^2}{\partial\theta_j} + \beta\frac{\partial^2\sigma_{t-1}^2}{\partial\beta\theta_j} \tag{17}$$

4 Stationarity of Time Series

4.1 Stationarity of σ_t^2

We know that $e_t = \sigma_t z_t$, since $z_t \sim i.i.d\ N(0,1)$,

$$\begin{aligned}
E[e_t^2] &= E[\sigma_t^2 z_t^2] \\
&= E[\sigma_t^2]E[z_t^2] \\
&= E[\sigma_t^2]
\end{aligned}$$

Note that $E[z_t^2]$ is the expectation of a chi-squared distribution with 1 degree of freedom.

Using the fact that $E[e_t^2] = E[\sigma_t^2]$,

$$\begin{aligned}
E[\sigma_t^2] &= E[\omega + \alpha\sigma_{t-1}^2 + \beta e_{t-1}^2 + \gamma x_t^2] \\
&= E[\omega] + \alpha E[\sigma_{t-1}^2] + \beta E[e_{t-1}^2] + \gamma E[x_t^2] \\
&= \omega + \gamma E[x_t^2] + \alpha E[\sigma_{t-1}^2] + \beta E[e_{t-1}^2] \\
&= \omega + \gamma k + (\alpha+\beta) E[\sigma_{t-1}^2] \\
&= \omega + \gamma k + (\alpha+\beta) E[\omega + \alpha\sigma_{t-2}^2 + \beta e_{t-2}^2 + \gamma x_{t-1}^2] \\
&= \dots \\
&= (\omega + \gamma k) + [1 + (\alpha+\beta) + (\alpha+\beta)^2 + (\alpha+\beta)^3 + \dots] \\
&= \frac{\omega + \gamma k}{1 - (\alpha+\beta)}
\end{aligned}$$

where $k = E[x_t^2]$, which is a constant, given that we assume x_t^2 is stationary. From this, we know that $|\alpha + \beta| < 1$.
Since x_t^2 is stationary, then $E[x_t^2, x_{t-s}^2]$ is not dependent on t.
Following that x_t^2 is stationary, we prove that

$$\begin{aligned}
Cov[\sigma_t^2, \sigma_{t-s}^2] &= E[\sigma_t^2\sigma_{t-s}^2] - E[\sigma_t^2]E[\sigma_{t-s}^2] \\
&= E[\sigma_t^2\sigma_{t-s}^2] - \left(\frac{\omega + \gamma k}{1 - (\alpha+\beta)}\right)^2
\end{aligned}$$

We found

$$\begin{aligned}
E[\sigma_t^2\sigma_{t-s}^2] &= E[(\omega + \alpha\sigma_{t-1}^2 + \beta e_{t-1}^2 + \gamma x_t^2)(\omega + \alpha\sigma_{t-2}^2 + \beta e_{t-2}^2 + \gamma x_{t-1}^2)] \\
&= \dots \\
&= (\omega^2 + 2\gamma\omega k) + (2\alpha\omega + 2\beta\omega + 2\alpha\gamma k + 2\beta\gamma k)\left(\frac{\omega + \gamma k}{1 - (\alpha+\beta)}\right) \\
&\quad + (\alpha+\beta)^2 E[\sigma_{t-1}^2\sigma_{t-s-1}^2] + \gamma^2 E[x_t^2 x_{t-s}^2]
\end{aligned}$$

Since we know x_t is stationary, $Cov[x_t^2, x_{t-s}^2] = E[x_t^2 x_{t-s}^2] - E[x_t^2]E[x_{t-s}^2]$ does not depend on t, thus we can conclude that $E[x_t^2 x_{t-s}^2]$ does not depend on t.

This shows that $Cov[\sigma_t^2, \sigma_{t-s}^2]$ does not depend on t.

We see that the following is a stationary AR(1) process now as $(\alpha+\beta) < 1 \implies (\alpha+\beta)^2 < 1$:

$$E[\sigma_t^2\sigma_{t-s}^2] = (\omega^2 + 2\gamma\omega k) + 2(\alpha\omega + \beta\omega + \alpha\gamma k + \beta\gamma k)\frac{\omega + \gamma k}{1-(\alpha+\beta)}$$
$$+ (\alpha+\beta)^2 E[\sigma_{t-1}^2\sigma_{t-s-1}^2] + \gamma^2 E[x_t^2 x_{t-s}^2]$$
$$\implies E[\sigma_t^2\sigma_{t-s}^2] \text{ does not depend on } t$$
$$\implies cov(\sigma_t^2, \sigma_{t-s}^2) = E[\sigma_t^2\sigma_{t-s}^2] - \left(\frac{\omega+\gamma k}{1-(\alpha+\beta)}\right)^2 \text{ does not depend on } t$$
$\therefore \{\sigma_t^2\}$ is stationary.

4.2 Stationarity of r_t

First, note that $z_t \sim i.i.d\ N(0,1) \implies E[z_t] = 0$.

To show $E[r_t]$ is constant:

$$\begin{aligned} E[r_t] &= E[\mu + \sigma_t z_t] \\ &= \mu + E[\sigma_t z_t] \\ &= \mu + E[\sigma_t]E[z_t] \\ &= \mu \end{aligned}$$

Next, we show finite variance of r_t:

$$\begin{aligned} Var(r_t) &= E[\sigma_t^2] \\ &= \frac{\omega+\gamma k}{1-(\alpha+\beta)} \end{aligned}$$

Lastly, we show that the autocovariance of r_t is not time dependent:

$$\begin{aligned} Cov(r_t, r_{t-r}) &= Cov(e_t, e_{t-r}) \\ &= E[e_t e_{t-r}] - E[e_t]E[e_{t-r}] \\ &= 0 \end{aligned}$$

4.3 Convergence of e_t

We know that $e_t = \sigma_t z_t$,

and z_t is

$$z_t \sim i.i.d\ N(0,1)$$

Therefore,

$$\begin{aligned} E[e_t|\mathcal{F}_{t-1}] &= E[\sigma_t z_t|\mathcal{F}_{t-1}] \\ &= E[z_t|\mathcal{F}_{t-1}]\ E[\sigma_t|\mathcal{F}_{t-1}] \\ &= 0 \end{aligned}$$

Thus, we conclude that e_t is an MDS sequence.

For simplicity, we assume our time series is ergodic.

By the Martingale CLT, as e_t is a stationary and ergodic Martingale difference,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t \xrightarrow{\text{d}} N(0,\sigma^2)$$

5 Consistency of the MLE

Let the score function be defined as $s(\theta) = \frac{\partial}{\partial\theta} l(\theta)$ and the information $I(\theta) = \frac{\partial^2}{\partial\theta^2} l(\theta)$.

Under suitable regularity conditions, and having shown that the time series is stationary and ergodic,

$$\hat{\theta} \xrightarrow{\text{p}} \theta_0$$

6 Asymptotic normality of the MLE

First, we consider the distribution of the true score $s(\theta_0)$. Under the regularity conditions, it can be shown that $E(s_n(\theta_0)) = 0$ and $Var(s_n(\theta_0)) = I_E(\theta_0)$. Furthermore, as we have shown that r_t converges in distribution by the Martingale CLT, the score function

$$\begin{aligned}\frac{1}{\sqrt{n}} s_n(\theta_0) &= \sum_{t=1}^{n} \frac{\partial}{\partial\theta} \log f(r_t|\theta_0) \\ &\xrightarrow{\text{d}} N(0, I_E(\theta_0))\end{aligned}$$

converges in distribution to a multivariate normal distribution $N(0, I_E(\theta_0))$.

Next, consider a first-order multivariate Taylor expansion of the score function at θ_0 about $\hat{\theta}$:

$$s_n(\theta_0) \approx s_n(\hat{\theta}) - I_n(\theta_0)(\theta_0 - \hat{\theta})$$

Since by definition $s_n(\hat{\theta}) = 0$,

$$\begin{aligned}s_n(\theta_0) &= -I_n(\theta_0)(\theta_0 - \hat{\theta}) \\ (\hat{\theta} - \theta_0) &= I_n(\theta_0)^{-1} s_n(\theta_0) \\ \sqrt{n}(\hat{\theta} - \theta_0) &= \sqrt{n} I_n(\theta_0)^{-1} s_n(\theta_0)\end{aligned}$$

As shown earlier,

$$\frac{1}{\sqrt{n}} s_n(\theta_0) \xrightarrow{\text{d}} N(0, I_E(\theta_0))$$

and by the law of large numbers,

$$\frac{1}{n} I_n(\theta_0) = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial \theta^2} \log(f(x_i|\theta)) \xrightarrow{\text{p}} I_E(\theta_0)$$

Therefore, by Slutsky's theorem,

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{\text{d}} I_E(\theta_0)^{-1} N(0, I_E(\theta_0)) = N(0, I_E(\theta_0)^{-1})$$