《Analysis of Financial Time Series》

Ch.3 Conditional Heteroscedastic Model

April 25, 2022

Outline

Conditional Heteroscedastic Model

- What is stock volatility?
  Answer: conditional standard deviation of stock returns.
- Common characteristics of volatility。
  - volatility clustering effect (ie. volatility 有群聚效應，在某一時段皆高，而在另一時段皆低)。
  - It is rare to have volatility jumps. (ie. volatility is a continuous variable of time 波動是時間的連續隨機變數)。
  - volatility is finite (ie. 波動值為有限不會發散到無窮)。
  - big price increase and big price drop exhibit different volatility behavior (EGARCH model is proposed to describe this property, 針對此性質而提出來的模型)。

Conditional Heteroscedastic Model (條件異質變異模型)

- Why is volatility important?
- Has many important applications.
- Option (derivative) pricing, e.g., Black-Scholes formula.
- Risk management, e.g. value at risk (VaR).
- Asset allocation, e.g., minimum-variance portfolio; see pages 184-185 of Campbell, Lo and MacKinlay (1997).
- Interval forecasts.
- **A key characteristic:** Not directly observable!!

- Black-Scholes risk-neutral Model

$$
\begin{aligned}
dS_t &= rS_t dt + \sigma S_t dW_t \\
d\ln S_t &= (r - \sigma^2/2)dt + \sigma dW_t
\end{aligned}
$$

$r =$ risk-free interest rate 。

- Black-Scholes European call option price:

$$
\begin{aligned}
C_t &= e^{-r(T-t)} E_Q[(S_T - K)^+ | S_t] \\
&= S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2) \\
d_1 &= \frac{\ln(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \\
d_2 &= d_1 - \sigma\sqrt{T-t}
\end{aligned}
$$

$\Phi(d_2) = P_Q(S_T > K | S_t)$ under risk neutral measure, the probability that the call option will be executed (在風險中立測度下，買權會被執行的機率)，where $K =$ strike price and $T =$ maturity time 。

How to calculate volatility ?

1. Use high-frequency data: French, Schwert & Stambaugh (1987); see Section 3.15.
    - Realized volatility of daily returns in recent literature.
    - Use daily high, low, and closing prices.
2. Econometric modeling.
3. Implied volatility of options data, e.g: VIX of CBOE.
4. VIX: volatility index，由 CBOE 所編譯的波動指數，2004/3/26 開始在期貨市場交易 。
5. CBOE (Chicago Board Options Exchange) 芝加哥選擇權交易所 。
6. 實証中發現，implied volatility 大多高於 GARCH 模型所估出的 volatility.
7. See Figure 1 and Figure 2.

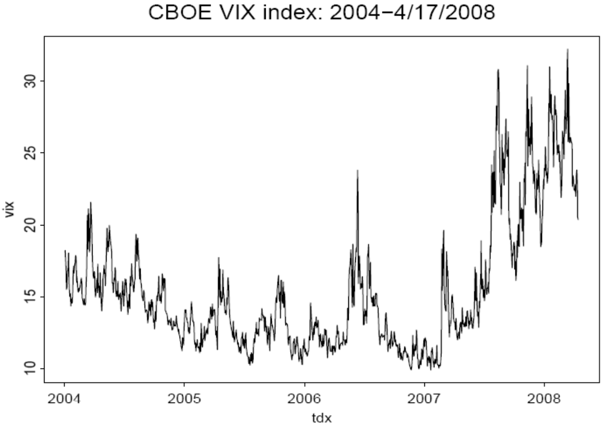


Figure 1: Time plot of the daily closing value of VIX of the CBOE: January 2, 2004 to April 17, 2008.

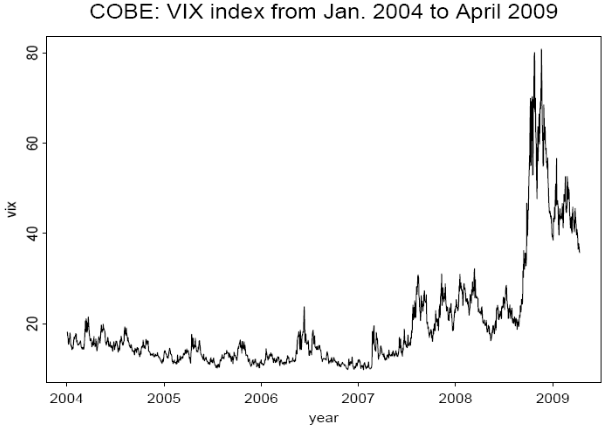


Figure 2: Time plot of the daily closing value of VIX of the CBOE: January 2, 2004 to April 16, 2009.

- We focus on the econometric modeling first.
- Use of high frequency data will be discussed later.
- **Basic idea** of econometric modeling Shocks of asset returns are NOT serially correlated, but dependent.
- That is, the dependence is nonlinear.
- As shown by the ACF of returns and absolute returns of some assets we discussed so far.
- For illustration, consider the monthly log stock returns of Intel Corporation from January 1973 to December 2008 shown in Figure 3.1 and Figure 3.2.

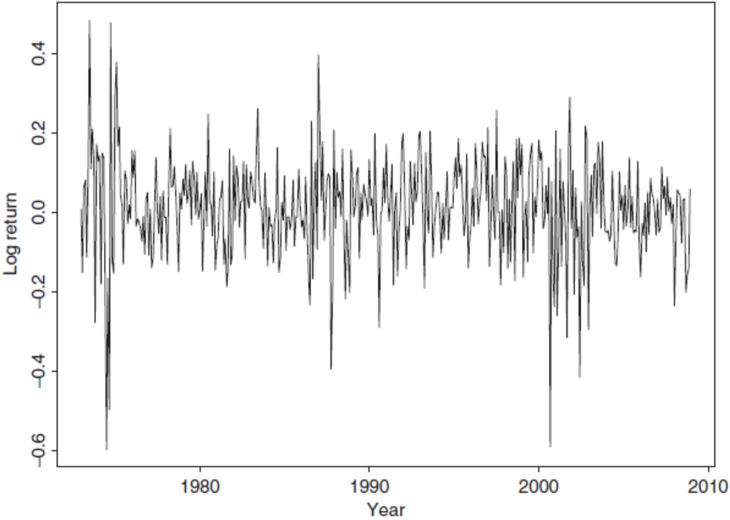


Figure 3.1 Time plot of monthly log returns of Intel stock from January 1973 to December 2008.

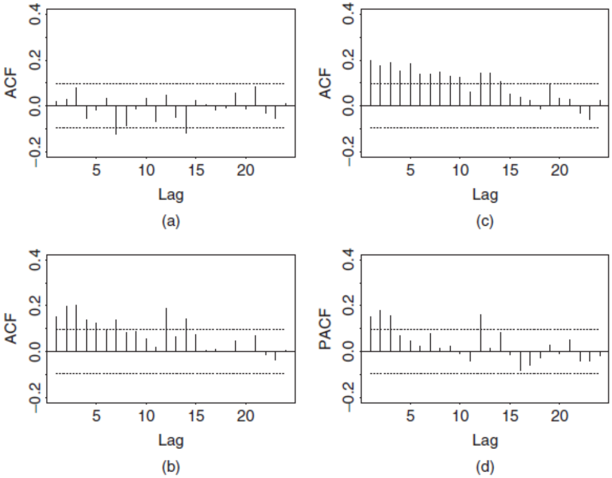


Figure 3.2 Sample ACF and PACF of various functions of monthly log stock returns of Intel Corporation from January 1973 to December 2008: (a) ACF of the log returns, (b) ACF of the squared log returns, (c) ACF of the absolute log returns, and (d) PACF of the squared log returns.

Basic structure

$$r_t = \mu_t + a_t, \quad \mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i},$$

Volatility models are concerned with time-evolution of

$$\sigma_t^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1}).$$

the conditional variance of a return.

Two general categories

- Fixed function
  e.g: GARCH 模型
- Stochastic function of the available information.
  e.g: SVM (stochastic volatility model) 模型

Univariate volatility models

1. Autoregressive conditional heteroscedastic (ARCH) model of Engle (1982)
2. Generalized ARCH (GARCH) model of Bollerslev (1986)
3. GARCH-M models
4. IGARCH models
5. Exponential GARCH (EGARCH) model of Nelson (1991)
6. Threshold GARCH model of Zakoian (1994) or GJR model of Glosten, Jagannathan, and Runkle (1993).
7. Conditional heteroscedastic ARMA (CHARMA) model of Tsay (1987),
8. Random coefficient autoregressive (RCA) model of Nicholls and Quinn (1982)
9. Stochastic volatility (SV) models of Melino and Turnbull (1990), Harvey, Ruiz and Shephard (1994), and Jacquier, Polson and Rossi (1994).

ARCH model (Engle,1982)

- ARCH model is the first model that provides a systematic framework for volatility modeling.
- The basic idea of ARCH models is that
  - (a) the shock (mean corrected return) a_t, of an asset return is serially uncorrelated, but dependent.
  - (b) the dependence of a_t can be described by a simple quadratic function of its lagged values.
- $$r_t = \mu_t + a_t, \quad \mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i}$$

ARCH(m) Model

- Let $a_t = r_t - \mu_t$ be the mean corrected return and $\sigma_t^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1})$ be the conditional variance of r_t. Assume

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 \end{aligned}$$

  - $\{\epsilon_t\}$ is a sequence of iid r.v's with mean 0 and variance 1.
  - $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$.
- The coefficients α_i must satisfy some regularity conditions to ensure that the unconditional variance of a_t is finite.
- Distribution of ϵ_t: Standard normal, standardized Student-t, generalized error dist (GED), or skewed Student-t.
- The ARCH effects also occurs in other financial time series such as the percentage change in exchange rate.

德國馬克對美元匯率的 10 分鐘報酬率

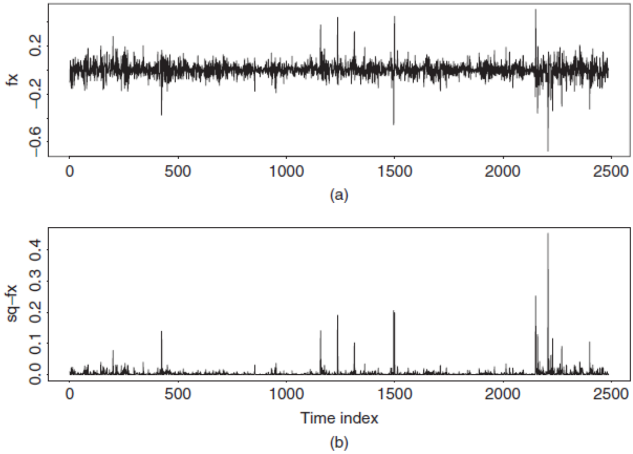


Figure 3.3 (a) Time plot of 10-minute returns of exchange rate between Deutsche mark and U.S. dollar from June 5, 1989, to June 19, 1989, and (b) the squared returns.

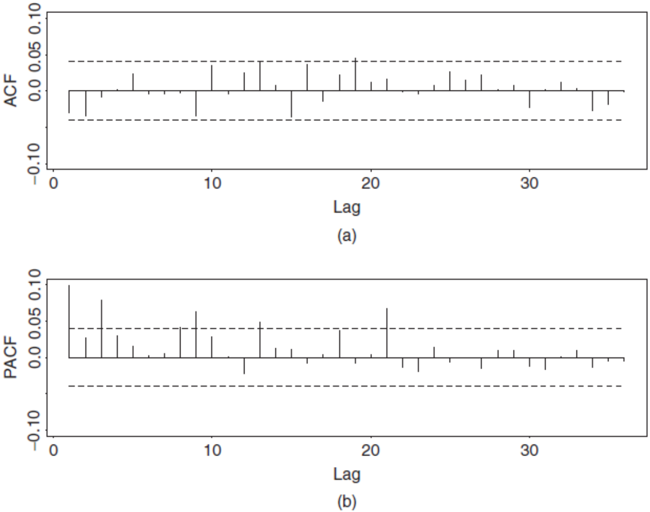


Figure 3.4 (a) Sample autocorrelation function of return series of mark/dollar exchange rate and (b) sample partial autocorrelation function of squared returns.

Properties of ARCH(1) models

- Consider an ARCH(1) model

$$a_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2,$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$.

-

$$\begin{aligned} Cov(a_t, a_{t+l}) &= E(a_t a_{t+l}) = E(\sigma_t \epsilon_t \sigma_{t+l} \epsilon_{t+l}) \\ &= E[\sigma_t \sigma_{t+l} E(\epsilon_t \epsilon_{t+l} | F_{t+l-1})] = 0 \end{aligned}$$

也就是說 a_t, a_{t+l} 為 uncorrelated，即 $\{a_t\}$ 為 White Noise。

- 但是

$$a_{t+1}^2 = \sigma_{t+1}^2 \epsilon_{t+1}^2 = (\alpha_0 + \alpha_1 a_t^2) \epsilon_{t+1}^2$$

顯然 a_{t+1}^2 與 a_t^2 為 correlated。因此，a_t 與 a_{t+1} 並非獨立。

- $$\begin{aligned} a_t^2 &= \sigma_t^2 \varepsilon_t^2 \\ &= \sigma_t^2(\varepsilon_t^2 - 1) + \sigma_t^2 \\ &= v_t + \alpha_0 + \alpha_1 a_{t-1}^2 \end{aligned}$$

  其中 $v_t = \sigma_t^2(\varepsilon_t^2 - 1)$ 。

- $$\begin{aligned} E(v_t) &= 0 \\ E(v_t v_{t+l}) &= E(\sigma_t^2 \sigma_{t+l}^2 E((\varepsilon_t^2 - 1)(\varepsilon_{t+l}^2 - 1))) \\ &= 0 \end{aligned}$$

  因此，v_t 為一個 White Noise Process 。
  $a_t^2 = v_t + \alpha_0 + \alpha_1 a_{t-1}^2 \Rightarrow a_t^2$ 有一個 AR(1) 的表示式 。

- Mean

$$\begin{aligned} E(a_t) &= E(\sigma_t \epsilon_t) \\ &= E[E(\sigma_t \epsilon_t | F_{t-1})] \\ &= E[\sigma_t E(\epsilon_t | F_{t-1})] \\ &= E[\sigma_t E(\epsilon_t)] \\ &= 0 \end{aligned}$$

- Variance

$$E(a_t^2) = E[\sigma_t^2 E(\epsilon_t^2 | F_{t-1})] = E(\sigma_t^2)$$

$$E(\sigma_t^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2)$$

$$\Rightarrow E(a_t^2) = \frac{\alpha_0}{1 - \alpha_1}$$

- Under normality assumption, i.e. $\epsilon_t \sim N(0,1)$:

$$E(a_t^4) = E[\sigma_t^4 E(\epsilon_t^4 | F_{t-1})] = 3E(\sigma_t^4)$$
$$\sigma_t^4 = \alpha_0^2 + 2\alpha_0\alpha_1 a_{t-1}^2 + \alpha_1^2 a_{t-1}^4$$
$$\frac{1}{3}E(a_t^4) = \alpha_0^2 + 2\alpha_0\alpha_1 \frac{\alpha_0}{1-\alpha_1} + \alpha_1^2 E(a_{t-1}^4)$$
$$\Rightarrow E(a_t^4) = \frac{\alpha_0^2[1-\alpha_1+2\alpha_1]}{(\frac{1}{3}-\alpha_1^2)(1-\alpha_1)} = \frac{3\alpha_0^2[1+\alpha_1]}{(1-3\alpha_1^2)(1-\alpha_1)}$$

- Excess kurtosis

$$\frac{E(a_t^4)}{E^2(a_t^2)} - 3 = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2} - 3$$
$$= \frac{6\alpha_1^2}{1-3\alpha_1^2} > 0$$

所以 ARCH(1) 模型的平穩分佈是厚尾 (heavy tailed)。

Summary of ARCH model property

- Consider an ARCH(1) model

$$a_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2,$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$.

1. $E(a_t) = 0$
2. $Var(a_t) = \alpha_0/(1-\alpha_1)$, if $0 \leq \alpha_1 < 1$
3. Under normality assumption of ϵ_t,

$$m_4 = E(a_t^4) = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$$

provided $0 \leq \alpha_1^2 < 1/3$.

- The 3rd property implies heavy tails.

Advantages & Weaknesses of ARCH models

- Advantages
  - Simplicity.
  - Generates volatility clustering.
    當 a_{t-1}^2 大時，$\sigma_t^2 = E(a_t^2|F_{t-1}) = \alpha_0 + \alpha_1 a_{t-1}^2$ 亦會變大。
  - Heavy tails (high kurtosis).
- Weaknesses.
  - Symmetric btw positive & negative prior returns.
    (即 a_{t-1} 的正負號，對 σ_t^2 的大小沒有差異。)
  - Restrictive 模型參數的限制較多!
    ARCH(1) model 為保證 $E(a_t^4) < \infty$, 要求 $0 < \alpha_1^2 < \frac{1}{3}$. 對於 $ARCH(p)$, $p > 1$，其對參數的限制將更複雜。
  - Provides no explanation.
    僅提供一個描述條件波動的模型，並未解釋 variation 的來源。
  - Not sufficiently adaptive (適應的) in prediction.
    ARCH 模型對於 large isolate stock 的反應較慢，故較易高估 volatility。

Test for ARCH effect

兩種檢定資料是否具有 ARCH effect 的方式 (即是否存在 conditional heteroscedastic 的效應)

1. 利用 $Q(m)$ 檢定 $\{a_t^2\}$ 的 acf 是否顯著。
2. 利用 Lagrange multiplier(LM) test 檢定線性模型的參數是否顯著.

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

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$$

$$\text{Test Stat. } F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)} \sim \chi_m^2$$

$$SSR_0 = \sum_{t=m+1}^{T} (a_t^2 - \varpi)^2, \quad \varpi = \frac{1}{T}\sum_{t=1}^{T} a_t^2$$

$$SSR_1 = \sum_{i=m+1}^{T} \hat{e}_t^2, \quad \hat{e}_t \text{ is the least-squares residual.}$$

Example

- Consider the monthly log stock returns of Intel Corporation from 1973 to 2008.
- The series does not have significant serial correlations so that it can be directly used to test for the ARCH effect.
- The $Q(m)$ statistics of the return series give $Q(12) = 18.26$ with a p value of 0.11, confirming no serial correlations in the data.
- The Lagrange multiplier test shows strong ARCH effects with test statistic $F \approx 53.62$, the p value of which 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.

Denote the return series by "intc". Note that the command "archTest" applies directly to the a_t series, not to a_t^2.

R Demonstration

```
> da=read.table("m-intc7308.txt",header=T)
> intc=log(da[,2]+1)
> Box.test(intc,lag=12,type='Ljung')
        Box-Ljung test

data:  intc
X-squared = 18.2635, df = 12, p-value = 0.1079

> at=intc-mean(intc)
> Box.test(at^2,lag=12,type='Ljung')
        Box-Ljung test

data:  at^2
X-squared = 89.8509, df = 12, p-value = 5.274e-14
```

R-code:

> library(FinTS)

> ArchTest(x,lags=12,demean=FALSE)

where

x: data

demean: If TRUE, remove the mean before computing the test statistic

Building an ARCH Model

1. Modeling the mean effect (ARMA model) and testing for ARCH effect
   Use Q-statistics of squared residuals; McLeod and Li (1983) & Engle (1982)
2. Order determination
   Use PACF of the squared residuals
3. Estimation: Conditional MLE
4. Model checking:
   - Q-stat of standardized residuals and squared standardized residuals.
   - Skewness & Kurtosis of standardized residuals. (檢定 $\{\frac{a_t}{\sigma_t}\}$ 是否為常態分佈)
5. Software: Many available. We use R in class.

Estimation

- Under normality, the likelihood function of an ARCH(m) model is

$$\begin{aligned} f(a_1, \cdots, a_T|\boldsymbol{\alpha}) &= f(a_T|F_{T-1})f(a_{T-1}|F_{T-2})\cdots f(a_{m+1}|F_m)f(a_1,\cdots,a_m|\boldsymbol{\alpha}) \\ &= \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{a_t^2}{2\sigma_t^2}\right) \times f(a_1,\cdots,a_m|\boldsymbol{\alpha}) \end{aligned}$$

  where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \cdots, \alpha_m)'$ and $f(a_1, a_2, \cdots, a_m|\boldsymbol{\alpha})$ is the joint pdf of $a_1, a_2, \cdots, a_m$.
- The conditional log likelihood function (ignoring the constant term) is

$$l(a_{m+1}, \cdots, a_T|\boldsymbol{\alpha}, a_1, \cdots, a_m) = -\sum_{t=m+1}^{T}\left[\frac{1}{2}\ln(\sigma_t^2) + \frac{1}{2}\frac{a_t^2}{\sigma_t^2}\right],$$

  where $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2$ can be evaluated recursively.

- Under the t-dist, the marginal pdf is

$$f_{T_\nu}(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \quad -\infty < t < \infty$$

$E(T_\nu) = 0$, $Var(T_\nu) = \frac{\nu}{\nu-2}$, $\nu > 2$.

- 若令 $T^* = \sqrt{\frac{\nu-2}{\nu}} T_\nu$，则 $Var(T^*) = 1$

$$\Rightarrow f_{T^*}(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\nu-2)\pi}} \left(1 + \frac{t^2}{\nu-2}\right)^{-(\nu+1)/2}, \nu > 2$$

- Under (a_t/σ_t is a scaled) t-dist, if ν is given, then the cond. log likelihood fun. is

$$
\begin{aligned}
& l(a_{m+1}, \cdots, a_T | \boldsymbol{\alpha}, A_m) \\
& = -\sum_{t=m+1}^{T} \left[\frac{\nu+1}{2} \ln\left(1 + \frac{a_t^2}{(\nu-2)\sigma_t^2}\right) + \frac{1}{2}\ln(\sigma_t^2) \right],
\end{aligned}
$$

where $A_m = (a_1, \cdots, a_m)$.

- If ν is unknown, then it becomes

$$
\begin{aligned}
& l(a_{m+1}, \cdots, a_T | \boldsymbol{\alpha}, A_m) \\
& = (T-m)\left\{ \ln\left[\Gamma\left(\frac{\nu+1}{2}\right)\right] - \ln\left[\Gamma\left(\frac{\nu}{2}\right)\right] - 0.5\ln((\nu-2)\pi) \right\} \\
& \quad + l(a_{m+1}, \cdots, a_T | \boldsymbol{\alpha}, A_m).
\end{aligned}
$$

Comparison with normal models:

- Using a heavy-tailed dist for ϵ_t reduces the ARCH effect
- You may also try the following generalized error distribution :

$$f(x) = \frac{v \times \exp(-\frac{1}{2}|x/\lambda|^v)}{\lambda 2^{(1+\frac{1}{v})}\Gamma(\frac{1}{v})}, -\infty < x < \infty, 0 < v < \infty.$$

$$\lambda = \left[2^{-\frac{2}{v}}\Gamma(\frac{1}{v})\Gamma(\frac{3}{v})\right]^{\frac{1}{2}}.$$

- When $v = 2 \Rightarrow$ Gaussian, $v < 2 \Rightarrow$ Heavy tail.

Model Checking

- For a properly specified ARCH model, the standardized residuals
$$\tilde{a}_t = \frac{a_t}{\sigma_t}$$
form a sequence of iid random variables. Therefore, one can check the adequacy of a fitted ARCH model by examining the series $\{\tilde{a}_t\}$.
- The Ljung–Box statistics of $\tilde{a}_t$ can be used to check the adequacy of the mean equation and that of $\tilde{a}_t^2$ can be used to test the validity of the volatility equation.
- The skewness, kurtosis, and quantile-to-quantile plot of $\{\tilde{a}_t\}$ can be used to check the validity of the distribution assumption.

Forecasting

- Consider an ARCH(m) model. At the forecast origin h, the 1-step-ahead forecast of σ^2_{h+1} is

$$\sigma^2_h(1) = \alpha_0 + \alpha_1 a^2_h + \cdots + \alpha_m a^2_{h+1-m}.$$

The 2-step-ahead forecast of σ^2_{h+2} is

$$\sigma^2_h(2) = \alpha_0 + \alpha_1 \sigma^2_h(1) + \alpha_2 a^2_h + \cdots + \alpha_m a^2_{h+2-m}.$$

- The l-step-ahead forecast for σ^2_{h+l} is

$$\sigma^2_h(l) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma^2_h(l-i),$$

where $\sigma^2_h(l-i) = a^2_{h+l-i}$ if $l - i \leq 0$.

Example 1

- We first apply the modeling procedure to build a simple ARCH model for the monthly log returns of Intel stock (1973 ∼ 2003 的月報酬率).
- According to Figure 3.2 (the PACF of a_t^2 are significant up to lag 4, p.11), we specify the model

$$r_t = \mu + a_t,\ a_t = \sigma_t \epsilon_t,\ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \alpha_3 a_{t-3}^2$$

  for the monthly log returns of Intel stock.
- Assuming that ϵ_t are iid standard normal, we obtain the fitted model

$$r_t = 0.0122 + a_t,\ \sigma_t^2 = 0.0106 + 0.2131 a_{t-1}^2 + 0.0770 a_{t-2}^2 + 0.0599 a_{t-3}^2$$

  where the standard errors of the parameters are 0.0057, 0.0010, 0.0757, 0.0480, and 0.0688, respectively.

- While the estimates meet the general requirement of an ARCH(3) model, the estimates of α_2 and α_3 appear to be statistically nonsignificant at the 5% level.
- Dropping the two nonsignificant parameters, we obtain the model

$$r_t = 0.0126 + a_t, \ \sigma_t^2 = 0.0111 + 0.3560 a_{t-1}^2, \qquad (3.12)$$

where the standard errors of the parameters are 0.0053, 0.0010, and 0.0761, respectively.
- All the estimates are highly significant.
- Figure 3.5 shows the standardized residuals $\{\tilde{a}_t\}$ and the sample ACF of some functions of $\{\tilde{a}_t\}$.
- The Ljung-Box statistics of standardized residuals $\{\tilde{a}_t\}$ give $Q(10) = 12.64$ with a p value of 0.24 and those of $\{\tilde{a}_t^2\}$ give $Q(10) = 14.75$ with a p value of 0.14.

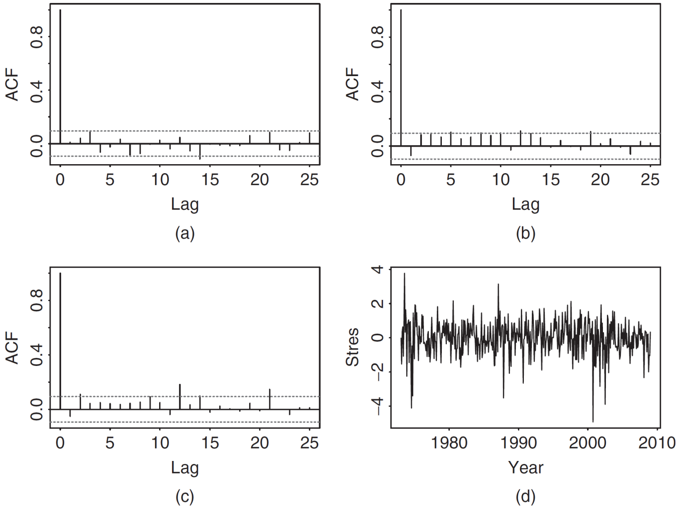


Figure 3.5 Model checking statistics of Gaussian ARCH(1) model in Eq. (3.12) for monthly log returns of Intel stock from January 1973 to December 2008: Parts (a), (b), and (c) show sample ACF of standardized residuals, their squared series, and absolute series, respectively; part (d) is time plot of standardized residuals.

R Demonstration

```
>library(fGarch)
>da=read.table("m-intc7308.txt",header=T)
>intc=as.vector(log(da[,2]+1))
>arch3.fit=garchFit(~ garch(3), data=intc, trace = FALSE)
>summary(arch3.fit)
Title:
 GARCH Modelling
Call:
 garchFit(formula = ~garch(3), data = intc, trace = FALSE)
Mean and Variance Equation:
 data ~ garch(3)
Conditional Distribution:
 norm
```

trace=FALSE: The parameter optimization process are not printed.

```
Coefficient(s):
        mu      omega     alpha1     alpha2     alpha3
0.011852   0.010588   0.237149   0.072747   0.053080
Std. Errors:
 based on Hessian
Error Analysis:
         Estimate   Std. Error   t value   Pr(>|t|)
mu       0.011852   0.005640     2.102     0.0356 *
omega    0.010588   0.001284     8.249   2.22e-16 ***
alpha1   0.237149   0.114734     2.067     0.0387 *
alpha2   0.072747   0.046990     1.548     0.1216
alpha3   0.053080   0.046526     1.141     0.2539
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Log Likelihood:
 291.8891    normalized:  0.6756692
```

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	192.312	0
Shapiro-Wilk Test	R	W	0.9638119	7.831191e-09
Ljung-Box Test	R	Q(10)	9.731162	0.4643876
Ljung-Box Test	R	Q(15)	18.39307	0.2425647
Ljung-Box Test	R	Q(20)	19.28568	0.5033325
Ljung-Box Test	R^2	Q(10)	7.2857	0.6982258
Ljung-Box Test	R^2	Q(15)	26.92023	0.02939306
Ljung-Box Test	R^2	Q(20)	27.70902	0.1164684
LM Arch Test	R	TR^2	24.88263	0.01538893

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-1.328190	-1.281102	-1.328454	-1.309600

```
> arch1.fit=garchFit(~ garch(1), data=intc, trace = FALSE)
> summary(arch1.fit)
Title:
 GARCH Modelling
Call:
 garchFit(formula = ~garch(1), data = intc, trace = FALSE)
Mean and Variance Equation:
 data ~ garch(1)
Conditional Distribution:
 norm
```

```
Coefficient(s):
        mu      omega     alpha1
  0.012637   0.011195   0.379492
Std. Errors:
 based on Hessian
Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu       0.012637    0.005428    2.328  0.01990 *
omega    0.011195    0.001239    9.034  < 2e-16 ***
alpha1   0.379492    0.115534    3.285  0.00102 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Log Likelihood:
 288.0589    normalized:  0.6668031
```

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	137.919	0
Shapiro-Wilk Test	R	W	0.9679255	4.025299e-08
Ljung-Box Test	R	Q(10)	12.54002	0.2505382
Ljung-Box Test	R	Q(15)	21.33508	0.1264607
Ljung-Box Test	R	Q(20)	23.19679	0.2792354
Ljung-Box Test	R^2	Q(10)	16.0159	0.09917815
Ljung-Box Test	R^2	Q(15)	36.08022	0.001721296
Ljung-Box Test	R^2	Q(20)	37.43683	0.01036728
LM Arch Test	R	TR^2	26.57744	0.008884587

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-1.319717	-1.291464	-1.319813	-1.308563

- For comparison, we also fit an ARCH(1) model with Student-t innovations to the series.
- The resulting model is

$$r_t = 0.0169 + a_t, \ \sigma_t^2 = 0.0120 + 0.2845 a_{t-1}^2, \tag{3.13}$$

where the standard errors of the parameters are 0.0053, 0.0017, and 0.1120, respectively. The estimated degrees of freedom is 6.01 with standard error 1.50.
- All the estimates are significant at the 5% level.
- The unconditional standard deviation of a_t is

$$\sqrt{0.0120/(1-0.2845)} \approx 0.1295,$$

which is close to that obtained under normality:

$$r_t = 0.0126 + a_t, \ \sigma_t^2 = 0.0111 + 0.3560 a_{t-1}^2.$$

Under normality, the unconditional stdev of a_t is

$$\sqrt{0.0111/(1-0.3560)} \approx 0.1312,$$

- The Ljung–Box statistics of the standardized residuals give $Q(12) = 14.88$ with a p value of 0.25, confirming that the mean equation is adequate.
- However, the Ljung–Box statistics for the squared standardized residuals show $Q(12) = 35.42$ with a p value of 0.0004. The volatility equation is inadequate at the 1% level.
- Further analysis shows that $Q(10) = 15.90$ with a p value of 0.10 for the squared standardized residuals.
- The inadequancy of the volatility equation is due to a large lag-12 ACF ($\rho_{12} = 0.188$) of the squared standardized residuals.

```
 % The next command fits an ARCH(1) model with Student-t dist.
> m3=garchFit(intc~garch(1,0),data=intc,trace=F,
   cond.dist='std')
> summary(m3) % Output shortened.

Call:
 garchFit(formula=intc~garch(1,0), data=intc, cond.dist="std",
    trace = F)
Mean and Variance Equation: data ~ garch(1, 0) [data = intc]
Conditional Distribution:  std   % Student-t distribution

Coefficient(s):
      mu     omega    alpha1     shape
0.016731  0.011939  0.285320  6.015195

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.016731    0.005302    3.155 0.001603 **
omega   0.011939    0.001603    7.449  9.4e-14 ***
alpha1  0.285320    0.110607    2.580 0.009892 **
shape   6.015195    1.562620    3.849 0.000118 ***
% Degrees of freedom
```

```
% The next command fits an ARCH(1) model with skew
%Student-t dist.
> m4=garchFit(intc~garch(1,0),data=intc,cond.dist='sstd',
   trace=F)
% Next, fit an ARMA(1,0)+GARCH(1,1) model with
% Gaussian noises.
> m5=garchFit(intc~arma(1,0)+garch(1,1),data=intc,trace=F)
```

Comparing models ARCH(1)-normal and ARCH(1)-t, we see that

- Using a heavy-tailed distribution for ϵ_t reduces the ARCH coefficient.
- The difference between the two models is small for this particular instance.
- Finally, a more appropriate conditional heteroscedastic model for the monthly log returns of Intel stock is a GARCH$(1, 1)$ model, which is discussed in the next section.

Example 2

- Consider the percentage changes of the exchange rate between Mark and Dollar in 10-minute intervals.
- The data (time plots of the returns and squared returns) are shown in Figure 3.3(a) on next page.
- As shown in Figure 3.4(a)(sample acf of the returns), the series has no serial correlations.
- The sample PACF of the squared series a_t^2 (Figure 3.4(b)) shows some big spikes, especially at lags 1 and 3.
- There are some large PACF at higher lags, but the lower order lags tend to be more important.

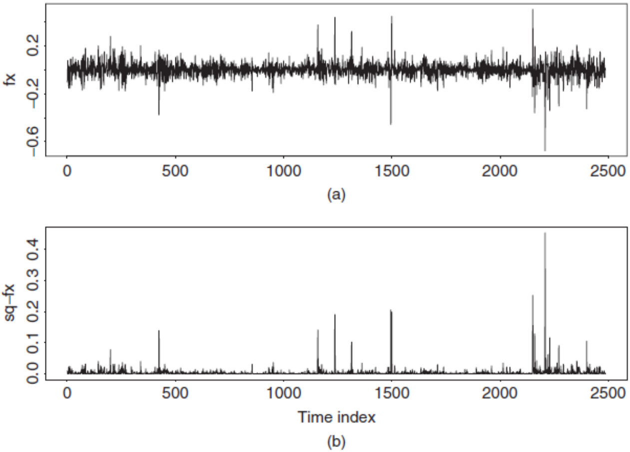


Figure 3.3 (a) Time plot of 10-minute returns of exchange rate between Deutsche mark and U.S. dollar from June 5, 1989, to June 19, 1989, and (b) the squared returns.

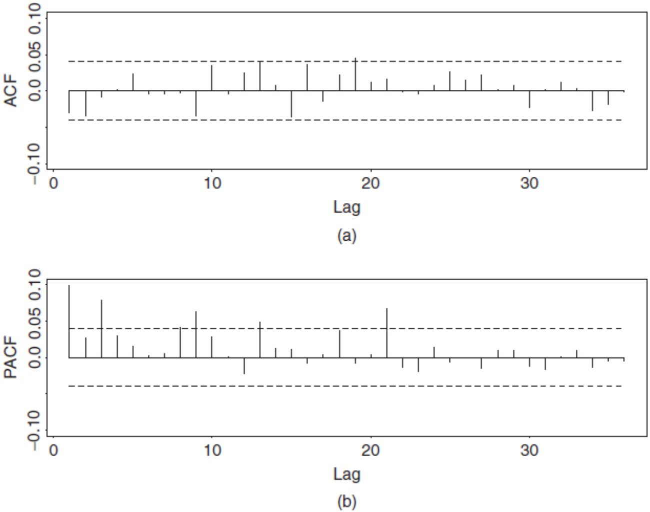


Figure 3.4 (a) Sample autocorrelation function of return series of mark/dollar exchange rate and (b) sample partial autocorrelation function of squared returns.

- We specify an ARCH(3) model for the series. Using the conditional Gaussian likelihood function, we obtain the fitted model

$$
\begin{aligned}
r_t &= 0.0018 + \sigma_t \epsilon_t, \\
\sigma_t^2 &= 0.22 \times 10^{-2} + 0.322 a_{t-1}^2 + 0.074 a_{t-2}^2 + 0.093 a_{t-3}^2,
\end{aligned}
$$

  where all the estimates in the volatility equation are statistically significant at the 5% significant level, and the standard errors of the parameters are 0.47×10^{-6}, 0.017, 0.016, and 0.014, respectively.
- Model checking, using the standardized residual $\tilde{a}_t$, indicates that the model is adequate.

GARCH(m,s) Model

- Let a_t denote the mean corrected return at time t,

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \end{aligned}$$

- $\{\epsilon_t\}$: iid $E(\epsilon_t) = 0,\ Var(\epsilon_t) = 1$
- $\alpha_0 > 0$
- $\alpha_i \geq 0,\ \beta_j \geq 0$ ($\alpha_i = 0$ for $i > m$ and $\beta_j = 0$ for $j > s$)
- $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$.
  $\Rightarrow$ implies that the unconditional variance of a_t is finite.

Re-parameterization:

- 當 a_t 為 GARCH(m,s) 時，a_t^2 有一個 ARMA(max(m,s),s) 的表達式。

$$
\begin{aligned}
a_t^2 &= a_t^2 - \sigma_t^2 + \sigma_t^2 = \eta_t + \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \\
&= \eta_t + \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j (-a_{t-j}^2 + \sigma_{t-j}^2) + \sum_{j=1}^{s} \beta_j a_{t-j}^2 \\
&= \alpha_0 + \sum_{j=1}^{\max(m,s)} (\alpha_j + \beta_j) a_{t-j}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}
\end{aligned}
$$

wher $\eta_t = a_t^2 - \sigma_t^2 = \sigma_t^2(\epsilon_t^2 - 1)$: martingale difference series.

- $E(\eta_t|\mathfrak{F}_{t-1}) = E(a_t^2|\mathfrak{F}_{t-1}) - \sigma_t^2 = 0$
  $\Rightarrow \eta_t$ 鞅差 (martingale difference)
- Use the ARMA form for the squared series a_t^2 to understand properties of GARCH models, e.g. moment equations, forecasting, etc.

Focus on a GARCH(1,1) model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- Weak stationary: $0 \leq \alpha_1, \beta_1 \leq 1, (\alpha_1 + \beta_1) < 1.$
  其中 $(\alpha_1 + \beta_1) < 1$ 的條件保證 a_t 的 unconditional variance 存在。
- Volatility clusters 的效應
- Heavy tails: if $1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2 > 0$, then

$$\frac{E(a_t^4)}{[E(a_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$

GARCH 模型的 kurtosis

- 令 K_ϵ 代表 ϵ_t 的 excess kurtosis，則

$$E(\epsilon_t^4) = K_\epsilon + 3$$

- $$Var(a_t) = E(\sigma_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

- $$\begin{aligned} E(a_t^4) &= (K_\epsilon + 3)E(\sigma_t^4) \\ \sigma_t^4 &= \alpha_0^2 + \alpha_1^2 a_{t-1}^4 + \beta_1^2 \sigma_{t-1}^4 + 2\alpha_0\alpha_1 a_{t-1}^2 \\ &\quad + 2\alpha_0\beta_1\sigma_{t-1}^2 + 2\alpha_1\beta_1\sigma_{t-1}^2 a_{t-1}^2 \end{aligned}$$

-

$$E(\sigma_t^4) = \alpha_0^2 + \alpha_1^2(K_\epsilon + 3)E(\sigma_{t-1}^4) + \beta_1^2 E(\sigma_{t-1}^4) + 2\alpha_0\alpha_1 E(a_{t-1}^2) + 2\alpha_0\beta_1 E(\sigma_{t-1}^2) + 2\alpha_1\beta_1 E(\sigma_{t-1}^4\epsilon_{t-1}^2) = \frac{\alpha_0^2(1+\alpha_1+\beta_1)}{[1-(\alpha_1+\beta_1)][1-\alpha_1^2(K_\epsilon+2)-(\alpha_1+\beta_1)^2]},$$

if $1 > \alpha_1 + \beta_1 > 0$ and $1 - \alpha_1^2(K_\epsilon + 2) - (\alpha_1 + \beta_1)^2 > 0.$

$$K_a = \frac{E(a_t^4)}{(E(a_t^2))^2} - 3 = \frac{(K_\epsilon + 3)[1-(\alpha_1+\beta_1)^2]}{1 - 2\alpha_1^2 - (\alpha_1+\beta_1)^2 - K_\epsilon\alpha_1^2} - 3$$

- 當 $K_\epsilon = 0$，e.q $\epsilon_t \sim$N(0,1)，則

$$K_a^{(g)} = \frac{6\alpha_1^2}{1 - 2\alpha_1^2 - (\alpha_1+\beta_1)^2}.$$

- 故當 $1-2\alpha_1^2-(\alpha_1+\beta_1)^2>0$ 時，GARCH(1,1) 有 heavy tail。
- 當 $\alpha_1=0$ 時，GARCH(1,1) 沒有 heavy tail。
- 當 ϵ_t 不是 normal，則將 K_a 化簡成可得

$$K_a=\frac{K_\epsilon+K_a^{(g)}+\frac{5}{6}K_\epsilon K_a^{(g)}}{1-\frac{1}{6}K_\epsilon K_a^{(g)}}$$

- 當 ϵ_t 為 standardized $t(\nu)$ 時，$E(\epsilon_t^4)=\frac{6}{\nu-4}+3 \Rightarrow K_\epsilon=\frac{6}{\nu-4}$ for $\nu>4$.
- 因此當假設 ν 為已知時，我們通常假設 $\nu=5$ 。此時

$$K_a=\frac{6+(\nu+1)K_a^{(g)}}{\nu-4-K_a^{(g)}}$$

if $1-2\alpha_1^2\frac{\nu-1}{\nu-4}-(\alpha_1+\beta_1)^2>0$.

- For 1-step ahead forecast,

$$\sigma_t^2(1) = \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_t^2$$

- For multi-step ahead forecasts,

$$\begin{aligned}\sigma_t^2(h) &= E_t(\sigma_{t+h}^2) \\ &= E_t(\alpha_0 + \alpha_1 a_{t+h-1}^2 + \beta_1 \sigma_{t+h-1}^2) \\ &= \alpha_0 + \alpha_1 E_t(\sigma_{t+h-1}^2(\epsilon_{t+h-1}^2 - 1)) + (\alpha_1 + \beta_1) E_t(\sigma_{t+h-1}^2) \\ &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2(h-1), \ h > 1\end{aligned}$$

where the last equality is by $E_t(\varepsilon_{t+h-1}^2 - 1) = 0$.

- This result is exactly the same as that of an ARMA(1,1) model with AR polynomial $1 - (\alpha_1 + \beta_1)$B.

- $\sigma_t^2(h) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2(h-1), \; h > 1$
- By repeated substitutions, we obtain

$$\sigma_t^2(h) = \frac{\alpha_0[1-(\alpha_1+\beta_1)^{h-1}]}{1-\alpha_1-\beta_1} + (\alpha_1+\beta_1)^{h-1}\sigma_t^2(1)$$

  Therefore

$$\sigma_t^2(h) \to \frac{\alpha_0}{1-\alpha_1-\beta_1} \text{ as } h \to \infty$$

  provided $\alpha_1 + \beta_1 < 1$.
- Consequently, the multistep ahead volatility forecasts of a GARCH(1,1) model converges to the unconditional variance of a_t as the forecast horizon increases to infinity provided that $Var(a_t)$ exists.

Example

- Consider the monthly log returns of Intel stock (1973∼2003 的月報酬率).
- We fit an ARCH(1) model to the series. However, the Ljung-Box statistics for the squared standardized residuals show $Q(12) = 35.42$ with a p value of 0.0004. The volatility equation is inadequate at the 1% level.
- We fit a a GARCH$(1, 1)$ model to the series. The volatility equation is adequate at the 5% level.

```
  % The next command fits a GARCH(1,1) model
> m2=garchFit(intc~garch(1,1),data=intc,trace=F)
> summary(m2) % output edited.
Coefficient(s):
        mu        omega       alpha1       beta1
0.01073352  0.00095445  0.08741989  0.85118414

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.0107335   0.0055289    1.941  0.05222 .
omega   0.0009544   0.0003989    2.392  0.01674 *
alpha1  0.0874199   0.0269810    3.240  0.00120 **
beta1   0.8511841   0.0393702   21.620  < 2e-16 ***
---
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  165.5740  0
 Shapiro-Wilk Test  R    W      0.9712087 1.626824e-07
 Ljung-Box Test     R    Q(10)  8.267633  0.6027128
 Ljung-Box Test     R    Q(15)  14.42612  0.4934871
 Ljung-Box Test     R    Q(20)  15.13331  0.7687297
 Ljung-Box Test     R^2  Q(10)  0.9891848 0.9998363
 Ljung-Box Test     R^2  Q(15)  11.36596  0.7262473
 Ljung-Box Test     R^2  Q(20)  12.68143  0.8906302
 LM Arch Test       R    TR^2   10.70199  0.5546164
```

Exmple of GARCH model:

- Monthly excess returns of S&P 500 index starting from 1926 for 792 observations.
- excess returns (超額報酬率): r_t 。
- e.g. The payoff of short selling treasury bill and long S&P 500 index. There is no need for net initial investment. 賣空國庫券，買進 S&P 500 的報酬，此種操作無須任何起始的投資金額 ce
- The ACF of the excess returns r_t see Fig3.7(a)，and the PACF see Fig3.7(b).
- The following is the fitted MA(3) model,

$$r_t = 0.062 + a_t + 0.0944a_{t-1} - 0.1407a_{t-3}, \quad \hat{\sigma}_a = 0.0576.$$

All the model coefficients are significant at 5 % level.
- For easy interpretation (為了方便說明), we use a Gaussian AR(3) model to fit the excess return data r_t (recall that an invertible MA process has an AR(∞) representation) :

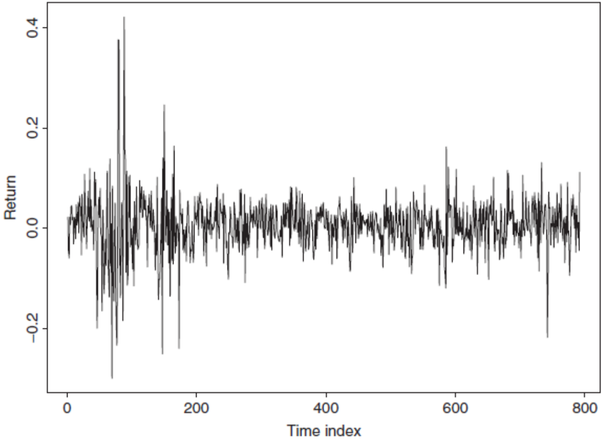


Figure 3.6 Time series plot of monthly excess returns of S&P 500 index from 1926 to 1991.

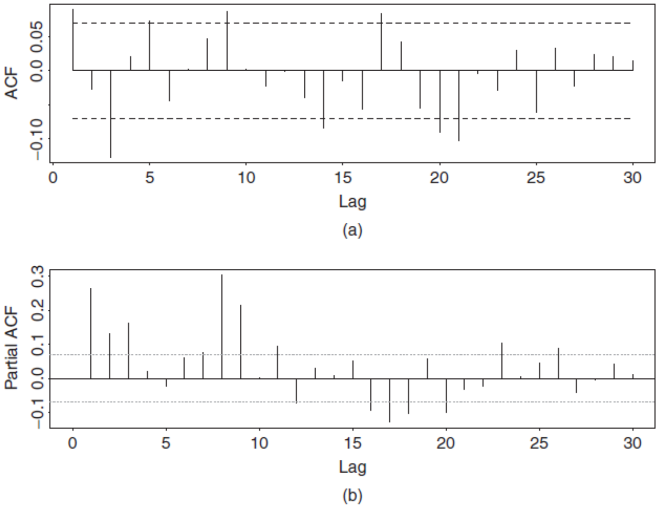


Figure 3.7 (a) Sample ACF of monthly excess returns of S&P 500 index and (b) sample PACF of squared monthly excess returns. Sample period is from 1926 to 1991.

- Key feature of the data is that the PACF of r_t^2 shows strong linear dependence.
- For the GARCH effects, use the joint estimation of AR(3)+GARCH$(1,1)$ model (R output):

$$r_t = 0.0075 + 0.0322 r_{t-1} - 0.0304 r_{t-2} - 0.0108 r_{t-3} + a_t$$
$$\sigma_t^2 = 0.000079 + 0.8552\sigma_{t-1}^2 + 0.1218 a_{t-1}^2.$$

- 此模型所對應的

$$Var(a_t) = \frac{0.000079}{1 - 0.8552 - 0.1218} = 0.00343$$

close to the expected value (from the Gaussian AR(3) model in the previous page, $\hat{\sigma}_a^2 = 0.0576^2 = 0.0033$).

- 對 S&P 500 index 的超額報酬率所配適的 AR(3)+GARCH(1,1) model，其 mean equation 中的 AR 參數皆不顯著。
- 考慮一個 simplified model：

$$r_t = 0.00745 + a_t, \ \sigma_t^2 = 0.00008 + 0.1223 a_{t-1}^2 + 0.8544 \sigma_{t-1}^2.$$

- For the estimated volatility process 見 Fig. 3.8.
- Model checking:
  For $\tilde{a}_t$: Q(10) = 11.22(0.34) and Q(20) = 24.29(0.23).
  For $\tilde{a}_t^2$: Q(10) = 9.92(0.27) and Q(20) = 16.74(0.54).
  $\Rightarrow$ GARCH(1,1) model is adequate.
- Note that in the fitted GARCH(1,1) model

$$\hat{\alpha}_1 + \hat{\beta}_1 = 0.9767 \approx 1.$$

實務中常觀察到這種現象，因而提出 IGARCH model (強制條件 $\alpha_1 + \beta_1 = 1$)。

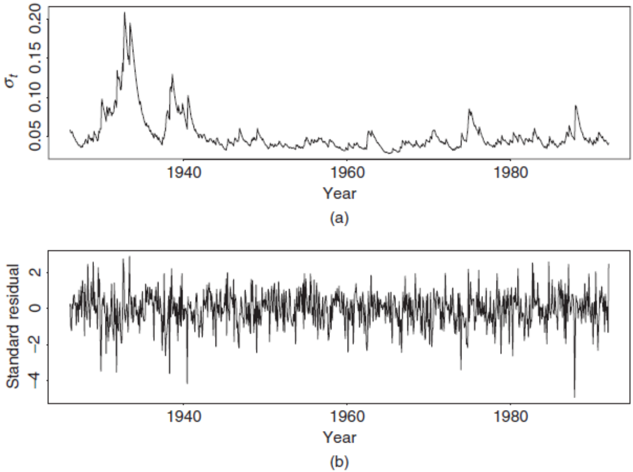


Figure 3.8 (a) Time series plot of estimated volatility (σ_t) for monthly excess returns of S&P 500 index and (b) standardized shocks of monthly excess returns of S&P 500 index. Both plots are based on GARCH(1,1) model in Eq. (3.19).

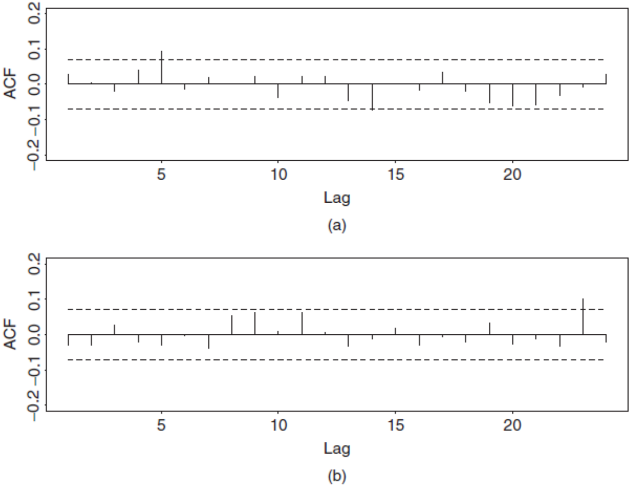


Figure 3.9 Model checking of GARCH(1,1) model in Eq. (3.19) for monthly excess returns of S&P 500 index: (a) Sample ACF of standardized residuals and (b) sample ACF of the squared standardized residuals.

- Forecast: 1-step ahead forecast:

$$\sigma_h^2(1) = 0.00008 + 0.1223a_h^2 + 0.8544\sigma_h^2$$

通常 σ_0^2 可假設為 0 或是 $Var(a_t)$.

- The forecast origin is $h = 792$, which corresponds to December, 1991.

Horizon	1	2	3	4	5	∞
Return	.0075	.0075	.0075	.0075	.0075	.0075
Volatility	.0538	.0539	.0540	.0542	.0543	.0586

- In the table, volatility denotes conditional standard deviation. Recall $\sqrt{Var(a_t)} = \sqrt{0.00342} = 0.0585$

R Commands Used in Example 3.3

```
> library(fGarch)
> sp5=scan(file='sp500.txt') % Load data
> plot(sp5,type='l')
% Below, fit an AR(3)+GARCH(1,1) model.
> m1=garchFit(~arma(3,0)+garch(1,1),data=sp5,trace=F)
> summary(m1)
% Below, fit a GARCH(1,1) model with Student-t distribution.
> m2=garchFit(~garch(1,1),data=sp5,trace=F,cond.dist="std")
> summary(m2)
% Obtain standardized residuals.
> stresi=residuals(m2,standardize=T)
> plot(stresi,type='l')
> Box.test(stresi,10,type='Ljung')
> predict(m2,5)
```

- Compare the Splus result with that of R.

$$\text{SPlus} \begin{cases} r_t = 0.0076 + a_t \\ \sigma_t^2 = 0.00008 + 0.121 a_{t-1}^2 + 0.851 \sigma_{t-1}^2 \end{cases}$$

$$\text{R} \begin{cases} r_t = 0.00745 + a_t \\ \sigma_t^2 = 0.00008 + 0.1223 a_{t-1}^2 + 0.8544 \sigma_{t-1}^2 \end{cases}$$

- Turn to Student-t innovation.

$$r_t = 0.0085 + a_t,$$
$$\sigma_t^2 = 0.000125 + 0.113 a_{t-1}^2 + 0.842 \sigma_{t-1}^2,$$

where the estimated degrees of freedom is 6.99.

Forecasting evaluation

- Not easy to do; see Andersen and Bollerslev (1998).
- 由於 volatility 並不是可以直接觀察，因此不容易比較不同 volatility model 的預測表現。
- 有些人採取 out-of sample 的預測，並利用比較

$$\sigma_t^2(h) \text{ and } a_{t+h}^2$$

來評估一個 volatility 模型的表現。然而他們總是發現 $\sigma_t^2(h)$ 與 a_{t+h}^2 的 correlation 偏低，故此法並不適合用來評估 volatility model 準確性或預測能力。

S&P 500 指數的月超額報酬, 1926–1991, 共 792 筆資料

```
> #load the package "rugarch"
> library(rugarch)
> sp5=scan(file="sp500.dat")
Read 792 items
> sp5=ts(sp5,frequency=12,start=c(1926,1))
> plot(sp5,type="l")
> m1=arima(sp5,order=c(3,0,0))
> m1

Call:
arima(x = sp5, order = c(3, 0, 0))

Coefficients:
         ar1      ar2      ar3  intercept
      0.0890  -0.0238  -0.1229     0.0062
s.e.  0.0353   0.0355   0.0353     0.0019

sigma^2 estimated as 0.00333:  log likelihood = 1135.25,  aic = -2260.5
```

In the following codes "sGARCH"= standard GARCH.

```
> #ARMA(3,0) model
> m.model=list(armaOrder=c(3,0),include.mean=T)
> #GARCH(1,1) model
> var.model=list(model="sGARCH",garchOrder=c(1,1))
> #The conditional density to use for the innovations.
> dist.model="norm"
> #model specification
> model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
> #fit a ARMA(3,0)+GARCH(1,1) model
> fit=ugarchfit(data=sp5,spec=model)
> fit
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(3,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.007475    0.001524  4.90313 0.000001
ar1     0.032197    0.038410  0.83825 0.401893
ar2    -0.030416    0.038469 -0.79064 0.429153
ar3    -0.010847    0.037658 -0.28804 0.773318
omega   0.000079    0.000028  2.80335 0.005058
alpha1  0.121842    0.022268  5.47153 0.000000
beta1   0.855169    0.021941 38.97568 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.007475    0.001696  4.40829 0.000010
ar1     0.032197    0.039454  0.81607 0.414457
ar2    -0.030416    0.038928 -0.78134 0.434605
ar3    -0.010847    0.038021 -0.28529 0.775426
omega   0.000079    0.000033  2.39744 0.016510
alpha1  0.121842    0.028092  4.33728 0.000014
beta1   0.855169    0.026321 32.48937 0.000000
```

```
LogLikelihood : 1269.843

Information Criteria
------------------------------------

Akaike       -3.1890
Bayes        -3.1477
Shibata      -3.1892
Hannan-Quinn -3.1731
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                  0.0004387  0.9833
Lag[2*(p+q)+(p+q)-1][8] 5.0511722  0.1812
Lag[4*(p+q)+(p+q)-1][14] 8.3030394  0.3214
d.o.f=3
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                       statistic p-value
Lag[1]                    0.6509  0.4198
Lag[2*(p+q)+(p+q)-1][5]   1.8139  0.6630
Lag[4*(p+q)+(p+q)-1][9]   3.7001  0.6400
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
             Statistic Shape Scale P-Value
ARCH Lag[3]     0.6235 0.500 2.000  0.4298
ARCH Lag[5]     1.4034 1.440 1.667  0.6181
ARCH Lag[7]     2.0255 2.315 1.543  0.7119
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.3635
Individual Statistics:
mu     0.09195
ar1    0.05486
ar2    0.08974
ar3    0.12219
omega  0.20301
alpha1 0.27398
beta1  0.17087

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value      prob sig
Sign Bias           3.1857 0.0015008 ***
Negative Sign Bias  0.8778 0.3802934
Positive Sign Bias  0.5925 0.5536841
Joint Effect       18.1518 0.0004093 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     28.00     0.083429
2    30     58.15     0.001043
3    40     51.43     0.087726
4    50     58.38     0.168700
```

```
> #fit a ARMA(0,0)+GARCH(1,1) model
> m.model=list(armaOrder=c(0,0),include.mean=T)
> var.model=list(model="sGARCH",garchOrder=c(1,1))
> dist.model="norm"
> model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
> fit=ugarchfit(data=sp5,spec=model)
> fit
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu       0.00745    0.001538   4.8450 0.000001
omega    0.00008    0.000028   2.8287 0.004674
alpha1   0.12226    0.022102   5.5315 0.000000
beta1    0.85435    0.021811  39.1708 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu       0.00745    0.001717   4.3382 0.000014
omega    0.00008    0.000034   2.3921 0.016754
alpha1   0.12226    0.028162   4.3412 0.000014
beta1    0.85435    0.026420  32.3374 0.000000
```

```
LogLikelihood : 1269.455

Information Criteria
------------------------------------

Akaike       -3.1956
Bayes        -3.1720
Shibata      -3.1956
Hannan-Quinn -3.1865
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                          statistic p-value
Lag[1]                       0.6270  0.4285
Lag[2*(p+q)+(p+q)-1][2]      0.6322  0.6347
Lag[4*(p+q)+(p+q)-1][5]      2.7319  0.4582
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                          statistic p-value
Lag[1]                        0.622  0.4303
Lag[2*(p+q)+(p+q)-1][5]       1.867  0.6502
Lag[4*(p+q)+(p+q)-1][9]       3.671  0.6450
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.6535 0.500 2.000  0.4189
ARCH Lag[5]    1.3390 1.440 1.667  0.6356
ARCH Lag[7]    1.9550 2.315 1.543  0.7269
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.0288
Individual Statistics:
mu     0.08925
omega  0.20221
alpha1 0.27132
beta1  0.16481

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:        1.07 1.24 1.6
Individual Statistic:   0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value      prob sig
Sign Bias           3.1028 0.0019855 ***
Negative Sign Bias  0.9191 0.3583038
Positive Sign Bias  0.7085 0.4788455
Joint Effect       17.6096 0.0005294 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     30.73      0.04324
2    30     40.27      0.07958
3    40     48.40      0.14373
4    50     61.41      0.10989
```

```
> sd(residuals(fit))
[1] 0.0584588
> var(residuals(fit))
            [,1]
[1,] 0.003417431
> resi=residuals(fit,standardize=T)
> Box.test(resi,lag=10,type="Ljung")

        Box-Ljung test

data:  resi
X-squared = 11.216, df = 10, p-value = 0.3409

> Box.test(resi,lag=20,type="Ljung")

        Box-Ljung test

data:  resi
X-squared = 24.292, df = 20, p-value = 0.2299

> Box.test(resi^2,lag=10,type="Ljung",fitdf=2)

        Box-Ljung test

data:  resi^2
X-squared = 9.9181, df = 8, p-value = 0.2708

> Box.test(resi^2,lag=20,type="Ljung",fitdf=2)

        Box-Ljung test

data:  resi^2
X-squared = 16.735, df = 18, p-value = 0.5414
```

```
> #forecast
> forc=ugarchforecast(fit, n.ahead=12)
> forc

*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 12
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=十二月 1991]:
      Series   Sigma
T+1  0.00745 0.05381
T+2  0.00745 0.05392
T+3  0.00745 0.05404
T+4  0.00745 0.05415
T+5  0.00745 0.05426
T+6  0.00745 0.05436
T+7  0.00745 0.05447
T+8  0.00745 0.05457
T+9  0.00745 0.05467
T+10 0.00745 0.05476
T+11 0.00745 0.05486
T+12 0.00745 0.05495
```

```
> #Turn to Student-t innovation
> m.model=list(armaOrder=c(0,0),include.mean=T)
> var.model=list(model="sGARCH",garchOrder=c(1,1))
> dist.model="std"
> model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
> fit=ugarchfit(data=sp5,spec=model)
> fit
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.008455    0.001515   5.5809 0.000000
omega   0.000125    0.000045   2.7494 0.005970
alpha1  0.113294    0.027031   4.1912 0.000028
beta1   0.842212    0.031957  26.3547 0.000000
shape   6.991570    1.677722   4.1673 0.000031

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.008455    0.001659   5.0974 0.000000
omega   0.000125    0.000039   3.1950 0.001398
alpha1  0.113294    0.030222   3.7487 0.000178
beta1   0.842212    0.025882  32.5398 0.000000
shape   6.991570    1.697841   4.1179 0.000038
```

```
LogLikelihood : 1283.406

Information Criteria
------------------------------------

Akaike       -3.2283
Bayes        -3.1988
Shibata      -3.2284
Hannan-Quinn -3.2170
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                          statistic p-value
Lag[1]                       0.6869  0.4072
Lag[2*(p+q)+(p+q)-1][2]      0.6921  0.6099
Lag[4*(p+q)+(p+q)-1][5]      2.8643  0.4324
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                          statistic p-value
Lag[1]                       0.2519  0.6158
Lag[2*(p+q)+(p+q)-1][5]      1.3619  0.7738
Lag[4*(p+q)+(p+q)-1][9]      3.2184  0.7224
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
             Statistic Shape Scale P-Value
ARCH Lag[3]     0.8794 0.500 2.000  0.3484
ARCH Lag[5]     1.3423 1.440 1.667  0.6347
ARCH Lag[7]     1.7507 2.315 1.543  0.7699
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.3011
Individual Statistics:
mu     0.13494
omega  0.18838
alpha1 0.48541
beta1  0.26807
shape  0.07581

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.28 1.47 1.88
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value      prob sig
Sign Bias           3.0829 0.0021218 ***
Negative Sign Bias  0.6407 0.5218863
Positive Sign Bias  0.3824 0.7022535
Joint Effect       16.4849 0.0009018 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     19.36       0.4337
2    30     35.20       0.1981
3    40     46.99       0.1778
4    50     51.31       0.3833
```

Weighted Ljung-Box Test & Weighted ARCH LM Test:

Fisher, T. J. and Gallagher, C. M. (2012). New Weighted Portmanteau Statistics for Time Series Goodness-of-Fit Testing. *Journal of the American Statistical Association*, **107**, 777-787.

Weighted Ljung-Box Test (I)

- Consider the problem discussed in Srivastava (2005) fortesting for an identity covariance matrix. Using his approach, let $\boldsymbol{\Sigma}_m$ denote the probability limit of matrix

$$\hat{\mathbf{R}}_m = \begin{pmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_m \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_{m-1} \\ \vdots & \ldots & \ddots & \vdots \\ \hat{\rho}_m & \cdots & \hat{\rho}_1 & 1 \end{pmatrix}$$

  where $\hat{\rho}_k$ is the lag k sample autocorrelation function.
- To test for an iid correlation structure, we test $H_0 : \boldsymbol{\Sigma}_m = \mathbf{I}$, which is equivalent to testing if each eigenvalue of $\boldsymbol{\Sigma}_m$ is one.
- In practice, we can use the matrix $\hat{\mathbf{R}}_m$ and its eigenvalues λ_i, $i = 1, \ldots, m+1$ to test H_0.

Weighted Ljung-Box Test (II)

- Note that

$$\frac{1}{m+1}\sum_{k=1}^{m+1}(\lambda_k - 1)^2 \geq 0$$

with equality when each $\lambda_k = 1$. An equivalent statement after some algebra is

$$\frac{1}{m+1}\mathrm{tr}(\hat{\mathbf{R}}_m^2) - 1 \geq 0.$$

Weighted Ljung-Box Test (III)

- From calculating the trace of the matrix $\hat{\mathbf{R}}_m^2$, we may consider testing the hypothesis based on the inequality

$$\sum_{k=1}^{m} \frac{2(m-k+1)}{m+1} \hat{\rho}_k^2 \geq 0.$$

- Noting the inequality will not change if each side is multiplied by $(m+1)/2m$, and replacing the squared autocorrelations with their standardized residual counterparts leads to the Weighted Ljung–Box statistic

$$\tilde{Q}_W = n(n+2) \sum_{k=1}^{m} \frac{m-k+1}{m} \frac{\hat{\rho}_k^2}{n-k}.$$

Recall the Ljung–Box statistic

$$Q(m) = n(n+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k}$$

Weighted ARCH LM Test

- Li and Mak (1994) showed that for an ARCH(b) model (GARCH($b,0$)), the residual autocorrelations $\hat{\rho}_k(\hat{a}_t^2/\hat{\sigma}_t^2)$ for $k = b+1,\ldots,m$ are asymptotically iid standard normal. They proposed the modified statistic

$$L(b,m) = n \sum_{k=b+1}^{m} \hat{\rho}_k^2(\hat{a}_t^2/\hat{\sigma}_t^2).$$

- Fisher and Gallagher (2012) recommend building a test based on the simpler Li–Mak test for fitted ARCH(b) processes. Consider the Weighted ARCH LM statistic:

$$L_W(b,m) = n \sum_{k=b+1}^{m} \frac{m-k+(b+1)}{m} \hat{\rho}_k^2(\hat{a}_t^2/\hat{\sigma}_t^2).$$

Nyblom stability test (I)

- Reference:
  Nyblom, J. (1989). Testing for the Constancy of Parameters Over Time. *Journal of the American Statistical Association*, **84**, 223-230.

Nyblom stability test (II)

- Consider the linear regression model with k variables

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + \epsilon_t, \quad t = 1, \ldots, n,$$

  where $\mathbf{x}_t = (x_{t,1}, \ldots, x_{t,k})'$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$.
- The time varying parameter alternative model assumes

$$\boldsymbol{\beta} = \boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t, \quad \eta_{t,i} \sim (0, \sigma^2_{\eta_i}),\ i = 1, \ldots, k.$$

- Individual coefficient tests:

$$H_0 : \beta_i \text{ is constant } (\sigma^2_{\eta_i} = 0) \qquad H_1 : \beta_i \text{ is not constant } (\sigma^2_{\eta_i} \neq 0)$$

- Joint test for all coefficients:

$$H_0 : \boldsymbol{\beta} \text{ is constant} \qquad H_1 : \boldsymbol{\beta} \text{ is not constant}$$

Bias Test (I)

- Reference:
  Engle, R.F. and Ng, V.K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, **48**, 1749-1778.
- This tests the presence of leverage effects in the standardized residuals (to capture possible misspecification of the GARCH model), by regressing the squared standardized residuals on lagged negative and positive shocks as follows:

$$\hat{\epsilon}_t^2 = c_0 + c_1 I_{\{\hat{a}_{t-1}<0\}} + c_2 I_{\{\hat{a}_{t-1}<0\}} \hat{a}_{t-1} + c_3 I_{\{\hat{a}_{t-1}\geq 0\}} \hat{a}_{t-1} + u_t$$

  where I is the indicator function and $\hat{a}_t$ is the estimated residuals from the GARCH process.

Bias Test (II)

$$\hat{\epsilon}_t^2 = c_0 + c_1 I_{\{\hat{a}_{t-1}<0\}} + c_2 I_{\{\hat{a}_{t-1}<0\}}\hat{a}_{t-1} + c_3 I_{\{\hat{a}_{t-1}\geq 0\}}\hat{a}_{t-1} + u_t$$

- Sign Bias Test:
  - $H_0 : c_1 = 0 \quad H_1 : c_1 \neq 0$
  - This test examines the impact of positive and negative return shocks on volatility not predicted by the model under consideration.
- Negative Size Bias Test:
  - $H_0 : c_2 = 0 \quad H_1 : c_2 \neq 0$
  - It focuses on the different effects that large and small negative return shocks have on volatility which are not predicted by the volatility model.

Bias Test (III)

$$\hat{\epsilon}_t^2 = c_0 + c_1 I_{\{\hat{a}_{t-1}<0\}} + c_2 I_{\{\hat{a}_{t-1}<0\}} \hat{a}_{t-1} + c_3 I_{\{\hat{a}_{t-1}\geq 0\}} \hat{a}_{t-1} + u_t$$

- Positive Size Bias Test:
  - $H_0 : c_3 = 0 \quad H_1 : c_3 \neq 0$
  - It focuses on the different impacts that large and small positive return shocks may have on volatility, which are not explained by the volatility model.
- Conduct these tests jointly

$$H_0 : c_1 = c_2 = c_3 = 0 \quad H_1 : \text{not } H_0$$

Adjusted Pearson Goodness-of-Fit Test (I)

- Reference:
  Vlaar, P. and Palm, F.C. (1993). The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps. *Journal of Business & Economic Statistics*, **11**, 351-359.

Adjusted Pearson Goodness-of-Fit Test (II)

- For i.i.d. observations, it can be shown that

$$\sum_{i=1}^{g} \frac{(n_i - En_i)^2}{En_i} \sim \chi^2(g-1)$$

  where g is the number of groups and n_i is the number of observations in each group.
- A problem with this test statistic for models with a time-varying variance and jumps is that their residuals are neither identically nor independently distributed.
- This problem can be solved by reclassifying the residuals.
- Instead of classifying the residuals according to their value, we calculate the probability of observing a value smaller than the residual. These probabilities should be identically uniformly distributed between 0 and 1.

Adjusted Pearson Goodness-of-Fit Test (III)

- The grouping mechanism for equally sized groups thus becomes

$$n_i = \sum_{t=1}^{n} I_{t,i}$$

where

$$I_{t,i} = \begin{cases} 1, & \text{if } (i-1)/g < F(X_t, \hat{\phi}) \leq i/g \\ 0, & \text{otherwise} \end{cases}$$

where $F(X_t, \hat{\phi})$ represents the value of the cumulative distribution function given the data and the estimated parameters.

IGARCH

- If the AR polynomial of the GARCH representation has a unit root, then we have an IGARCH model.
- A key feature of IGARCH models is that the impact of past squared shocks $\eta_{t-i} = a_{t-i}^2 - \sigma_{t-i}^2$ for $i > 0$ on a_t^2 is persistent.
- An IGARCH(1,1) model:

$$a_t = \sigma_t \epsilon_t, \;\; \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2,$$

$0 < \beta_1 < 1.$

- For the monthly excess returns of the S&P 500 index, an estimated IGARCH(1,1) model is

$$\begin{aligned} r_t &= 0.0074 + a_t,\ a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 0.000051 + 0.857\sigma_{t-1}^2 + 0.143 a_{t-1}^2 \end{aligned}$$

- The parameters estimates are close to those of the GARCH(1,1) model shown before.
- But the unconditional variance of a_t is not defined for the IGARCH(1,1) model, which seems hard to justify for an excess return series.
- From a theoretical point of view, the IGARCH phenomena might be caused by occasional level shifts in volatility.

IGARCH Forecast

- For the l-step forecast of an IGARCH(1,1) model,

$$\sigma_h^2(l) = \sigma_h^2(1) + (l-1)\alpha_0, \quad l \leq 1.$$

  where h is the forecast origin.
- Effect of $\sigma_h^2(1)$ on future volatilities is persistent, and the volatility forecasts form a straight line with slope α_0.
- See Nelson (1990) for more info. (研究 IGARCH 模型中 σ_t^2 的機率性質。)

- Special case $\alpha_0 = 0 \Rightarrow \sigma_h^2(l) = \sigma_h^2(1),\ \forall\ h \text{ and } l.$ 此時

$$E(\sigma_{t+1}^2 | F_{t-1}) = (1-\beta_1)E(a_t^2 | F_{t-1}) + \beta_1 \sigma_t^2 = \sigma_t^2$$

$\Rightarrow \{\sigma_{t+1}^2, F_{t-1}\}$ 為 martingale 。
This special IGARCH model is the volatility model used in RiskMetrics to VaR calculation.

- This special model is also an exponential smoothing model for $\{a_t^2\}$:

$$\begin{aligned}
\sigma_t^2 &= (1-\beta_1)a_{t-1}^2 + \beta_1\sigma_{t-1}^2 \\
&= (1-\beta_1)a_{t-1}^2 + \beta_1[(1-\beta_1)a_{t-2}^2 + \beta_1\sigma_{t-2}^2] \\
&= (1-\beta_1)a_{t-1}^2 + (1-\beta_1)\beta_1 a_{t-2}^2 + \beta_1^2\sigma_{t-2}^2 \\
&= (1-\beta_1)(a_{t-1}^2 + \beta_1 a_{t-2}^2 + \beta_1^2 a_{t-3}^2 + \cdots),
\end{aligned}$$

which is the well known exponential smoothing formulation with β_1 being the discounting factor.

Example

- An IGARCH(1,1) model for the monthly excess returns of S&P500 index from 1926 to 1991 is given below via R.

$$r_t = 0.0074 + a_t,\ a_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = 0.000051 + 0.143 a_{t-1}^2 + 0.857 \sigma_{t-1}^2$$

An IGARCH(1,1) model for the monthly excess returns of S&P 500 index from 1926 to 1991.

```
> #fit a ARMA(0,0)+IGARCH(1,1) model
> m.model=list(armaOrder=c(0,0),include.mean=T)
> var.model=list(model="iGARCH",garchOrder=c(1,1))
> dist.model="norm"
> model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
> fit=ugarchfit(data=sp5,spec=model)
> #forecast
> forc=ugarchforecast(fit, n.ahead=12)
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : iGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.007417    0.001525   4.8621 0.000001
omega   0.000051    0.000018   2.9238 0.003458
alpha1  0.142951    0.021443   6.6667 0.000000
beta1   0.857049          NA       NA       NA

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.007417    0.001587   4.6726 0.000003
omega   0.000051    0.000019   2.6913 0.007118
alpha1  0.142951    0.024978   5.7230 0.000000
beta1   0.857049          NA       NA       NA
```

```
LogLikelihood : 1268.238

Information Criteria
------------------------------------

Akaike       -3.1950
Bayes        -3.1773
Shibata      -3.1951
Hannan-Quinn -3.1882
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.5265  0.4681
Lag[2*(p+q)+(p+q)-1][2]    0.5304  0.6795
Lag[4*(p+q)+(p+q)-1][5]    2.5233  0.5009
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.166  0.2803
Lag[2*(p+q)+(p+q)-1][5]     2.672  0.4702
Lag[4*(p+q)+(p+q)-1][9]     4.506  0.5054
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.4608 0.500 2.000  0.4972
ARCH Lag[5]    1.3891 1.440 1.667  0.6219
ARCH Lag[7]    2.2325 2.315 1.543  0.6682
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  0.4498
Individual Statistics:
mu      0.08818
omega   0.14331
alpha1  0.22409

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          0.846 1.01 1.35
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value     prob sig
Sign Bias           3.2677 0.001131 ***
Negative Sign Bias  1.3431 0.179638
Positive Sign Bias  0.9411 0.346937
Joint Effect       19.5642 0.000209 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     31.84      0.03259
2    30     38.61      0.10944
3    40     50.73      0.09888
4    50     57.49      0.18957
```

```
*---------------------------------*
*       GARCH Model Forecast      *
*---------------------------------*
Model: iGARCH
Horizon: 12
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=十二月 1991]:
       Series   Sigma
T+1  0.007417 0.05595
T+2  0.007417 0.05641
T+3  0.007417 0.05686
T+4  0.007417 0.05731
T+5  0.007417 0.05775
T+6  0.007417 0.05820
T+7  0.007417 0.05864
T+8  0.007417 0.05907
T+9  0.007417 0.05950
T+10 0.007417 0.05993
T+11 0.007417 0.06036
T+12 0.007417 0.06078
```

The GARCH-M model

- GARCH in Mean 的模型 (當 $E(r_t \mid \mathcal{F}_{t-1})$ 與 σ_t^2 有相關性時的模型)：

$$r_t = \mu + c\sigma_t^2 + a_t, \quad a_t = \sigma_t \epsilon_t,$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

  where c is referred to as risk premium (風險溢酬), which is expected to be positive.

- 假設 $\mu = 0$，在此模型下，$E(r_t r_{t+1}) = c^2 E(\sigma_t^2 \sigma_{t+1}^2) \neq 0$ 。
- 因此加入風險溢酬項，可用來解釋 $\{r_t\}$ 具有一步相關的現象。
- 文獻中，亦有考慮

$$r_t = \mu + c\sigma_t + a_t \quad \text{和} \quad r_t = \mu + c\log(\sigma_t^2) + a_t$$

  的 GARCM-M 的模型。

Example

- A GARCH(1,1)-M model for the monthly excess returns of S&P 500 index from January 1926 to December 1991.
- The fitted model is

$$r_t = 0.0054 + 1.01\sigma_t^2 + a_t, \ \sigma_t^2 = 0.00008 + 0.123a_{t-1}^2 + 0.852\sigma_{t-1}^2$$

Std. err. of risk premium is 0.888.

- The estimated risk premium for the index return is positive but is not statistically significant at the 5% level.

An GARCH(1,1)-M model for S&P 500 index from January 1926 to December 1991.
In the following codes, "archm=TRUE, archpow=2" implies fitting GARCH-m model with σ_t^2.

```
#fit a GARCH(1,1)-M model
m.model=list(armaOrder=c(0,0),include.mean=T,archm=TRUE,archpow=2)
var.model=list(model="sGARCH",garchOrder=c(1,1))
dist.model="norm"
model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
fit=ugarchfit(data=sp5,spec=model)
#forecast
forc=ugarchforecast(fit, n.ahead=12)
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.005421    0.002360   2.2969 0.021623
archm   1.007797    0.888466   1.1343 0.256664
omega   0.000083    0.000029   2.8294 0.004663
alpha1  0.123118    0.022285   5.5248 0.000000
beta1   0.852274    0.022397  38.0526 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.005421    0.002440   2.2221 0.026276
archm   1.007797    0.891207   1.1308 0.258130
omega   0.000083    0.000035   2.3849 0.017082
alpha1  0.123118    0.027801   4.4286 0.000009
beta1   0.852274    0.027057  31.4994 0.000000
```

```
LogLikelihood : 1270.102

Information Criteria
------------------------------------

Akaike       -3.1947
Bayes        -3.1652
Shibata      -3.1948
Hannan-Quinn -3.1834
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.7246  0.3946
Lag[2*(p+q)+(p+q)-1][2]    0.7501  0.5869
Lag[4*(p+q)+(p+q)-1][5]    3.0242  0.4027
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.7587  0.3837
Lag[2*(p+q)+(p+q)-1][5]    1.8688  0.6497
Lag[4*(p+q)+(p+q)-1][9]    3.4841  0.6771
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.5049 0.500 2.000  0.4774
ARCH Lag[5]    1.0893 1.440 1.667  0.7066
ARCH Lag[7]    1.6863 2.315 1.543  0.7833
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.3267
Individual Statistics:
mu     0.06463
archm  0.03106
omega  0.21406
alpha1 0.26214
beta1  0.16478

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.28 1.47 1.88
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value      prob sig
Sign Bias            3.583 3.600e-04 ***
Negative Sign Bias   1.160 2.466e-01
Positive Sign Bias   0.621 5.348e-01
Joint Effect        22.181 5.982e-05 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     23.86      0.20163
2    30     41.11      0.06741
3    40     40.63      0.39861
4    50     56.86      0.20554
```

```
*---------------------------------*
*       GARCH Model Forecast      *
*---------------------------------*
Model: sGARCH
Horizon: 12
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=十二月 1991]:
       Series   Sigma
T+1  0.008344 0.05385
T+2  0.008356 0.05396
T+3  0.008367 0.05407
T+4  0.008378 0.05417
T+5  0.008389 0.05427
T+6  0.008399 0.05436
T+7  0.008410 0.05446
T+8  0.008420 0.05455
T+9  0.008430 0.05464
T+10 0.008439 0.05472
T+11 0.008448 0.05481
T+12 0.008458 0.05489
```

EGARCH

- The EGARCH model (exponential GARCH 模型, Nelson, 1991 提出)
- An EGARCH$(m,\ s)$ model:

$$a_t = \sigma_t \epsilon_t,\ \ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + \cdots + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B - \cdots - \alpha_m B^m} g(\epsilon_{t-1})$$

- Asymmetry in responses to $+$ $\&$ $-$ returns

$$g(\epsilon_t) = \theta \epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)],$$

with $E[g(\epsilon_t)] = 0$.
- 可視為 $\{\epsilon_t\}$ 與 $\{|\epsilon_t| - E|\epsilon_t|\}$ 的加權，即一種 weighted innovation; 其中 $\{\epsilon_t\}$ 與 $\{|\epsilon_t| - E|\epsilon_t|\}$ 兩者皆為 zero-mean 的 iid noise.

To see asymmetry of $g(\epsilon_t)$, rewrite it as

$$g(\epsilon_t) = \begin{cases} (\theta+\gamma)\epsilon_t - \gamma E(|\epsilon_t|) & if\ \epsilon_t \geq 0, \\ (\theta-\gamma)\epsilon_t - \gamma E(|\epsilon_t|) & if\ \epsilon_t < 0 \end{cases}$$

- $$\begin{cases} \epsilon_t \sim N(0,1) \ \Rightarrow E|\epsilon_t| = \sqrt{\frac{2}{\pi}} \\ \epsilon_t \sim Standardized\ t \ \Rightarrow E|\epsilon_t| = \frac{2\sqrt{\nu-2}\Gamma\left(\frac{\nu+1}{2}\right)}{(\nu-1)\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}} \end{cases}$$
- 當 $\gamma \neq 0$ 時，EGARCH 模型為非線性。

- Some features of EGARCH models:
  - uses logged conditional variance to relax the positiveness constraint
  - asymmetric responses
- Consider an EGARCH(1,1) model

$$a_t = \sigma_t \epsilon_t, \quad (1 - \alpha B) \ln(\sigma_t^2) = (1 - \alpha)\alpha_0 + g(\epsilon_{t-1})$$

where the subscript of α_1 is omitted.

- Under normality, $E(|\epsilon_t|) = \sqrt{2/\pi}$ and the model becomes

$$(1-\alpha B)\ln(\sigma_t^2) = \begin{cases} \alpha_* + (\theta+\gamma)\epsilon_{t-1} & \text{if } \epsilon_{t-1} \geq 0, \\ \alpha_* + (\theta-\gamma)\epsilon_{t-1} & \text{if } \epsilon_{t-1} < 0 \end{cases}$$

where

$$\begin{aligned} \alpha_* &= (1-\alpha)\alpha_0 - \sqrt{\frac{2}{\pi}}\gamma, \text{when } \epsilon_t \sim N(0,1), \\ \alpha_* &= (1-\alpha)\alpha_0 - \frac{2\sqrt{\nu-2}\,\Gamma\left(\frac{\nu+1}{2}\right)}{(\nu-1)\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}}\gamma, \text{when } \epsilon_t \sim \text{Standardized } t. \end{aligned}$$

- This is a nonlinear fun. similar to that of the threshold AR model of Tong (1978, 1990).

- Specifically, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\alpha} \exp(\alpha_*) \begin{cases} \exp[(\theta+\gamma)\frac{a_{t-1}}{\sigma_{t-1}}], & \text{if } a_{t-1} \geq 0, \\ \exp[(\theta-\gamma)\frac{a_{t-1}}{\sigma_{t-1}}], & \text{if } a_{t-1} < 0. \end{cases}$$

- The coefs $(\theta+\gamma)$ & $(\theta-\gamma)$ show the asymmetry in response to positive and negative a_{t-1}.
- The model is nonlinear if $\gamma \neq 0$. The parameter γ denotes the magnitude effect.
- The parameter θ is referred to as the leverage parameter, since θ shows the effect of the sign of a_{t-1}.
  Leverage effect which corresponds to a negative correlation between past returns and future volatility.
  槓桿效應：過去的報酬率與未來波動為負相關的效應。
- See Nelson (1991) for an example of EGARCH model.

Example of EGARCH: (text p.145) (Nelson, 1991)
Daily value-weight market index 的超額報酬 (excess return)，1962/7 ~ 1987/12，共有 6408 筆資料。

- The excess returns are obtained by removing monthly Treasury bill return from the value-weighted index returns.
- Treasury bill return was assumed to be constant for each calendar day within a given month.

$$r_t = \phi_0 + \phi_1 r_{t-1} + c\sigma_t^2 + a_t$$
$$\ln(\sigma_t^2) = \alpha_0 + \ln(1+\omega N_t) + \frac{1+\beta B}{1-\alpha_1 B - \alpha_2 B^2} g(\epsilon_{t-1})$$

- σ_t^2 is the conditional variance of a_t given F_{t-1}。
- $N_t =$ 從 $(t-1)$ 到 t 天，沒有交易的日數。
- $\epsilon_t \sim$ generalized error distribution。

Parameter Estimates

TABLE 3.3: Estimated AR(1)–EGARCH$(2,2)$ Model for Daily Excess Returns of Value-Weighted CRSP Market Index: July 1962-December 1987

Parameter	α_0	w	γ	α_1	α_2	β
Estimate	-10.06	0.183	0.156	1.929	-0.929	-0.978
Error	0.346	0.028	0.013	0.015	0.015	0.006
Parameter	θ	ϕ_0	ϕ_1	c	v	
Estimate	-0.118	$3.5 \cdot 10^{-4}$	0.205	-3.361	1.576	
Error	0.009	$9.9 \cdot 10^{-5}$	0.012	2.026	0.032	

結果顯示 risk premium $c < 0$，但不顯著 。

Example:

Monthly log returns of IBM stock from January 1926 to December 1997 for 864 observations.

- The following AR(1)-EGARCH(1,1) is obtained by RATS program:

$$r_t = 0.0108 + 0.0932 r_{t-1} + a_t, \quad a_t = \sigma_t \epsilon_t$$
$$\ln(\sigma_t^2) = -5.467 + \frac{g(\epsilon_{t-1})}{1 - 0.92386B}$$
$$g(\epsilon_{t-1}) = -0.049112\epsilon_{t-1} + 0.206452[|\epsilon_{t-1}| - \sqrt{2/\pi}]$$

- Model checking: $\widetilde{a}_t = \frac{a_t}{\sigma_t}$ (standardized residuals).
  For $\widetilde{a}_t$: $Q(10) = 6.65(0.67)$ and $Q(20) = 20.52(0.36)$
  For $\widetilde{a}_t^2$: $Q(10) = 3.37(0.91)$ and $Q(20) = 11.15(0.89)$

Discussion:

- Using $\sqrt{2/\pi} \approx 0.7979 \approx 0.8$, we obtain

$$\ln(\sigma_t^2) = -0.581 + 0.92386 \ln(\sigma_{t-1}^2) + \begin{cases} 0.1573\epsilon_{t-1} & if\ \epsilon_{t-1} \geq 0 \\ -0.2556\epsilon_{t-1} & if\ \epsilon_{t-1} < 0. \end{cases}$$

- Taking anti-log transformation, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\times 0.92386} e^{-0.581} \times \begin{cases} e^{0.1573\epsilon_{t-1}} & if\ \epsilon_{t-1} \geq 0 \\ e^{-0.2556\epsilon_{t-1}} & if\ \epsilon_{t-1} < 0. \end{cases}$$

- For a standardized shock with magnitude 2, (i.e. two standard deviations), we have

$$\frac{\sigma_t^2(\epsilon_{t-1}=-2)}{\sigma_t^2(\epsilon_{t-1}=2)} = \frac{\exp[-0.2556 \times (-2)]}{\exp(0.1573 \times 2)} = e^{0.1966} = 1.217.$$

- Therefore, the impact of a negative shock of size two-standard deviations is about 21.7% higher than that of a positive shock of the same size.

Forecasting using EGARCH$(1,1)$ model

- EGARCH(1,1) Forecasting: some recursive formula available

$$
\begin{cases}
a_t = \sigma_t \epsilon_t \\
\ln(\sigma_t^2) = \alpha_0 + \frac{g(\epsilon_{t-1})}{1-\alpha_1 B}
\end{cases}
$$

$$
\begin{aligned}
\ln(\sigma_t^2) &= \alpha_0(1-\alpha_1) + \alpha_1 \ln(\sigma_{t-1}^2) + g(\epsilon_{t-1}) \\
\sigma_t^2 &= \sigma_{t-1}^{2\alpha_1} e^{(1-\alpha_1)\alpha_0} e^{g(\epsilon_{t-1})}, \\
g(\epsilon_{t-1}) &= \theta\epsilon_{t-1} + \gamma\left(|\epsilon_{t-1}| - \sqrt{2/\pi}\right)
\end{aligned}
$$

- Let h be the forecast origin. For the 1-step-ahead forecast, we have

$$\sigma_{h+1}^2 = \sigma_h^{2\alpha_1} e^{(1-\alpha_1)\alpha_0} e^{g(\epsilon_h)}$$

Thus, the 1-step-ahead volatility forecast at the forecast origin h is $\hat{\sigma}_h^2(1) = \sigma_{h+1}^2$.

- For the 2-step-ahead forecast,

$$\sigma_{h+2}^2 = \sigma_{h+1}^{2\alpha_1} e^{(1-\alpha_1)\alpha_0} e^{g(\epsilon_{h+1})}$$

Taking conditional expectation at time h, we have

$$\hat{\sigma}_h^2(2) = \sigma_{h+1}^{2\alpha_1} e^{(1-\alpha_1)\alpha_0} E_h[e^{g(\epsilon_{h+1})}]$$

where E_h denotes a conditional expectation taken at the time origin h.

The prior expectation can be obtained as follows:

$$
\begin{aligned}
E[e^{g(\epsilon)}] &= \int_{-\infty}^{\infty} e^{\theta\epsilon+\gamma\left(|\epsilon|-\sqrt{2/\pi}\right)} f(\epsilon)d\epsilon \\
&= e^{-\gamma\sqrt{2/\pi}}\left[\int_{0}^{\infty} e^{(\theta+\gamma)\epsilon}\frac{1}{\sqrt{2\pi}}e^{-\frac{\epsilon^2}{2}}d\epsilon + \int_{-\infty}^{0} e^{(\theta-\gamma)\epsilon}\frac{1}{\sqrt{2\pi}}e^{-\frac{\epsilon^2}{2}}d\epsilon\right] \\
&= e^{-\gamma\sqrt{2/\pi}}\left[e^{\frac{(\theta+\gamma)^2}{2}}\Phi(\theta+\gamma) + e^{\frac{(\theta-\gamma)^2}{2}}\Phi(\theta-\gamma)\right]
\end{aligned}
$$

where $f(\epsilon)$ and $\Phi(x)$ are the probability density function and CDF of the standard normal distribution, respectively.

$$
\Rightarrow \hat{\sigma}_h^2(2) = \hat{\sigma}_h^{2\alpha_1}(1)\ e^{(1-\alpha_1)\alpha_0-\gamma\sqrt{2/\pi}}\left[e^{\frac{(\theta+\gamma)^2}{2}}\Phi(\theta+\gamma) + e^{\frac{(\theta-\gamma)^2}{2}}\Phi(\gamma-\theta)\right]
$$

同理可得

$$
\hat{\sigma}_h^2(j) = \widehat{\sigma_h^{2\alpha_1}}(j-1)e^{(1-\alpha_1)\alpha_0-\gamma\sqrt{2/\pi}}\left[e^{\frac{(\theta+\gamma)^2}{2}}\Phi(\theta+\gamma) + e^{\frac{(\theta-\gamma)^2}{2}}\Phi(\gamma-\theta)\right]
$$

Example

- For illustration, consider the AR(1)–EGARCH$(1,1)$ model of the previous section for the monthly log returns of IBM stock, ending December 1997.
- Using the fitted EGARCH$(1,1)$ model, we can compute the volatility forecasts for the series.
- At the forecast origin $t = 864$, the forecasts are

$$\begin{aligned} \hat{\sigma}^2_{864}(1) &= 6.44 \times 10^{-3}, \\ \hat{\sigma}^2_{864}(2) &= 6.24 \times 10^{-3}, \\ \hat{\sigma}^2_{864}(3) &= 6.06 \times 10^{-3}, \\ \hat{\sigma}^2_{864}(10) &= 5.89 \times 10^{-3}. \end{aligned}$$

- These forecasts converge gradually to the sample variance 4.37×10^{-3} of the shock process a_t.

R demonstration:

```
ibm=read.table(file="m-ibmln.dat")
ibm=ts(ibm,frequency=12,start=c(1926,1))

#fit a ARMA(1,0)+EGARCH(1,1) model
m.model=list(armaOrder=c(1,0),include.mean=T)
var.model=list(model="eGARCH",garchOrder=c(1,1))
dist.model="norm"
model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
fit=ugarchfit(data=ibm,spec=model)
#forecast
forc=ugarchforecast(fit, n.ahead=12)
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(1,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.011895    0.002060   5.7741 0.000000
ar1     0.093214    0.030332   3.0731 0.002118
omega  -0.416250    0.180463  -2.3066 0.021079
alpha1 -0.049112    0.026739  -1.8367 0.066259
beta1   0.923860    0.032702  28.2513 0.000000
gamma1  0.206452    0.048310   4.2735 0.000019

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.011895    0.001952   6.0951 0.000000
ar1     0.093214    0.024854   3.7505 0.000176
omega  -0.416250    0.268814  -1.5485 0.121510
alpha1 -0.049112    0.035761  -1.3733 0.169651
beta1   0.923860    0.048983  18.8607 0.000000
gamma1  0.206452    0.064783   3.1868 0.001438
```

```
LogLikelihood : 1166.037

Information Criteria
------------------------------------

Akaike       -2.6853
Bayes        -2.6522
Shibata      -2.6854
Hannan-Quinn -2.6726
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1108  0.7392
Lag[2*(p+q)+(p+q)-1][2]    0.8758  0.8115
Lag[4*(p+q)+(p+q)-1][5]    1.7610  0.7787
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.5559  0.4559
Lag[2*(p+q)+(p+q)-1][5]    0.8203  0.8988
Lag[4*(p+q)+(p+q)-1][9]    1.5354  0.9526
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.0458 0.500 2.000  0.8305
ARCH Lag[5]    0.6108 1.440 1.667  0.8505
ARCH Lag[7]    0.8948 2.315 1.543  0.9301
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.0998
Individual Statistics:
mu     0.29940
ar1    0.05665
omega  0.18557
alpha1 0.07523
beta1  0.16646
gamma1 0.15579

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.49 1.68 2.12
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias            1.223 0.2215
Negative Sign Bias   1.335 0.1821
Positive Sign Bias   1.015 0.3105
Joint Effect         2.814 0.4212


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     20.40      0.37099
2    30     40.65      0.07382
3    40     46.46      0.19191
4    50     68.18      0.03630
```

```
> resi=residuals(fit,standardize=T)
> Box.test(resi,lag=10,type="Ljung",fitdf=1)

        Box-Ljung test

data:  resi
X-squared = 6.6536, df = 9, p-value = 0.6731

> Box.test(resi,lag=20,type="Ljung",fitdf=1)

        Box-Ljung test

data:  resi
X-squared = 20.515, df = 19, p-value = 0.3642

> Box.test(resi^2,lag=10,type="Ljung",fitdf=2)

        Box-Ljung test

data:  resi^2
X-squared = 3.3714, df = 8, p-value = 0.9089

> Box.test(resi^2,lag=20,type="Ljung",fitdf=2)

        Box-Ljung test

data:  resi^2
X-squared = 11.146, df = 18, p-value = 0.8881
```

```
*---------------------------------*
*       GARCH Model Forecast      *
*---------------------------------*
Model: eGARCH
Horizon: 12
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=十二月 1997]:
       Series   Sigma
T+1  0.006541 0.08026
T+2  0.011396 0.07898
T+3  0.011849 0.07782
T+4  0.011891 0.07676
T+5  0.011895 0.07579
T+6  0.011895 0.07491
T+7  0.011895 0.07410
T+8  0.011895 0.07337
T+9  0.011895 0.07269
T+10 0.011895 0.07208
T+11 0.011895 0.07151
T+12 0.011895 0.07099
```

Remark:

Before you are comfortable with changing commands in R for EGARCH model estimation, you may use the GJR model discussed below to estimate the leverage effect.

- The Threshold GARCH (TGARCH) or GJR Model.
- TGARCH(s, m) or GJR(s, m) model is defined as

$$\begin{aligned} r_t &= \mu_t + a_t,\ a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2, \end{aligned}$$

where N_{t-i} is an indicator variable such that

$$N_{t-i} = \begin{cases} 1 & if\ a_{t-i} < 0, \\ 0 & if\ a_{t-i} \geq 0. \end{cases}$$

- One expects γ_i to be positive so that prior negative returns have higher impact on the volatility.

Example of TGARCH (GJR) model for the series of monthly IBM log returns from 1926 to 1997.

- The fitted GJR model is

$$
\begin{aligned}
r_t &= 0.012 + a_t,\ a_t = \sigma_t \epsilon_t \\
\sigma_t^2 &= 3.8 \times 10^{-4} + (0.053 + 0.082 N_{t-1}) a_{t-1}^2 + 0.815 \sigma_{t-1}^2,
\end{aligned}
$$

  where all estimates are significant, and model checking indicates that the fitted model is adequate.

- The sample variance of the IBM log returns is about 0.004 and the empirical 2.5% percentile of the data is about -0.119. If we use these two quantities for σ^2_{t-1} and a_{t-1}, respectively, then we have

$$\begin{aligned}\frac{\sigma_t^2(-)}{\sigma_t^2(+)} &= \frac{0.00038 + 0.135 \times 0.119^2 + 0.815 \times 0.004}{0.00038 + 0.053 \times 0.119^2 + 0.815 \times 0.004} \\ &\approx 1.26328\end{aligned}$$

- In this particular case, the negative prior return has about 26% higher impact on the conditional variance.

$$\begin{aligned}r_t &= 0.012 + a_t,\ a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 3.8 \times 10^{-4} + (0.053 + 0.082 N_{t-1}) a_{t-1}^2 + 0.815 \sigma_{t-1}^2,\end{aligned}$$

R demonstration:

```
> ibm=read.table(file="m-ibmln.dat")
> ibm=ts(ibm,frequency=12,start=c(1926,1))
>
> var(ibm)
            V1
V1 0.004397524
> quantile(ibm,0.025)
      2.5%
-0.1188095
>
> #fit a ARMA(0,0)+TGARCH(1,1) model
> m.model=list(armaOrder=c(0,0),include.mean=T)
> var.model=list(model="gjrGARCH",garchOrder=c(1,1))
> dist.model="ged"
> model=ugarchspec(variance.model=var.model,mean.model=m.model,distribution.model=dist.model)
> fit=ugarchfit(data=ibm,spec=model)
> #forecast
> forc=ugarchforecast(fit, n.ahead=12)
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.012253    0.002009   6.0995 0.000000
omega   0.000380    0.000153   2.4798 0.013145
alpha1  0.052856    0.027433   1.9268 0.054008
beta1   0.815244    0.054374  14.9932 0.000000
gamma1  0.082255    0.046203   1.7803 0.075028
shape   1.519483    0.096189  15.7969 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.012253    0.002238   5.4748 0.000000
omega   0.000380    0.000155   2.4498 0.014293
alpha1  0.052856    0.024980   2.1159 0.034351
beta1   0.815244    0.051719  15.7630 0.000000
gamma1  0.082255    0.045750   1.7979 0.072190
shape   1.519483    0.091797  16.5527 0.000000
```

```
LogLikelihood : 1173.4

Information Criteria
------------------------------------

Akaike       -2.7023
Bayes        -2.6692
Shibata      -2.7024
Hannan-Quinn -2.6897
```

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic  p-value
Lag[1]                      8.788 0.003032
Lag[2*(p+q)+(p+q)-1][2]     9.905 0.002025
Lag[4*(p+q)+(p+q)-1][5]    10.962 0.005198
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1143  0.7353
Lag[2*(p+q)+(p+q)-1][5]    0.3793  0.9746
Lag[4*(p+q)+(p+q)-1][9]    1.2316  0.9746
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.1318 0.500 2.000  0.7166
ARCH Lag[5]    0.6735 1.440 1.667  0.8315
ARCH Lag[7]    1.1446 2.315 1.543  0.8889
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.9236
Individual Statistics:
mu     0.3405
omega  0.1958
alpha1 0.2729
beta1  0.2328
gamma1 0.1816
shape  0.8839

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.49 1.68 2.12
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias            1.146 0.2522
Negative Sign Bias   1.082 0.2794
Positive Sign Bias   1.051 0.2934
Joint Effect         2.277 0.5169


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     15.31       0.7030
2    30     23.71       0.7432
3    40     40.81       0.3906
4    50     55.91       0.2315
```

```
*---------------------------------*
*       GARCH Model Forecast      *
*---------------------------------*
Model: gjrGARCH
Horizon: 12
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=十二月 1997]:
      Series   Sigma
T+1  0.01225 0.07070
T+2  0.01225 0.07018
T+3  0.01225 0.06970
T+4  0.01225 0.06926
T+5  0.01225 0.06886
T+6  0.01225 0.06849
T+7  0.01225 0.06815
T+8  0.01225 0.06785
T+9  0.01225 0.06757
T+10 0.01225 0.06731
T+11 0.01225 0.06708
T+12 0.01225 0.06687
```

The CHARMA model

- Conditional heteroscedastic ARMA model 採用 random coefficient 來製造 conditional heteroscedastic.
- Make use of "interaction" btw past shocks (i.e. 模型包含 $a_{t-1}a_{t-2}$ 等項)
- A CHARMA model is defined as

$$r_t = \mu_t + a_t, \; a_t = \delta_{1t}a_{t-1} + \delta_{2t}a_{t-2} + \cdots + \delta_{mt}a_{t-m} + \eta_t,$$

where $\{\eta_t\}$ is iid $N(0, \sigma_\eta^2)$, $\{\delta_t\} = \{(\delta_{1t}, \ldots, \delta_{mt})'\}$ is a sequence of iid random vectors $D(0, \Omega)$, $\{\delta_t\} \perp \{\eta_t\}$.

- The model can be written as

$$\begin{aligned} a_t &= \mathbf{a}'_{t-1}\delta_t + \eta_t \\ &= (a_{t-1}, \ldots, a_{t-m}) \begin{pmatrix} \delta_{1t} \\ \vdots \\ \delta_{mt} \end{pmatrix} + \eta_t \end{aligned}$$

with conditional variance

$$\begin{aligned} Var(a_t|F_{t-1}) = \sigma_t^2 &= \sigma_\eta^2 + \mathbf{a}'_{t-1}Cov(\delta_t)\mathbf{a}_{t-1} \\ &= \sigma_\eta^2 + (a_{t-1}, \ldots, a_{t-m})\Omega(a_{t-1}, \ldots, a_{t-m})'. \end{aligned}$$

$(Var(a_t|F_{t-1}) \geq 0\ \because\ \Omega$ 為 positive defined matrix)

- E.q $m = 1$ 時，$\sigma_t^2 = \sigma_\eta^2 + \omega_{11}a_{t-1}^2$ 與 ARCH(1) 相同。
- 當 $m = 2$ 時，$\sigma_t^2 = \sigma_\eta^2 + \omega_{11}a_{t-1}^2 + 2\omega_{12}a_{t-1}a_{t-2} + \omega_{22}a_{t-2}^2$，比 ARCH(2) 多了交叉項。

Example

Monthly excess returns of S&P 500 index (1926.01-1991.12).

- 若考慮 $r_t = \phi_0 + a_t$, $a_t = \delta_{1t} a_{t-1} + \delta_{2t} a_{t-2} + \eta_t$ 則 $\hat{\Omega}$ 中的交叉項的係數僅為 marginal significant。
- 所以我們改考慮下面模型

$$
\begin{aligned}
r_t &= \phi_0 + a_t \\
a_t &= \delta_{1t} a_{t-1} + \delta_{2t} a_{t-2} + \delta_{3t} a_{t-3} + \eta_t
\end{aligned}
$$

並強制要求 δ_{3t} 與 $(\delta_{2t}, \delta_{1t})$ uncorrelated 。

- A fitted model is

$$r_t = 0.0068 + a_t$$

$$\sigma_t^2 = 0.00136 + (a_{t-1}, a_{t-2}, a_{t-3})\hat{\Omega}(a_{t-1}, a_{t-2}, a_{t-3})'$$

$$= 0.00136 + 0.12a_{t-1}^2 - 0.12a_{t-1}a_{t-2} + 1.19a_{t-2}^2 + 0.299a_{t-3}^2$$

$$\hat{\Omega} = \begin{bmatrix} 0.121(0.036) & -0.062(0.028) & 0 \\ -0.062(0.028) & 1.191(0.025) & 0 \\ 0 & 0 & 0.299(0.042) \end{bmatrix}.$$

係數均顯著 (std errors in parentheses)。

- 當 $a_{t-1}a_{t-2} < 0$ 時，σ_t^2 變大。
- 當 Ω 為對角矩陣時，CHARMA(m) 模型的 σ_t^2 與 ARCH(m) 有相同的結構。
- 當 m 大的時候，我們可採用一些限制式，使大部分的交互作用項為零。

Including effects of explanatory variables in CHARMA models

- Can be used in the same manner, i.e. with random coefs.
- 以下將 CHARMA 推廣使得 r_t depends on some explanatory variable $\{x_{it}\}_{i=1}^{m}$,

$$r_t = \mu_t + a_t$$

$$\begin{aligned} a_t &= \sum_{i=1}^{m} \delta_{it} x_{i,t-1} + \eta_t,\ \delta = (\delta_{1t}, \ldots, \delta_{mt}) \perp \{\eta_t\} \\ \sigma_t^2 &= \sigma_\eta^2 + (x_{1,t-1}, \ldots, x_{m,t-1})\Omega(x_{1,t-1}, \ldots, x_{m,t-1})' \end{aligned}$$

其中 $x_{i,t-1}$ 可包含 $a'_{t-j}s$。

RCA model (random coefficient autoregressive)

- A time series r_t is a RCA(p) model if

$$r_t = \phi_0 + \sum_{i=1}^{p} (\phi_i + \delta_{it}) r_{t-i} + a_t \quad \text{(mean equation)}$$

$\{\delta_t\} = \{(\delta_{1t}, \ldots, \delta_{pt})'\}$ a sequence of indep random vectors with mean zero & covariance Ω_δ 且與 $\{a_t\}$ indep.

- For the RCA(p) model, we have

$$\mu_t = E(r_t|F_{t-1}) = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i},$$
$$\sigma_t^2 = \sigma_a^2 + (r_{t-1}, \ldots, r_{t-p})\Omega_\delta(r_{t-1}, \ldots, r_{t-p})'.$$

- 在 CHARMA 模型則假設 $\sigma_t^2 = \mathbf{a}'_{t-1}\Omega\mathbf{a}_{t-1}$ 。
- RCA 與 CHARMA 模型的主要差異：
  - RCA 的 volatility 是 r_{t-i} (observed lagged value) 的 quadratic 函數
  - CHARMA 則為 a_{t-i} (lagged innovation) 的 quadratic 函數。

Stochastic volatility model

- 在 conditional variance equation 中加入 innovation 的項
- A (simple) SV model is

$$a_t = \sigma_t \epsilon_t, \; (1 - \alpha_1 B - \cdots - \alpha_m B^m) \ln(\sigma_t^2) = \alpha_0 + v_t$$

where

- $\epsilon_t' s$ are iid N$(0, 1)$
- $v_t' s$ are iid N$(0, \sigma_v^2)$
- $\{\epsilon_t\}$ and $\{v_t\}$ are independent
- 假設 $1 - \sum_{i=1}^{m} \alpha_i B^i = 0$ 的根落在單位圓之外。

Stochastic volatility model

- 實証上發現 SV 模型通常可以改善 model fitting。
- 但 SV 模型對於 out of sample volatility 的預測未必有較好的結果。
- SV 模型可增加模型的 flexibility，但使得參數的估計變困難。
- 文獻上有兩種參數估計法
  1. quasi-likelihood method via Kalman filtering.
  2. Monte Carlo Markov Chain (MCMC)

Long-memory SV model

- A simple LMSV model is

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t = \sigma \exp(u_t/2), \quad (1-B)^d u_t = \eta_t$$

where $\sigma > 0$, $\epsilon_t \sim$ iid N$(0,1)$, $\eta_t \sim$ iid $N(0, \sigma_\eta^2)$, ϵ_t and η_t are independent, $0 < d < 0.5$.
- $\ln(a_t^2)$ series is a Gaussian long-memory signal plus a non-Gaussian white noise

$$\begin{aligned} \ln(a_t^2) &= \ln(\sigma^2) + u_t + \ln(\epsilon_t^2) \\ &= [\ln(\sigma^2) + E(\ln \epsilon_t^2)] + u_t + [\ln(\epsilon_t^2) - E(\ln \epsilon_t^2)] \\ &\equiv \mu + u_t + e_t \end{aligned}$$

其中 u_t 為 $I(d)$ 過程，e_t 為非常態分布的 i.i.d. innovations
- Breidt, Crato and de Lima (1998). "The detection and estimation of long memory in stochastic volatility" Journal of Econometrics.

Fractional difference d for companies in the S&P 500 index

- Estimation of the LMSV model is complicated, but the fractional difference d can be estimated by a quasi-maximum-likelihood method or a regression method.
- Bollerslev and Jubinski (1999) and Ray and Tsay (2000) found that the median estimate of d is about 0.38 for companies in the S&P 500 index (using the log series of squared daily return).
- Ray and Tsay (2000) studied common long-memory components in daily stock volatilities of groups of companies classified by various characteristics.
- They found that companies in the same industrial or business sector (e.g big U.S. national banks and financial institutions) tend to have more common long-memory components.

利用高頻資料估計低頻資料的波動

- French, Schwert & Stambaugh (1987) 曾提出利用高頻資料來估計低頻資料波動的想法。
- 近年來由於高頻資料較以前容易取得 (特別是在 foreign exchanges market 外匯市場)，此種方法再度引起討論 (Anderson et al, 1999.)

以 daily returns 來估計 monthly volatility

- r_t^m = 資產在第 t 月的對數報酬率.
- r_{t_i} = 在第 t 月中第 i 天的 log return
- 若在第 t 個月中有 n 個交易日，則

$$r_t^m = \sum_{i=1}^{n} r_{t_i}$$

- 假設 variance 及 covariance 均存在，則

$$Var(r_t^m|F_{t-1}) = \sum_{i=1}^{n} Var(r_{t_i}|F_{t-1}) + 2\sum_{i<j} Cov[r_{t_i}, r_{t_j}|F_{t-1}]$$

其中 F_{t-1} 代表到 (含) 第 $(t-1)$ 個月之前的資訊集合。

$$Var(r_t^m|F_{t-1}) = \sum_{i=1}^{n} Var(r_{t_i}|F_{t-1}) + 2\sum_{i<j} Cov[r_{t_i}, r_{t_j}|F_{t-1}]$$

(a) If $\{r_{t_i}\} \sim$ WN, then $Var(r_t^m|F_{t-1}) = nVar(r_{t_1})$. Use $\hat{Var}(r_{t_1}) = \frac{1}{n-1}\sum_{i=1}^{n}(r_{t_i} - \overline{r}_t)^2$, where $\overline{r}_t = \frac{1}{n}\sum_{i=1}^{n} r_{t_i}$. Then the volatility estimate of the t-th month is

$$\hat{\sigma}_m^2 = \widehat{Var}(r_t^m|F_{t-1}) = \frac{n}{n-1}\sum_{i=1}^{n}(r_{t_i} - \overline{r}_t)^2$$

(b) If $\{r_{t_i}\} \sim$ MA(1), then

$$Var(r_t^m|F_{t-1}) = nVar(r_{t_1}) + 2(n-1)Cov(r_{t_1}, r_{t_2})$$

$$\Rightarrow \hat{\sigma}_m^2 = \frac{n}{n-1}\sum_{i=1}^{n}(r_{t_i} - \overline{r}_t)^2 + 2\sum_{i=1}^{n-1}(r_{t_i} - \overline{r}_t)(r_{t_{i+1}} - \overline{r}_t)$$

此方法的弱點如下:

1. 必須先知道 daily return $\{r_{t_i}\}$ 的模型，否則無法估計其 covariance 結構。
2. 一個月中僅有 21 個交易日，樣本數太少，不易得到 variance 及 covariance 的有效估計值。
3. 當日報酬 $\{r_{t_i}\}$ 具有較大的 excess kurtosis 及較強的 serial correlation，則上述的 $\hat{\sigma}_m^2$ 可能為 inconsistent (Bou, Russell & Tiao, 2000)。

Example (以高頻資料估計低頻波動)

S&P 500 月對數報酬率的波動 (1926.1∼1999.12 (time plot); 1980.1 ∼ 1999.12 (modelling))

1. 假設 $\{r_{t,i}\}$ daily log return 為 WN，

$$\hat{\sigma}_m^2 = \frac{n}{n-1}\sum_{i=1}^{n}(r_{t,i} - \overline{r}_t)^2$$

2. 假設 daily log return 為 MA(1)，

$$\hat{\sigma}_m^2 = \frac{n}{n-1}\sum_{i=1}^{n}(r_{t,i} - \overline{r}_t)^2 + 2\sum_{i=1}^{n-1}(r_{t,i} - \overline{r}_t)(r_{t,i+1} - \overline{r}_t)$$

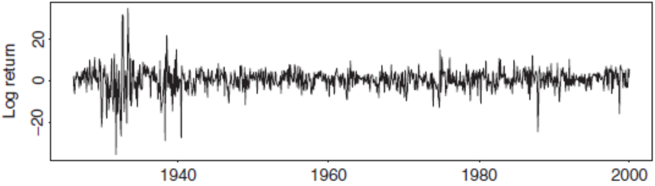
3. 直接對月資料配適模型 GARCH$(1,1)$ + Gaussian noise。

$$r_t^m = 0.658 + a_t, \;\; a_t = \sigma_t\epsilon_t$$

$$\sigma_t^2 = 3.349 + 0.086a_{t-1}^2 + 0.735\sigma_{t-1}^2$$

-

Figure: S&P500 月對數報酬率 (含股息) 的 time plot,1926.01-1999.12



(b)

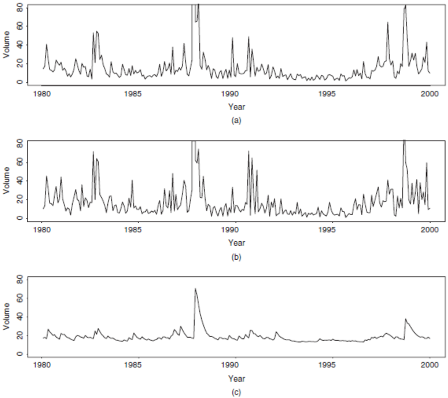


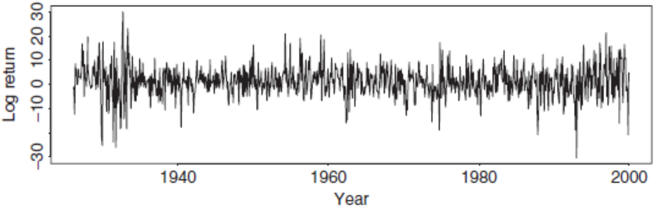
Figure 3.12 Time plots of estimated monthly volatility for log returns of S&P 500 index from January 1980 to December 1999: (a) assumes that daily log returns form a white noise series, (b) assumes that daily log returns follow an MA(1) model, and (c) uses monthly returns from January 1962 to December 1999 and a GARCH(1,1) model.

- (a) WN (b) MA(1) (c) GARCH。日報酬率所估計的波動 (a)&(b) 遠高於以月對數報酬率配適 GARCH$(1,1)$ 模型 (c) 所估計出來的波動

Example:

探討 IBM 的月對數報酬率在夏天的波動是否低於其他季節，1926.1 $\sim$ 1999.12.

Figure: IBM 月對數報酬率 (含股息), 1926.01-1999.12



(i) 先配適 GARCH-AR(1) 模型，係數皆顯著，且 $\{\tilde{a}_t\}$ 與 $\{\tilde{a}^2\}$ 均不具相關性 $\Rightarrow$ model is adequate。

(ii) 若加入考慮夏天的效應，

$$u_t = \begin{cases} 1, & \text{if } t \text{ is June, July, August} \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + u_t(\alpha_{00} + \alpha_{10} a_{t-1}^2 + \beta_{10} \sigma_{t-1}^2) \\ &= 4.539 + 0.113 a_{t-1}^2 + 0.816 \sigma_{t-1}^2 - 5.15 u_t \end{aligned}$$

- 以兩個 GARCH(1,1) 來描述 IBM 報酬的波動度。
- 其中 α_{10} 與 β_{10} 均不顯著。
- u_t 的係數顯著代表 summer effect is significant。
- $\{\tilde{a}_t\}$ 與 $\{\tilde{a}^2\}$ 仍為 uncorrelated ，故模型 adequate。

- $$\sigma_t^2 = \begin{cases} -0.615 + 0.113a_{t-1}^2 + 0.816\sigma_{t-1}^2, & \text{if } t \text{ is June, July, August} \\ 4.539 + 0.113a_{t-1}^2 + 0.816\sigma_{t-1}^2, & \text{otherwise} \end{cases}$$

  但 summer effect 的常數項 $-0.615 < 0$ 為一個不合理的數值，但由於 $-0.615 = 4.539 - 5.15$ (兩個 GARCH(1,1) 的常數項之合)，但原來兩個 GARCH(1,1) 常數項的標準差分別為 1.071 & 1.9，因此 -0.615 可能為一個不顯著的常數。
- 所以考慮以下模型：
  強制限制 summer volatility equation 的常數項為。

  $$\sigma_t^2 = \alpha_1 a_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \gamma(1 - u_t)$$

- 所得到的模型如下:

  $$\sigma_t^2 = 0.114a_{t-1}^2 + 0.811\sigma_{t-1}^2 + 4.552(1 - u_t)$$

  仍為 adequate model 。

- 在 1999 年 IBM 報酬率的 a_t^2 與 σ_t^2 的中位數，分別為 29.4 及 75.1 代入模型中，可得

  $$\sigma_t^2 = 0.114 \times 29.4 + 0.881 \times 75.1 = 64.3 \text{ for summer month}$$
  $$\sigma_t^2 = 0.114 \times 29.2 + 0.881 \times 75.1 + 4.552 = 68.8 \text{ for other month}$$

  $\frac{64.3}{68.8} \approx 93\% \Rightarrow$ IBM stock 的月對數報酬率的波動在夏季約較其他月份少了近 7%。

Example. S&P500 index(1926.1∼1999.12)

- S&P500 index 時常被用來當作衍生性商品的標的物
- 我們想探討是否 S&P500 index 中的個股過去報酬會影響 S&P500 index 報酬的波動模型。
- 我們考慮的個股為 IBM stock 的 past return。
- 令 r_t 代表 S&P500 index 的月對數報酬。
- GARCH(2,1) 模型，

$$r_t = 0.609 + a_t, \;\; a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = 0.717 + 0.147 a_{t-2}^2 + 0.859 \sigma_{t-1}^2$$

is an adequate model。

- 為了考慮 IBM 股票的效應，將模型修正為

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-2}^2 + \beta_1 \sigma_{t-1}^2 + \gamma(x_{t-1} - 1.24)^2$$

其中 x_{t-1} = IBM 股票的對數報酬率，其中 1.24 為 x_t 的樣本平均。

- 配適後得到下列模型:

$$r_t = 0.616 + a_t,\ a_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = 1.069 + 0.148 a_{t-2}^2 + 0.834 \sigma_{t-1}^2 - 0.007(x_{t-1} - 1.24)^2$$

其中 IBM 股票效應的係數 $\gamma = -0.007$ 為顯著，且模型為 adequate 。

- $\gamma = -0.007 < 0$，負號代表若在 S&P500 index 的 volatility 模型加入 IBM 股票的前一期報酬，會降低對 S&P500 index σ_t^2 的估計值。

Table: Fitted volatility for the monthly log returns of the S&P500 index form July to December 1999 using model s with and without the past log return of IBM stock.

Month	7/99	8/99	9/99	10/99	11/99	12/99
Model without IBM	26.30	26.01	24.73	21.69	20.71	22.46
Model with IBM	23.32	23.13	22.46	20.00	19.45	18.27

Alternative Approaches to Volatility

Some alternative methods:

- Moving window estimates
- Use of high-frequency financial data
- Use of daily open, high, low and closing prices

Moving window estimates

- A simple approach to capture time-varying feature of the volatility. There is no simple answer to the choice of window size.
- **Demonstration**: Consider the daily log returns of the S&P 500 index from 1950 to 2008. An R script, mvwindow.R, is available on the course web.

Instructions

1. Download the file and save it in your R working directory.
2. Compile the program using the command:
   source("mvwindow.R")
   The R script "mvwindow.R" can be downloaded from the website "`http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2016/`".
3. To run the program: mvol=mvwindow(rt,size), where "rt" denotes the return series and "size" is the size of the window.
4. The output is the volatility, i.e., σ_t , stored in "sigma.t".

Use of high-frequency financial data: Realized integrated volatility

- If the sample mean $\overline{r}_t$ is zero, then $\hat{\sigma}_m^2 \approx \sum_{i=1}^{n} r_{t,i}^2$.
  $\Rightarrow$ Use cumulative sum of squares of daily log returns within a month as an estimate of monthly volatility.
- Apply the idea to intra-daily log returns and obtain realized integrated volatility.
- Assume daily log return $r_t = \sum_{i=1}^{n} r_{t,i}$. The quantity

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2,$$

is called the *realized* volatility of r_t.

- **Advantages:** simplicity and using intra-daily information
- **Weaknesses:**
  - The higher the frequency at which prices are sampled, the larger the RV, indicating the possible presence of the microstructure noise (see figure next page).
  - Effects of market microstructure (noises) which is due to the imperfections of trading processes, including
    - price discreteness
    - infrequent trading
    - bid-ask bounce effects, $\cdots$
  - Overlook overnight volatilities.

RV v.s. sampling frequency (Fan and Wang, JASA 2007)

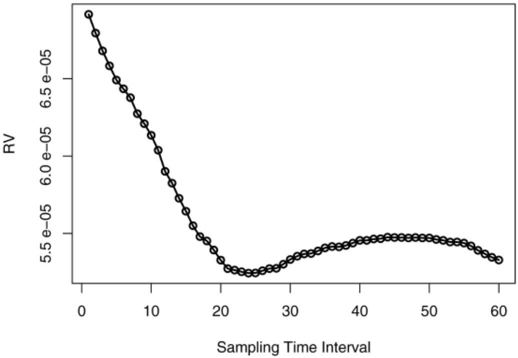


Figure 1. Plot of RV as a function of sampling time interval, in minutes. The horizon axis is the time interval in minutes that the data are sampled from euro–dollar exchange rates on January 7, 2004, for computing the RV. The shorter the sampling time interval, the higher the sampling frequency. The RV increases as the sampling time interval decreases, which suggests the presence of the microstructure noise.

Further discussion for the realized integrated volatility

- In-filled asymptotic argument.
  - Let Δ be the sampling interval, as $\Delta \to 0$, the sample size goes to infinity.
  - Under the assumption that the Δ-interval log returns, e.g. 5-minute returns, are i.i.d., then $\sum_{j=1}^{n} r_{t,j}^2$ converges to the variance of the daily log return r_t.
- In practice, however, there are microstructure noises that affect the estimate such as the bid-ask bounce.
- In fact, it can be shown that as Δ goes to zero, the observed sum of squares of Δ-interval returns goes to infinity. (see the figure of the previous page)

What next? How to solve the problem?

Two approaches have been proposed:

(a) Find an optimal sampling interval Δ (Bandi and Russell (2006)):

- Equivalently, the optimal sample size

$$n^* = 6.5\text{hours}/\Delta \approx \left[\frac{Q}{(\hat{\sigma}^2_{noise})^2}\right]^{1/3},$$

- $Q = \frac{M}{3}\sum_{j=1}^{M} r_{t,j}^4$
- $\hat{\sigma}^2_{noise} = \frac{1}{M}\sum_{j=1}^{M} r_{t,j}^2$
- $M =$ the number of daily quotes available for the underlying stock
- $r_{t,j} =$ returns computed from the mid-point of the bid and ask quotes.

(b) Subsampling method: Zhang et al. (2006).

- Choose Δ between 10 to 20 minutes.
- Compute the integrated volatility for each of the possible Δ-interval return series.
- Then, compute the average.

Use of Daily Open, High, Low and Close Prices

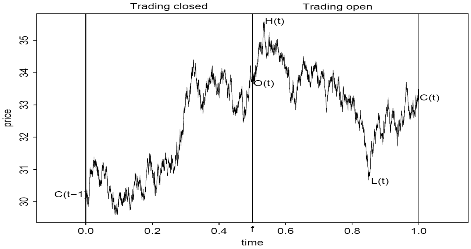


Figure: Time plot of intra-daily price over time

- C_t = the closing prices of the t-th trading day;
- O_t = the opening price of the t-th trading day;
- H_t = the highest price of the t-th trading period;
- L_t = the lowest price of the t-th trading period;

- The conventional conditional variance (or volatility) is

$$\sigma_t^2 = E[(C_t - C_{t-1})^2 | F_{t-1}],$$

  where F_{t-1} = public information available up to time $t-1$.
- Some alternatives:
  - $\hat{\sigma}^2_{0,t} = (C_t - C_{t-1})^2$;
  - $\hat{\sigma}^2_{1,t} = \frac{(O_t - C_{t-1})^2}{2f} + \frac{(C_t - O_t)^2}{2(1-f)},\ 0 < f < 1,$
  - $\hat{\sigma}^2_{2,t} = \frac{(H_t - L_t)^2}{4\ln(2)}$
  - $\hat{\sigma}^2_{3,t} = 0.17\frac{(O_t - C_{t-1})^2}{f} + 0.83\frac{\hat{\sigma}^2_{2,t}}{(1-f)}$
  - $\hat{\sigma}^2_{5,t} = 0.5(H_t - L_t)^2 - 2[2\ln(2) - 1](C_t - O_t)^2$
  - $\hat{\sigma}^2_{6,t} = 0.12\frac{(O_t - C_{t-1})^2}{f} + 0.88\frac{\hat{\sigma}^2_{5,t}}{1-f}$
  - f = fraction of the day (in interval [0,1]) that trading is closed.
  - A more precise, but complicated, estimator $\hat{\sigma}^2_{4,t}$ was also considered. But it is close to $\hat{\sigma}^2_{5,t}$.

- Defining the efficiency factor of a volatility estimator as

$$\text{Eff}(\hat{\sigma}^2_{i,t}) = \frac{Var(\hat{\sigma}^2_{0,t})}{Var(\hat{\sigma}^2_{i,t})}$$

- Garman and Klass (1980) found that $\text{Eff}(\hat{\sigma}^2_{i,t})$ is approximately $2,\ 5.2,\ 6.2,\ 7.4$ and 8.4 for $i = 1,\ 2,\ 3,\ 5$ and 6, respectively, for the simple diffusion model entertained.

Define

- $o_t = \ln(O_t) - \ln(C_{t-1})$ be the normalized open;
- $u_t = \ln(H_t) - \ln(O_t)$ be the normalized high;
- $d_t = \ln(L_t) - \ln(O_t)$ be the normalized low;
- $c_t = \ln(C_t) - \ln(O_t)$ be the normalized close;

- Suppose that there are n days of data available and the volatility is constant over the period.
- Yang and Zhang (2000) recommend the estimate

$$\hat{\sigma}_{yz}^2 = \hat{\sigma}_o^2 + k\hat{\sigma}_c^2 + (1-k)\hat{\sigma}_{rs}^2$$

as a robust estimator of the volatility, where

$$\hat{\sigma}_o^2 = \frac{1}{n-1}\sum_{t=1}^{n}(o_t - \overline{o})^2, \; with\ \overline{o} = \frac{1}{n}\sum_{t=1}^{n} o_t,$$

$$\hat{\sigma}_c^2 = \frac{1}{n-1}\sum_{t=1}^{n}(c_t - \overline{c})^2, \; with\ \overline{c} = \frac{1}{n}\sum_{t=1}^{n} c_t,$$

$$\hat{\sigma}_{rs}^2 = \frac{1}{n}\sum_{t=1}^{n}[u_t(u_t - c_t) + d_t(d_t - c_t)],$$

$$k = \frac{0.34}{1.34 + (n+1)/(n-1)}.$$

This estimate seems to perform well.

- Some alternative approaches to volatility estimation is currently under intensive study. It is rather early to assess the impact of these methods.
- It is a good idea in general to use more information. However, regulations and institutional effects need to be considered.