Lecture 2: The ARCH Model

Zheng Tian

1 The Volatility of Asset Returns

This lecture focuses on the volatility of asset returns. Here volatility refers to the **conditional variance** of a time series. That is, for a return series $\{r_t\}$, we are now interested in

$$\sigma_t^2 = \text{Var}(r_t \mid F_{t-1})$$

where F_{t-1} is the information set at time $t-1$.

1.1 Characteristics of volatility

In practice, we have observed some stylized facts about volatility, and some volatility models are proposed to characterize them. These properties of volatility include:

1. **There exist volatility clusters.** That is, volatility may be high for certain time periods and low for other periods. Figure 1 shows the daily changes in the log of the NYSE U.S. 100 stock price index. As seen, a cluster of tranquil periods from 2003 to 2007 is followed by a cluster of drastic volatile periods from 2008 to 2010.

2. **Volatility evolves over time in a continuous manner.** That is, volatility jumps are rare. In Figure 2, the change of the volatility of the annualized real GDP growth rate of the U.S. is relatively smooth.

3. **Volatility does not diverge to infinity.** That is, volatility varies within some fixed range. Statistically speaking, this means that volatility is often stationary.

4. Volatility seems to react differently to a big price increase or a big price drop, referred to as **the leverage effect**.

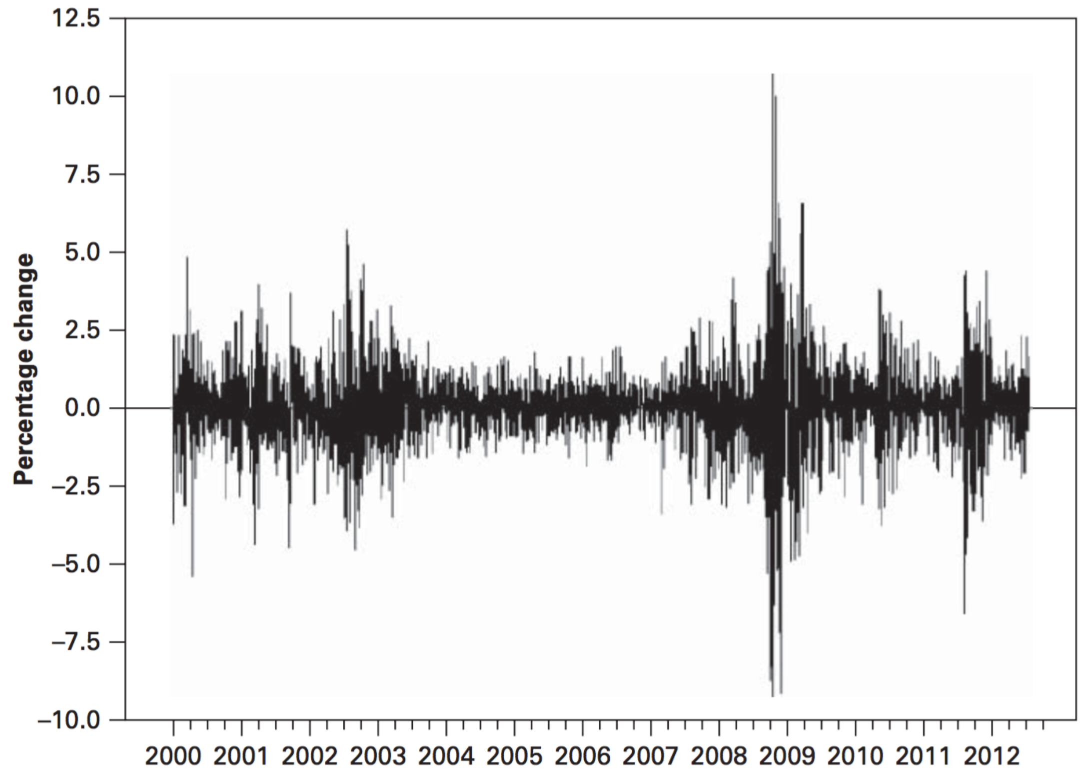

Figure 1: Percentage Change in the NYSE U.S. 100

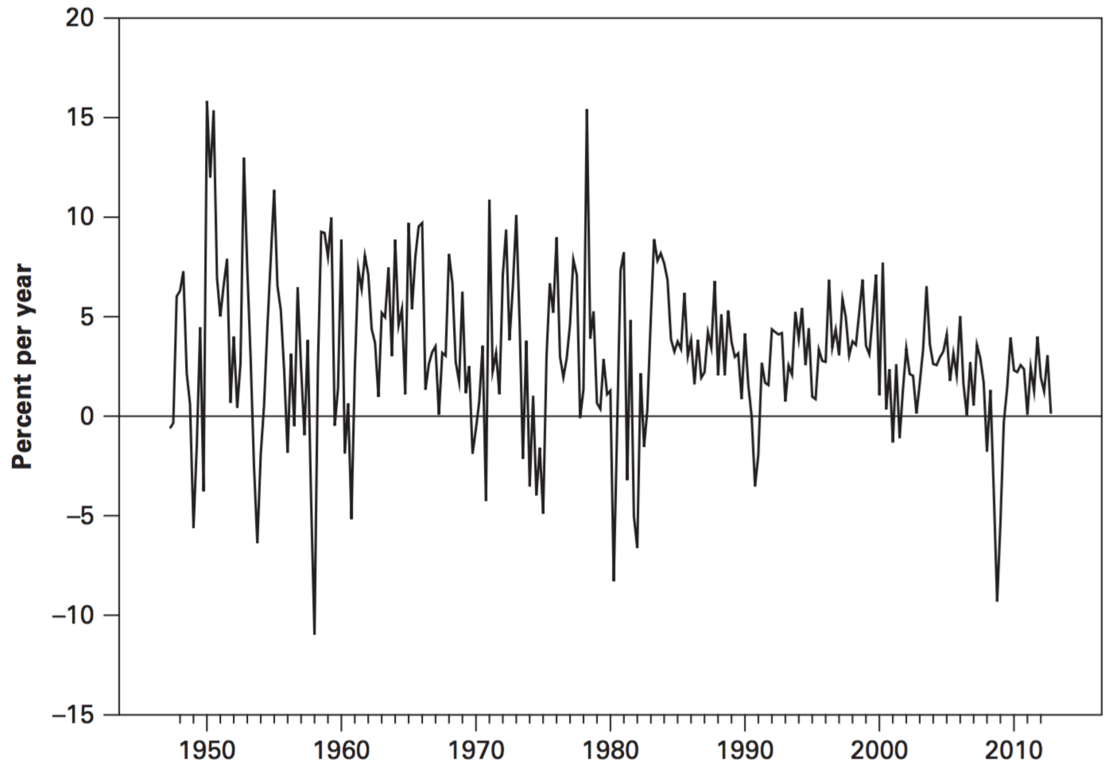

Figure 2: Annualized Growth Rate of Real GDP

2 The Structure of a Volatility Model

2.1 The basic idea of building a volatility model

Consider the log return series $\{r_t\}$. The basic idea of a volatility model is that $\{r_t\}$ may appear to be either serially uncorrelated or serially correlated with a minor order, but $\{r_t\}$ is a dependent series and the dependence arises from its conditional variance.

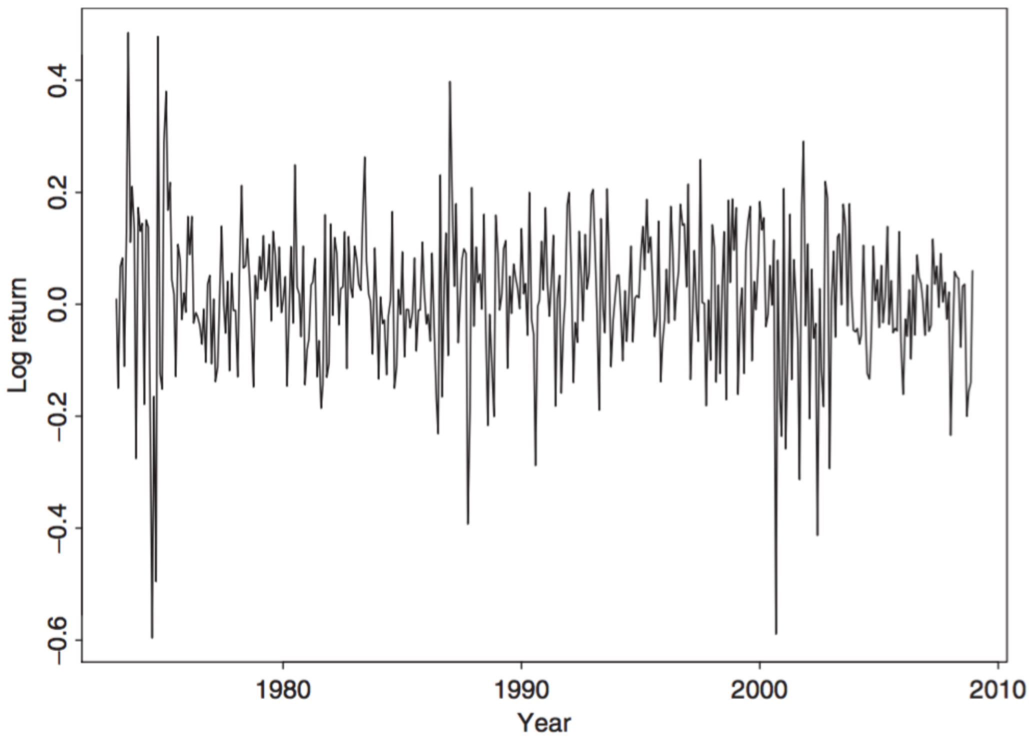

Figure 3: Time plot of monthly log returns of Intel stock from January 1973 to December 2008

For illustration, consider the monthly log stock returns of Intel Corporation from January 1973 to December 2008 shown in Figure 3.

Figure 4 displays the sample ACF and PACF of the log return.

- Figure 4(a) shows the sample ACF of the log return series $\{r_t\}$, which suggests no significant serial correlations except for a minor one at lag 7.
- Figure 4(b) shows the sample ACF of the squared log returns $\{r_t^2\}$, which shows strong autocorrelation in the first few lags.
- Figure 4(c) shows the sample ACF of the absolute log returns, also showing strong autocorrelation.

These two plots clearly suggest that the monthly log returns are not serially independent. Combining the three plots, it seems that the log returns are indeed serially uncorrelated but dependent. Volatility models attempt to capture such dependence in the return series.

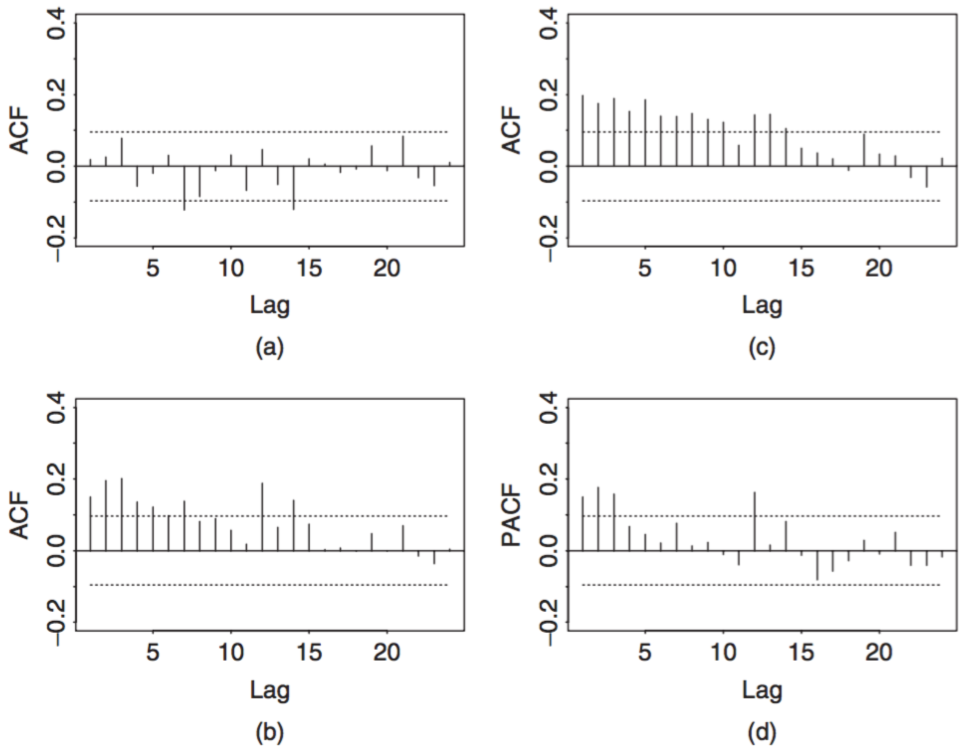

Figure 4: Sample ACF and PACF of various functions of monthly log stock returns of Intel Corporation from January 1973 to December 2008: (a) ACF of the log returns, (b) ACF of the squared log returns, (c) ACF of the absolute log returns, and (d) PACF of the squared log returns.

2.2 The mean equation and the volatility equation

To capture the dependence in a time series through its second moment but not the mean, we model the mean process and the variance process separately.

For a return series $\{r_t\}$, we can model it as

$$r_t = \mu_t + a_t \tag{1}$$

where μ_t represents the conditional mean and a_t is modeled to capture the conditional variance.

2.2.1 The mean equation

The mean equation is essentially an ARMA(p, q) model that is defined in terms of the conditional mean.

$$\mu_t = E(r_t \mid F_{t-1}) = \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i} \tag{2}$$
$$y_t = r_t - \phi_0 - \sum_{i=1}^{k} \beta_i x_{it}$$

F_{t-1} is the information set at time $t-1$. If you combine these two equations, and let $\mu_t = r_t - a_t$, you will find that it is just an ARMA(p, q) model with additional regressors x_{it}.

2.2.2 The variance equation

Let's see what is the conditional mean of r_t. Denote the conditional variance of r_t with σ_t^2.

$$\begin{aligned} \sigma_t^2 = \mathrm{Var}(r_t \mid F_{t-1}) &= E\left((r_t - E(r_t|F_{t-1}))^2|F_{t-1}\right) \\ &= E\left((r_t - \mu_t)^2 \mid F_{t-1}\right) \\ &= \mathrm{Var}(a_t \mid F_{t-1}) \end{aligned}$$

If we assume that $E(a_t \mid F_{t-1}) = 0$, we can see that $\sigma_t^2 = E(a_t^2 \mid F_{t-1})$. We can use the lagged value of a_t^2 to represent the information set F_{t-1}, and write σ_t^2 as a linear model as follows

$$\sigma_t^2 = E(a_t^2|F_{t-1}) = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2$$

The equation above captures the essence of an ARCH model as we will explain later.

2.3 The procedure of building a volatility model

Building a volatility model for an asset return series consists of four steps:

1. Specify a mean equation by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.

   The goal of this step is to get a series of residuals that does not display any autocorrelation.

2. Use the squared residuals of the mean equation to test for ARCH effects.

3. Specify a volatility model if ARCH effects are statistically significant, and perform a joint estimation of the mean and volatility equations.

4. Check the fitted model carefully and refine it if necessary.

2.4 Testing for the presence of ARCH effect

2.4.1 The Ljung-Box test for the series of a_t^2

Upon obtaining the residuals from the estimation of an adequate mean equation, we can use the squared residuals $\{\hat{a}_t^2\}$ to test the existence of autocorrelation.

The Ljung-Box test is used to test the null hypothesis $H_0 : \rho_1 = \cdots = \rho_m = 0$. The $Q(m)$ statistic is calculated and compared with the critical value from $\chi^2(m)$ distribution at the desired significance level. The rejection of the null hypothesis implies that there is autoregressive conditional heteroskedastic (ARCH) effect.

2.4.2 The LM test

We estimate a AR(m) model regarding $\{\hat{a}_t^2\}$, that is,

$$\hat{a}_t^2 = \alpha_0 + \alpha_1 \hat{a}_{t-1}^2 + \cdots + \alpha_m \hat{a}_{t-m}^2 + e_t$$

With this model, we test the joint hypothesis

$$H_0 : \alpha_1 = \cdots = \alpha_m = 0$$

The LM statistic is NR^2 where N is the sample size of this regression and R^2 is the coefficient of the determination of this regression. Given the null hypothesis is true, this statistic follows a $\chi^2(m)$ distribution.

Alternatively, we can use F statistic to test the joint hypothesis.

- Let $SSR_0 = \sum_{t=m+1}^{T} (\hat{a}_t^2 - \bar{\omega})^2$, where $\bar{\omega} = (1/T) \sum_{t=1}^{T} \hat{a}_t^2$. SSR_0 is in fact the restricted sum of squared residuals from the above regression with the m restrictions $\alpha_1 = \cdots = \alpha_m = 0$.
- Let $SSR_1 = \sum_{t=m+1}^{T} \hat{e}_t^2$ where $\hat{e}_t$ is the residuals from the regression. SSR_1 is the unrestricted SSR. The degree of freedom of SSR_1 is $T - 2m - 1 = (T - m) - (m + 1)$.
- The F statistic is

  $$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)} \sim F(m, T - 2m - 1)$$

  When $T \to \infty$, we know mF is asymptotically distributed as a $\chi^2(m)$ distribution.
- Rejecting the null hypothesis motivates us to model the possible ARCH effect.

2.4.3 An example

Go back to Figure 4. Since the return series is already stationary, we directly test the squared return series to check the ARCH effect.

In the LM test of the ARCH effect, $F = 53.62$ and the p value is close to zero. The Ljung–Box statistics of the a_t^2 series also shows strong ARCH effects with $Q(12) = 89.85$, the p value of which is close to zero. Therefore, we can confirm that the return series of Intel stock has an ARCH effect, and next we need to model such an effect.

3 The ARCH Model

3.1 The ARCH(m) Model

3.1.1 The basic idea of an ARCH model

Consider a series of shocks $\{a_t\}$ in a return series $\{r_t\}$. The basic idea of an Autoregressive Conditional Heteroskedasticity (ARCH) model is

1. the shock a_t of the return series is serially uncorrelated but dependent; and,
2. the dependence of a_t can be modeled through an autoregressive process of a_t^2.

3.1.2 The ARCH(m) model

An ARCH(m) model takes the following form

$$a_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 \tag{3}$$

where $\epsilon_t \sim i.i.d.(0, 1)$, $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i = 1, \ldots, m$.

- The assumption of $\mathrm{Var}(\epsilon_t) = 1$ is to make the analysis regarding the properties of the ARCH(m) model easy;
- The assumption of $\alpha_0 > 0$ and $\alpha_i \geq 0$ is to ensure the conditional variance of a_t is positive.
- $\alpha_1, \ldots, \alpha_m$ should also satisfy some regularity conditions to ensure the unconditional variance of a_t is finite.

3.2 The Properties of an ARCH Model

Let's take an ARCH(1) model as an example to discuss the properties of ARCH model and see how such a model can capture the basic idea mentioned above and the stylized fact that highly volatile periods tend to be followed by high volatility periods.

Assume an ARCH(1) model as follows

$$a_t = \sigma_t \epsilon_t,\ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2,\ \epsilon_t \sim i.i.d.(0,1) \tag{4}$$

where $a_0 > 0$ and $a_1 \geq 0$.

3.2.1 The unconditional mean of a_t

$$E(a_t) = E(\sigma_t \epsilon_t) = E(\sigma_t)E(\epsilon_t) = 0$$

The second equality is ensured because σ_t and ϵ_t are independent, and the third equality comes from the assumption of $E(\epsilon_t) = 0$.

3.2.2 The unconditional variance of a_t

$$\begin{aligned}\mathrm{Var}(a_t) &= E(a_t^2) = E(\sigma_t^2 \epsilon_t^2)\\ &= E(\alpha_0 + \alpha_1 a_{t-1}^2) \cdot 1 = \alpha_0 + \alpha_1 \mathrm{Var}(a_{t-1})\end{aligned}$$

Assuming the unconditional mean of a_t is a constant(why?), we can have

$$\mathrm{Var}(a_t) = \frac{\alpha_0}{1-\alpha_1}$$

Since the variance should be positive and finite, we must have $0 \leq \alpha_1 < 1$.

The reason that we need to assume the unconditional mean of a_t to be constant and finite is that we assume the return series $\{r_t\}$ itself is constant. Keep in mind that a complete ARCH model also includes a mean equation for the return series, say, an ARMA model.

3.2.3 The unconditional covariance of a_t

Since ϵ_t and ϵ_{t-i} for $i \neq 0$ are independent,

$$\begin{aligned}\mathrm{Cov}(a_t, a_{t-i}) &= E(a_t a_{t-i}) = E(\sigma_t \epsilon_t \sigma_{t-i} \epsilon_{t-i})\\ &= E(\sigma_t \sigma_{t-i})E(\epsilon_t \epsilon_{t-i}) = 0\end{aligned}$$

Now we know that a_t has constant unconditional mean and variance, and it is serially uncorrelated, satisfying the basic idea of building an ARCH model.

3.2.4 The kurtosis of a_t

Sometimes, we may also require the fourth moment of a_t to be finite so that the variance of a_t will not go wild without bounds.

Assume that $\epsilon \sim N(0,1)$, implying that $E(\epsilon_t^4) = 3$. Thus, we have

$$\begin{aligned} E(a_t^4) &= E(\sigma_t^4\epsilon_t^4) = E(\sigma_t^4)E(\epsilon_t^4) = 3E(\sigma_t^4) \\ &= 3E\left(E_{t-1}(\sigma_t^4)\right) = 3E\left(E_{t-1}(\alpha_0 + \alpha_1 a_{t-1}^2)^2\right) \\ &= 3E(\alpha_0 + \alpha_1 a_{t-1}^2)^2 \\ &= 3\left(\alpha_0^2 + 2\alpha_0\alpha_1 E(a_{t-1}^2) + \alpha_1^2 E(a_{t-1}^4)\right) \end{aligned}$$

Assume that a_t is fourth-order stationary so that we can define $m_4 = E(a_t^4) = E(a_{t-1}^4)$. Then, using the fact that $E(a_t^2) = \alpha_0/(1-\alpha_1)$, we can solve m_4 from the above equation.

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$$

This result has two important implications:

1. Since the fourth moment of a_t is positive, we see that α_1 must also satisfy the condition $1 - 3\alpha_1^2 > 0$, that is, $0 \le \alpha_1^2 < \frac{1}{3}$.
2. The kurtosis of a_t is

$$\text{kurtosis} = \frac{E(a_t^4)}{E(a_t^2)^2} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \cdot \frac{(1-\alpha_1)^2}{\alpha_0^2} = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2} > 3$$

   Thus, the the excess kurtosis of a_t is positive and the tail distribution of a_t is heavier than that of a normal distribution.

   In other words, the shock a_t of a conditional Gaussian ARCH(1) model is more likely than a Gaussian white noise series to produce "outliers".

3.2.5 The conditional mean and variance

From nown on, let's write $E_{t-1}(a_t)$ to represent the conditional mean given the information set F_{t-1}, i.e., $E(a_t \mid F_{t-1})$. Since ϵ_t is i.i.d, we have $E_{t-1}(\epsilon_t) = E(\epsilon_t) = 0$. Thus,

$$\begin{aligned} E_{t-1}(a_t) &= E_{t-1}(\sigma_t\epsilon_t) = E_{t-1}\left((\alpha_0 + \alpha_1 a_{t-1}^2)^{1/2}\epsilon_t\right) \\ &= (\alpha_0 + \alpha_1 a_{t-1}^2)^{1/2} E_{t-1}(\epsilon_t) = 0 \end{aligned}$$

The conditional variance of a_t is

$$\begin{aligned}
\mathrm{Var}_{t-1}(a_t) &= E_{t-1}(a_t^2) = E_{t-1}\left(\sigma_t^2\epsilon_t^2\right) \\
&= E_{t-1}\left((\alpha_0+\alpha_1 a_{t-1}^2)\epsilon_t^2\right) \\
&= E_{t-1}(\alpha_0+\alpha_1 a_{t-1}^2)E_{t-1}(\epsilon_t^2) \\
&= (\alpha_0+\alpha_1 a_{t-1}^2)E(\epsilon_t^2) \\
&= \alpha_0+\alpha_1 a_{t-1}^2 = \sigma_t^2
\end{aligned}$$

By calculating the conditional variance, we see that if the realized value of a_{t-1}^2 is large, the conditional variance of a_t will be large as well. This essentially captures the stylized fact of volatility that high-volatility periods tend to follow previous high-volitility periods.

4 Estimation and Forecasting

4.1 Order determination

Before estimating an ARCH(m) model, we need to determine the order m. The basic idea is that we treat an ARCH(m) model as an AR process of $\{a_t^2\}$, and apply the partial autocorrelation function (PACF) to determine m.

We justify the use of the PACF of $\{a_t^2\}$ to determine m through two perspectives.

1. We can consider a_t^2 as an unbiased estimator of σ_t^2 given the sample data because $E_{t-1}(a_t^2) = \sigma_t^2$. Therefore, we use a_t^2 as an approximate to σ_t^2.
2. We can define $\eta_t = a_t^2 - \sigma_t^2$. It can be shown that
   - $E(\eta_t) = 0$ and $E(\eta_t\eta_{t-s}) = 0$ for $s > 0$.
   - But η_t is not i.i.d. because a_t^2 is dependent.

   So an ARCH(m) model is essentially an AR(m) model, except that η_t is not i.i.d. That is,

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 + \eta_t$$

Therefore, we can use the PACF of $\{a_t^2\}$ to determine the order m. However, the PACF may not be effective when the sample size is small.

4.2 Maximum likelihood estimation

4.2.1 A brief introduction of maximum likelihood function

Likelihood function when y_t is independent The likelihood function is the joint density function $f(\mathbf{y}|\boldsymbol{\theta})$ when we consider it as a function of the parameters $\boldsymbol{\theta}$ given a set of data $\mathbf{y}$.

- $\mathbf{y} = (y_1, y_2, \ldots, y_T)$ represents the observations, which are assumed to be identically independently distributed.
- $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ represents all parameters in the model (i.e., data generating process) that generates $\mathbf{y}$.

For each observation, y_t, the marginal density function is $f(y_t|\boldsymbol{\theta})$ and because y_i is independent, **the likelihood function** is

$$L(\boldsymbol{\theta}|\mathbf{y}) = \prod_{t=1}^{T} f(y_t|\boldsymbol{\theta}) \tag{5}$$

And the **log-likelihood** function is

$$l(\boldsymbol{\theta}|\mathbf{y}) = \sum_{t=1}^{T} l_t(y_t|\boldsymbol{\theta}) \tag{6}$$

where $l_t(y_t|\boldsymbol{\theta}) = \ln\left(f(y_t|\boldsymbol{\theta})\right)$ is the **contribution** to the loglikelihood function.

Likelihood function when y_t is dependent

Maximum likelihood estimation

4.2.2 Maximum likelihood estimation of ARCH(m)