Chapter 2

Volatilities and Correlations

Volatilities and correlations are important concepts in quantitative finance, statistics and risk management. They are important parameters in forecasting the evolution of stock price, simulating stochastic models, pricing options and derivatives, calculating Value at Risk etc. In this chapter, we are going to explain how historical data can be used to estimate the volatilities and correlations. In particular, we shall consider commonly used models such as exponentially weighted moving average (EWMA), autoregressive conditional heteroscedasticity (ARCH), and generalized ARCH (GARCH). We shall also illustrate these models by real data. An important characteristic of these models is that volatilities and correlations are not constant, but change with time. It is commonly believed that this is closer to the reality than a constant volatility model. Before we go deep into the troubled waters, we first recall some notations, definitions of some technical terms.

2.1 Moving Standard Deviation

Recall that the basic assumption of **Black-Scholes-Merton** model is that the percentage change in the stock price in a short period of time is normally distributed. More precisely,

$$\delta S / S \sim N(\mu\,\delta t, \sigma^2\,\delta t),$$

where δS is the change in the stock price S over $(t, t+\delta t)$, $\mu(\delta t)$ is the mean of percentage change and $\sigma\sqrt{\delta t}$ is the standard deviation (s.d.) of this percentage change. Define

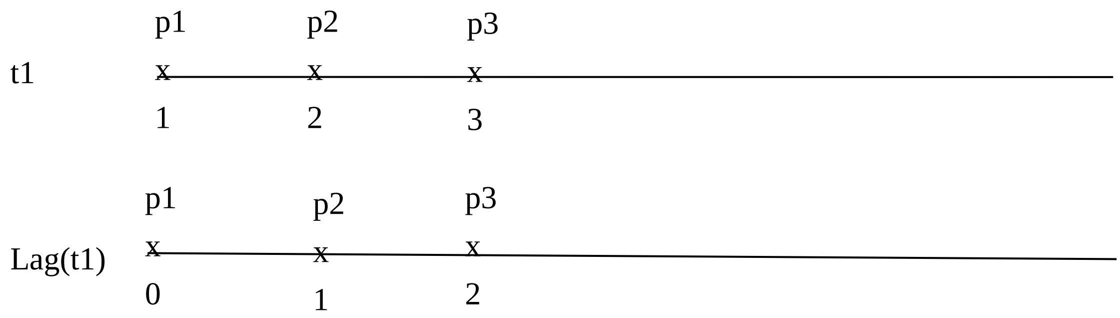
$$u_i = (S_i - S_{i-1}) / S_{i-1} \sim N(\mu\tau, \sigma^2\tau), \quad i = 1, \cdots, n, \tag{2.1}$$

where τ*=1/252*. Equation (2.1) suggests that the usual sample variance of u_i,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2$$

is an unbiased estimate of $\sigma^2\,\tau$ and hence we can estimate σ by $\hat{\sigma} = s/\sqrt{\tau}$. How to choose an appropriate value for n is not easy. Generally speaking, the estimate is more accurate when *n* is large. However, σ does change over time. Including too many irrelevant observations from the past may be useless or even harmful in predicting the future volatility. Using the most recent 90 days to 180 days seems to work reasonably well. As a rule of thumb, n should be equal to the number of days to which the volatility is applied. For example, if we want to apply the estimate to value a one-year option, then the daily stock price for last year should be used. More sophisticated methods for estimating volatility will be discussed in later sections. Let us first reconsider the stock example in Chapter 1. The file *"stock.csv"* contains adjusted daily closing price for the stock HSBC (0005), CLP (0002) and Cheung Kong (0001) from 1/1/1999 to 31/12/2002. Let us first read in these stock prices and compute the u_i for each stock respectively according to (2.1).

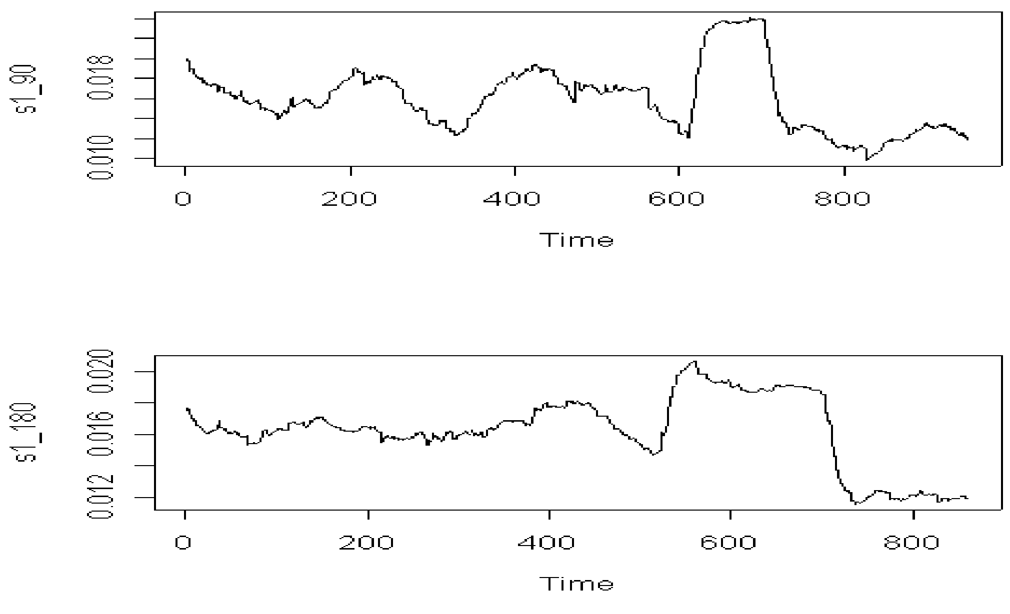
```
> d<-read.csv("stock.csv")
> t1<-as.ts(d$HSBC)          # save as time series
> t2<-as.ts(d$CLP)
> t3<-as.ts(d$CK)
> u1<-(lag(t1)-t1)/t1        # compute u
> u2<-(lag(t2)-t2)/t2
> u3<-(lag(t3)-t3)/t3
```



We have seen these time series plots from Chapter 1 that u_i is fluctuating around zero (roughly). The volatility is small in certain periods and large in other periods. Obviously the volatility varies with time. We focused on computing the "moving" standard deviation of u_i based on a window of length *w*. First, we simply compute the s.d. of the first *w* observations $u_1, \ldots, u_w$. Next, we compute the s.d. of $u_2, \ldots, u_{w+1}$, and so on. Then we have a series of s.d. This is similar to the concept of moving average in a time series except that we are computing the s.d. instead of the average.

To facilitate the computation, we first write the following R function *msd(t,w)* to calculate the "moving" s.d. of window width *w*. R has a built-in function *window(t,i,j)* to extract a subset from *i* to *j* of the time series *t* (see *help(window)* for details). However, we need a loop for varying the starting point *i* and end point *j* of the window. Then we use this *msd()* function to compute the 90 days and 180 days moving s.d. and plot the moving s.d. The following commands are contained in the file *"msd.r"* and can be executed using *source("msd.r")*.

```
msd<-function(t,w) {        # function to compute moving s.d.
   n<-length(t)-w+1
   out<-c()                 # initialize a null vector to store the output
   for (i in 1:n) {
      j<-i+w-1
      s<-sd(window(t,i,j))     # compute the sd of t(i) to t(j)
      out<-append(out,s)       # append the result to out
   }
   out<-as.ts(out)             # convert to time series
}
s1_90<-msd(u1,90)     # compute 90-day moving sd of u1
s1_180<-msd(u1,180)   # compute 180-day moving sd of u1
par(mfrow=c(2,1))     # time series plots
plot(s1_90)
plot(s1_180)
```



Notice that different window width *w* has very different effects on the volatility estimates. The minimum and maximum of *s1_90* is respectively 0.0098 and 0.02426. Hence the minimum and maximum annual volatility is $(\sqrt{252})(0.0098) = 15.56\%$ and $(\sqrt{252})(0.02426) = 38.51\%$. Similarly, the minimum and maximum of *s1_180* is

0.0116 and 0.0208 and the annual volatility is 18.41% and 33.02%. Empirical study shows that the annual volatility of stock is usually lies within 15% to 50%.

2.2 Exponentially Weighted Moving Average (EWMA)

In last section, the square of the 'moving' s.d. at time *t* is $s_t^2 = \frac{1}{w-1}\sum_{i=t-w+1}^{t}(u_i - \bar{u})^2$ which is an average of most recent *w* terms of $(u_i - \bar{u})^2$. This estimate actually carries equal weight of these *w* terms even though the most recent tem $(u_t - \bar{u})^2$ should have a greater effect on the volatility. For example, an extreme effect from long ago may still have effects on the estimate until this extreme value excluded from the moving window. To overcome this problem, we introduce the Exponentially Weighted Moving Average (EWMA) model which gives the highest weight to the most recent term and gradually down-weighted the terms as we move back in time. Before we go into the precise definition of this model, there are few remarks should be made.

1. In estimating the volatility, instead of using $u_i = \ln(S_i / S_{i-1})$, we usually use the percentage change in the stock price $u_i = (S_i - S_{i-1})/S_{i-1}$. In fact, it can be shown that these two definitions of u_i are close if u_i is small.
2. The mean of u_i, $\bar{u}$, is assumed to be zero for simplicity. In practice, the daily percentage change is small and fluctuates around zero. Assuming the mean of u_i equals zero also greatly simplifies the model as well.
3. Replace *w-1* in the denominator of (2.2) by w. This will give us a **Maximum Likelihood Estimate** (**MLE**) instead of unbiased estimate. This is also consistent with the ML estimation method we will use later on, i.e.,

$$s_t^2 = \frac{1}{w}\sum_{i=t-w+1}^{t} u_i^2 \,. \tag{2.2}$$

Now the EWMA model for estimating the variance rate is defined recursively as

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)\, u_{n-1}^2 \tag{2.3}$$

where $\lambda < 1$ is the parameter need to be estimated in this model. To see why this model is called EWMA, let us substitute σ_{n-1}^2 in (2.3) to obtain

$$\begin{aligned}\sigma_n^2 &= \lambda\,[\lambda\sigma_{n-2}^2 + (1-\lambda)u_{n-2}^2] + (1-\lambda)u_{n-1}^2 \\ &= (1-\lambda)(u_{n-1}^2 + \lambda\, u_{n-2}^2) + \lambda^2\sigma_{n-2}^2\end{aligned}$$

Continue substituting for σ_{n-2}^2,

$$\begin{aligned} \sigma_n^2 &= (1-\lambda)(u_{n-1}^2 + \lambda\, u_{n-2}^2 + \lambda^2\, u_{n-3}^2) + \lambda^3 \sigma_{n-3}^2 \\ &\vdots \\ \sigma_n^2 &= (1-\lambda)\sum_{i=1}^{m} \lambda^{i-1} u_{n-i}^2 + \lambda^m \sigma_{n-m}^2 \\ &\vdots \end{aligned} \tag{2.4}$$

Taking limit when m tends to infinity, the last term of (2.4) vanishes and the coefficient of the u_{n-i}^2 term is $(1-\lambda)\lambda^{i-1}$. These weights decrease exponentially as we move backward in time and they add up to 1. The Riskmetrics database developed by J. P. Morgran uses this EWMA model with λ=0.94 for updating the daily volatility estimate. Empirical studies found that this value of λ gives reasonably volatility estimate for most stocks.

Of course, we can also estimate this λ from the historical data by ML method. However, we defer the details of MLE of λ to the latter section. In next section, we shall introduce the famous GARCH(1,1) model which includes the EWMA model as a special case. We then describe the ML estimation method of the parameters in GARCH model which can also be applied to estimating λ in that EWMA model.

2.3 ARCH and GARCH model

First, let us consider the autoregressive conditional heteroscedasticity ARCH(m) model proposed by Engle (1982). This model assumes that the current variance rate σ_n^2 depends on the m most recent values of u_i^2 as well as a long-run average variance rate V_L. That is,

$$\sigma_n^2 = \gamma\, V_L + \alpha_1\, u_{n-1}^2 + \ldots + \alpha_m\, u_{n-m}^2 \tag{2.5}$$

where $\gamma + \alpha_1 + \ldots + \alpha_m = 1$. The advantage of this model is that it introduces a long-run variance rate. The disadvantages are there are too many parameters need to be estimated if m is large and it lacks of the recursive relationship found in EWMA model. However, combining the ideas in EWMA model in (2.3) and ARCH model in (2.5), this leads to the famous GARCH model.

The generalized autoregressive conditional heteroscedasticity GARCH(1,1) model was proposed by Bollerslev in 1986. Recall that in the EWMA model in (2.3), σ_n^2 is defined recursively in terms of σ_{n-1}^2 and u_{n-1}^2. In the ARCH(m) model in (2.5), σ_n^2

is defined in terms of *m* most recent values of u_i^2 and a long-run variance rate V_L . Now we define the GARCH(1,1) model as

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2, \tag{2.6}$$

where $\gamma + \alpha + \beta = 1$ and V_L is the long-run average variance rate. We shall explain why we called V_L as the long-run average variance rate later. Note that when $\gamma = 0$, GARCH(1,1) reduces to an EWMA model. The most general GARCH model is the GARCH(p,q) model:

$$\sigma_n^2 = \gamma V_L + \alpha_1 u_{n-1}^2 + \ldots + \alpha_p u_{n-p}^2 + \beta_1 \sigma_{n-1}^2 + \ldots + \beta_q \sigma_{n-q}^2 \quad .$$

However, in practice, we seldom use models other than the GARCH(1,1) model. GARCH(1,1) model is by far the most popular model used in estimating volatility.

2.4 Maximum Likelihood estimation in GARCH(1,1) model

We turn to the problem of estimating the parameters in the GARCH(1,1) model in (2.6). There are four unknown parameters V_L, γ, α, β need to be estimated with one constraint: $\gamma + \alpha + \beta = 1$. Therefore the GARCH(1,1) model can be rewritten as

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2, \tag{2.7}$$

where $\omega = \gamma V_L = (1-\alpha-\beta)V_L$. Recall that $u_i \sim N(0, \sigma_i^2)$. For convenience, we write σ_i^2 as v_i. The likelihood function of u_i is

$$\prod_{i=1}^{n} [\frac{1}{\sqrt{2\pi v_i}} \exp(\frac{-u_i^2}{2v_i})]$$

and the log-likelihood function of u_i is

$$\frac{n}{2}\ln(2\pi) + \frac{1}{2}\sum_{i=1}^{n} [-\ln(v_i) - \frac{u_i^2}{v_i}].$$

Therefore, we need to choose the parameters ω, α, β such that the function

$$l(\omega, \alpha, \beta) = \sum_{i=1}^{n} [-\ln(v_i) - \frac{u_i^2}{v_i}] \tag{2.8}$$

is maximized. Note that the parameters entered into the function (2.8) through v_i with the recursive relation defined in (2.7). This is a very complicated maximization problem. Fortunately R has a built-in function `garch()` located in a library "`tseries`" for computing the MLE of GARCH(1,1) model. This library does not come with the standard R package. We need to install this library by choosing `packages -> install package(s)` in the menu. Let us illustrate this by using the HSBC stock data. Recall that we first define $u_i = (S_i - S_{i-1})/S_{i-1}$ as the percentage change in stock price. Then we fit the u_i by a GARCH(1,1) model and obtain the MLE of ω, α, β.

```
> library(tseries)              # load library "tseries"
> res<-garch(u1,order=c(1,1))  # fit GARCH(1,1) model and save it to res
> names(res)                    # see what is in res
 [1] "order"         "coef"          "n.likeli"      "n.used"
 [5] "residuals"     "fitted.values" "series"        "frequency"
 [9] "call"          "asy.se.coef"
> round(res$coef,6)             # display the coefficient using 6 digits
      a0       a1       b1
0.000009 0.029318 0.934555
> -2*res$n.likeli               # compute log-likelihood value
[1] 7590.325
```

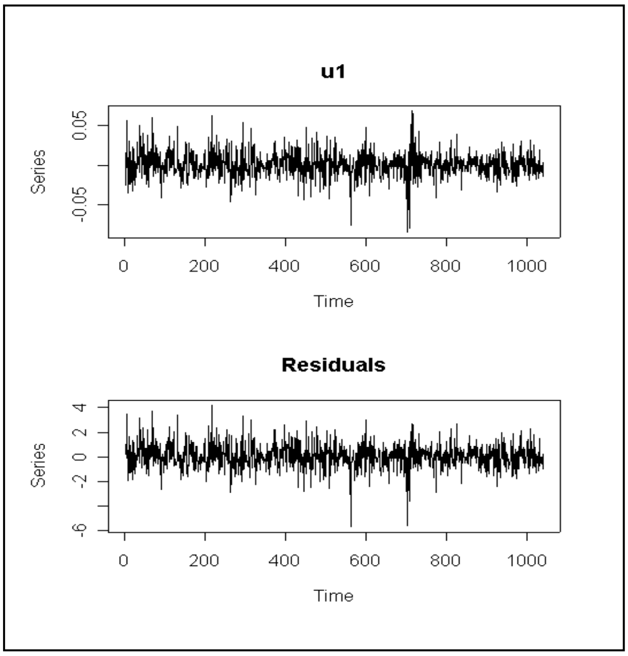
From the output, the MLE of ω, α, β are respectively 0.000009, 0.029318 and 0.934555. The maximum log-likelihood value in (2.8) is 7590.325. The estimated long-run variance rate is V_L = ω / (1−α−β) = 0.000252. We can display the summary of the results by the command `summary(res)`:

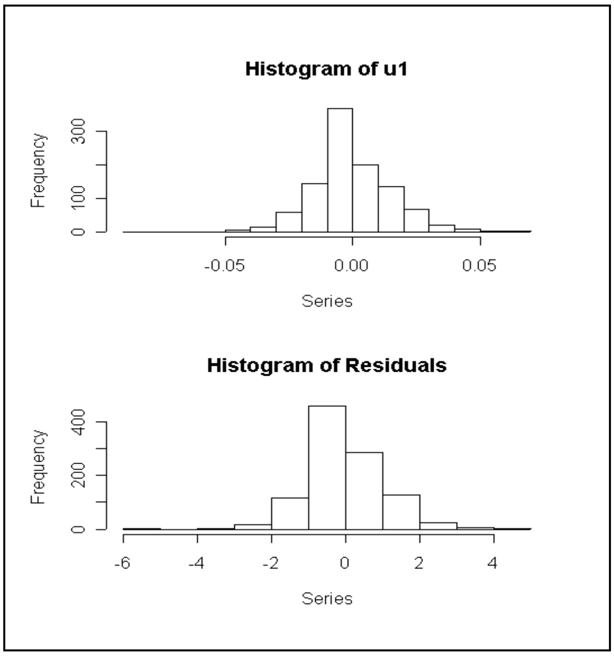
```
Residuals:
    Min      1Q  Median      3Q     Max
-5.7230 -0.5521  0.0000  0.5908  4.2013

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 8.716e-06   3.113e-06     2.80  0.00511 **
a1 2.932e-02   5.616e-03     5.22 1.79e-07 ***
b1 9.346e-01   1.653e-02    56.55  < 2e-16 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 291.4218, df = 2, p-value < 2.2e-16

        Box-Ljung test

data:  Squared.Residuals
X-squared = 0.5677, df = 1, p-value = 0.4512
```

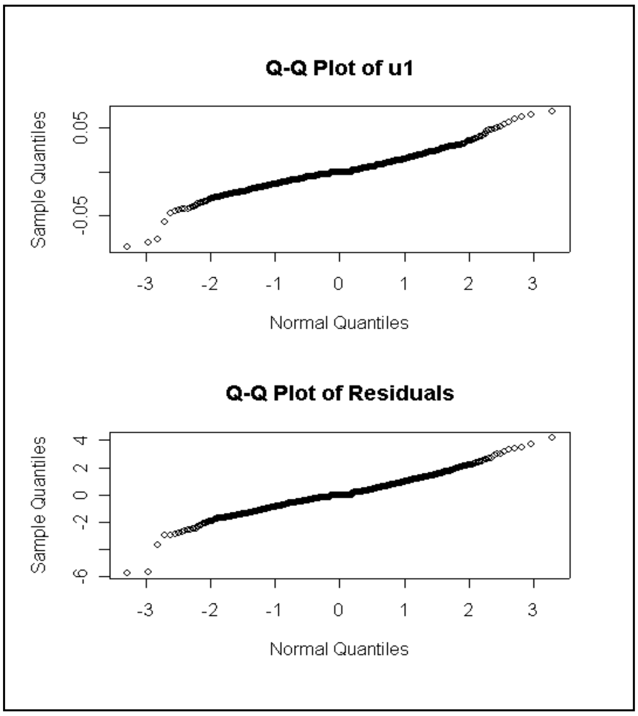
From the output, all the p-values are small and hence the coefficients are significantly different from zero. We defer the discussion of the diagnostic tests in next section.

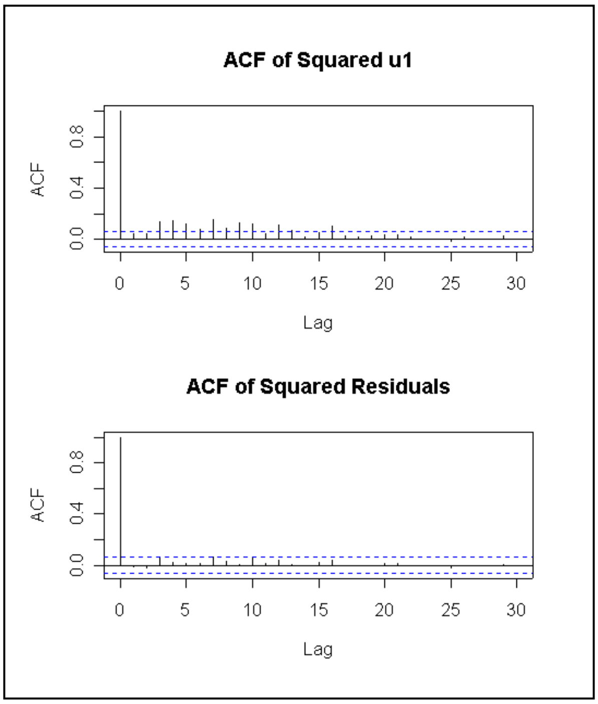
2.5 Model diagnostic in GARCH

We can also plot the results by the command > `plot(res)`. This gives several diagnostic plots as follow:









Recall that $u_i \sim N(0, \sigma_i^2)$. If GARCH(1,1) is correct, then residual $u_i / \sigma_i \sim N(0,1)$. From these plots, both u_i and the residuals are approximately normally distributed except for few outliers. More importantly, the auto-correlation function (acf) plots clearly indicate that the serial correlations exist in u_i^2 has been successfully removed by fitting this GARCH(1,1). The Box-Ljung test in the summary output is used to test whether the time series a_t has autocorrelations. The test statistic is

$$\chi_K^2 = n\sum_{k=1}^{K}\frac{n+2}{n-k}r_k^2 \quad \text{where} \quad r_k = \sum_{i=1}^{n-k} a_i a_{i+k} \Big/ \sum_{i=1}^{n} a_i^2 \quad \text{is the autocorr. of } a_t \text{ of lag k.}$$

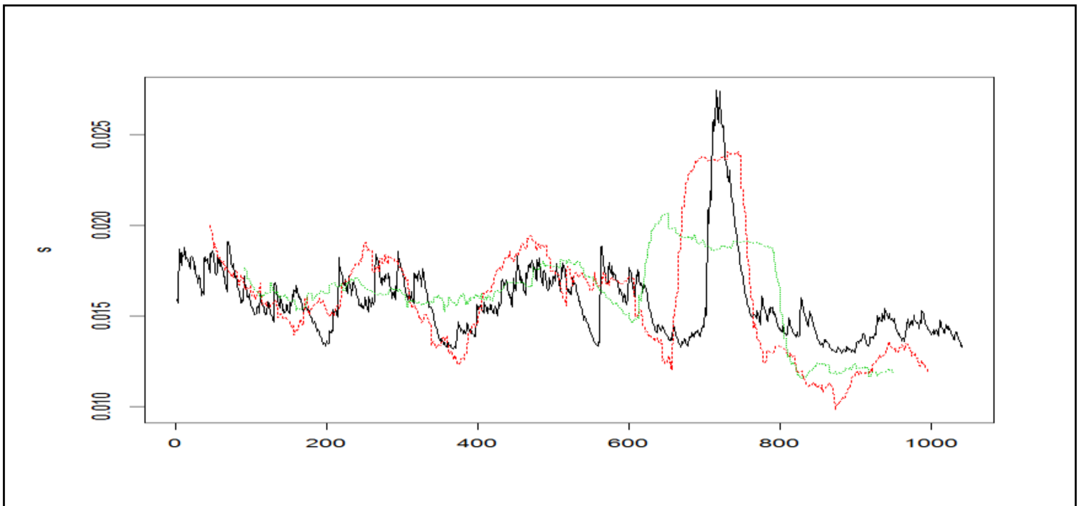
This test statistic has a Chi-square distribution of K degrees of freedom is no autocorrelation exist in a_t. The proof is given in the appendix. Let us compute the Box-Ljung test statistic for K=15 of u_i^2 and the squared residuals u_i^2/σ_i^2 respectively.

```
> Box.test(u1^2,lag=15,type="Ljung")

        Box-Ljung test
data:  u1^2
X-squared = 151.1576, df = 15, p-value < 2.2e-16

> Box.test(res$resid^2,lag=15,type="Ljung")

        Box-Ljung test
data:  res$resid^2
X-squared = 17.7587, df = 15, p-value = 0.2756
```

The P-value of u_i^2 being small means that there exist autocorrelations in u_i^2 while the P-value of u_i^2/σ_i^2 being large means that there is no autocorrelations in u_i^2/σ_i^2. It also clearly indicates that the autocorrelation of u_i^2 is removed by the GARCH model. This means that σ_i^2 is a good estimate of the variance rate.

Finally we can plot the fitted values to see some unique features in the volatilities of GARCH(1,1) model. The fitted values is simply the $\sigma_i = \sqrt{v_i}$, where σ_i^2 is computed recursively using (2.7) with the MLE of ω, α, β.

```
par(mfrow=c(1,1))
t90<-as.ts(c(rep(NA,45),s1_90))                # add 45 NA in front of s1_90
t180<-as.ts(c(rep(NA,90),s1_180))          # add 90 NA in front of s1_180
s<-cbind(res$fitted.values[,1],t90,t180)
matplot(s,type="l")
```



Estimated volatilities from GARCH(1,1) model, "moving" s.d. with 90 days and 180 days window are plotted together.. The plot shows that some clustered and spiky patterns in the estimated volatilities from GARCH(1,1) while the plot of the estimated volatilities using the "moving" s.d. tends to smooth out these spiky patterns and the effects of relatively high or low volatilities will last for a long period of time.

2.6 Forecasting future volatilities

Once the MLEs of ω, α, β are obtained, we can forecast the future volatilities using this GARCH(1,1) model. Recall that in (2.7),

$$\begin{aligned}\sigma_n^2 &= \omega + \alpha\, u_{n-1}^2 + \beta\, \sigma_{n-1}^2 \\ &= (1-\alpha-\beta)\, V_L + \alpha\, u_{n-1}^2 + \beta\, \sigma_{n-1}^2 \\ \Rightarrow \sigma_n^2 - V_L &= \alpha(u_{n-1}^2 - V_L) + \beta(\sigma_{n-1}^2 - V_L).\end{aligned}$$

On day $n+k$ in the future,

$$\sigma_{n+k}^2 - V_L = \alpha(u_{n+k-1}^2 - V_L) + \beta(\sigma_{n+k-1}^2 - V_L).$$

Replacing u_{n+k-1}^2 by its expectation, i.e. $E(u_{n+k-1}^2 \mid \sigma_{n+k-1}^2) = \sigma_{n+k-1}^2$, we have

$$\begin{aligned}E(\sigma_{n+k}^2 - V_L) &= (\alpha+\beta) E(\sigma_{n+k-1}^2 - V_L) \\ &= (\alpha+\beta)^2 E(\sigma_{n+k-2}^2 - V_L) \\ &\ \vdots \\ &= (\alpha+\beta)^k (\sigma_n^2 - V_L).\end{aligned}$$

Therefore, $E(\sigma_{n+k}^2) = V_L + (\alpha+\beta)^k (\sigma_n^2 - V_L).$ (2.9)

There are several remarks worth mentioning:

1. Note that where k is large and $\alpha+\beta<1$, the expected variance rate tends to V_L and that is the reason why we called V_L the long-run variance rate.
2. EWMA model can be considered as a special case of the GARCH(1,1) with $\omega = 0$ and $\lambda = \beta$ so that $\alpha+\beta = 1$. From (2.9), the expected variance rate is constant at the long-run variance rate V_L. This is an undesirable property of the EWMA model.
3. The expected variance rate exhibits a mean reversion property with reversion level V_L. If the current variance rate is above V_L, the estimated future variance rate in (2.9) will be pushed down to V_L when we look further and further ahead.

Similarly when the current variance rate is below V_L, the estimated future variance rate will be pulled up to V_L as we look forward in time. (Also see 4 below)

4. It can be shown that the GARCH(1,1) is equivalent to a model with the variance rate *V* follows the stochastic process with

$$dV = (1-\alpha-\beta)(V_L - V)dt + \alpha\sqrt{2}\,V\ dz$$

where z follows a Wiener process, note that $V_n - V_{n-1} = (1-\alpha-\beta)(V_L - V_{n-1}) + \alpha(\frac{u_{n-1}^2}{V_{n-1}} - 1)V_{n-1}$ and $Var(Z^2) = 2$ for $Z \sim N(0,1)$ In fact, this model is very similar to the **Vasicek** Model used in the term-structure model.

Example: To illustrate how to use (2.9) to forecast future volatilities, let us go back to our HSBC example. From the R's GARCH output, the MLE of ω, α, and β are respectively 0.000009, 0.029318 and 0.934555. The current daily volatility can be obtained from the last value of res$fitted.values and is equal to 0.0134. Hence $\sigma_n^2 = 0.0134^2 = 0.00018$ and $V_L = \omega/(1-\alpha-\beta) = 0.000249$.

Let $V(t) = E(\sigma_{n+t}^2), \quad a = -\ln(\alpha+\beta)$. Equation (2.9) becomes

$$V(t) = V_L + e^{-at}[V(0) - V_L].$$

V(t) is the estimate of the *t*-th day instantaneous variance. The average of the estimate variance from *n* to *n*+*T* days is $\frac{1}{T}\int_0^T V(t)dt = V_L + \frac{1-e^{-aT}}{aT}[V(0)-V_L]$.

Now suppose we have a 10-day option for HSBC, we need a single estimate of the annual volatility for valuing the option through the use of the Black-Scholes formula.

$V(10) = V_L + \frac{1-e^{-10a}}{10a}[V(0)-V_L] = 0.0001909$. Note that this is the estimated average daily variance rate. We have to convert this to annual volatility. The estimated daily volatility is $\sqrt{0.0001909} = 0.01382$ and the estimated annual volatility is $(\sqrt{252})(0.01382) = 0.2194$ or 22%.

Alternatively, we can use the discrete version, i.e., using the average of $E(\sigma_{n+1}^2), \ldots, E(\sigma_{n+T}^2)$ to estimate *V(T)*.

2.7 Correlation

Correlation plays an important role in reducing risk in a portfolio. It is also an important parameter in calculating VaR and in portfolio management. If the variance rate varies over time, it is not surprising that the correlation also varies with time. We can extent the EWMA model and GARCH(1,1) model to estimate correlation. First let us review the definition of covariance and correlation. Let

$$x_i = \frac{X_i - X_{i-1}}{X_{i-1}} \quad \text{and} \quad y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}$$

be the daily percentage change of two stocks X and Y respectively. The estimated covariance between X and Y based on the most recent m observations is

$$\text{cov}_{n-1} = \frac{1}{m}\sum_{i=1}^{m} x_{n-i} y_{n-i} \ . \tag{2.10}$$

The correlation between X and Y on day n is

$$\rho_{xy,n-1} = \frac{\text{cov}_{n-1}}{\sigma_{x,n-1}\sigma_{y,n-1}} \quad \text{where} \quad \sigma^2_{x,n-1} = \frac{1}{m}\sum_{i=1}^{m} x^2_{n-i} \quad \sigma^2_{y,n-1} = \frac{1}{m}\sum_{i=1}^{m} y^2_{n-i} \tag{2.11}$$

We have the current estimate of cov_{n-1} and ρ_{n-1} from (2.10) and (2.11). These estimates can be updated daily similarly by using EWMA or GARCH(1,1) model.

For EWMA model,

$$\text{cov}_n = \lambda\,\text{cov}_{n-1} + (1-\lambda) x_{n-1} y_{n-1} \quad \text{and} \quad \rho_\text{n} = \frac{\text{cov}_n}{\sigma_{x,n}\sigma_{y,n}}$$

where $\sigma^2_{x,n}$ and $\sigma^2_{y,n}$ are updated accordingly to the EWMA model by using (2.3).

For GARCH(1,1) model,

$$\text{cov}_n = \omega + \alpha\, x_{n-1} y_{n-1} + \beta\,\text{cov}_{n-1} \quad \text{and} \quad \rho_\text{n} = \frac{\text{cov}_n}{\sigma_{x,n}\sigma_{y,n}} \tag{2.12}$$

where $\sigma^2_{x,n}$ and $\sigma^2_{y,n}$ are updated according to the GARCH(1,1) model by using (2.7).

```
> u=cbind(u1,u2,u3)       # combine u1,u2 and u3 to form the matrix u
> u[1042,]                # display the current value of u
         u1           u2          u3
0.002941176 -0.003174603  0.019076305

> cor(u[953:1042,])       # compute the corr of u using the current 90 days
            u1          u2          u3
u1  1.00000000 -0.02368854  0.62052080
u2 -0.02368854  1.00000000 -0.01419078
u3  0.62052080 -0.01419078  1.00000000

> var(u[953:1042,])       # compute the var-cov matrix of u
              u1            u2            u3
u1  1.437215e-04 -2.287668e-06  1.446623e-04
u2 -2.287668e-06  6.489141e-05 -2.222993e-06
u3  1.446623e-04 -2.222993e-06  3.781606e-04
```

It is important to note that the estimated value of the parameters used in updating the variance rate and the covariance should be the same. Let us illustrate this by our example. First let us compute the correlation matrix of the daily percentage return of HSBC, CLP and CK.

Note that HSBC and CK are highly correlated while CLP has no or little correlation with HSBC and CK. The current value of u and the numbers in the covariance matrix will be used in (2.7) and (2.12) for updating.

Next we fit the GARCH(1,1) model to u1, u2 and u3 separately. This of course gives three different sets of ω, α and β. For simplicity we use the mean of these three sets of estimated parameter as the combined estimate for ω, α and β. Then we will use this set of combined estimate for updating the variance rates as well as the correlations. This will ensures the correlation matrix obtained is **positive definite**.

```
> res1=garch(u1)     # fit and save the garch(1,1) result
> res2=garch(u2)     # default order is (1,1)
> res3=garch(u3)
> (coef=rbind(res1$coef,res2$coef,res3$coef)) # combine and display coef.
               a0         a1        b1
[1,] 8.715587e-06 0.02931832 0.9345546
[2,] 6.318063e-06 0.12744105 0.8374826
[3,] 2.784795e-05 0.10194960 0.8489493

> round(apply(coef,2,mean),6)      # compute and display the column mean
      a0       a1       b1
0.000014 0.086236 0.873662
```

Now we have all the numbers we need to update the variance rates and the correlations according to (2.7) and (2.12).

$\omega = 0.000014 \quad \alpha = 0.086236 \quad \beta = 0.873662$

$\sigma_1^2 = \omega + \alpha(0.00294)^2 + \beta(0.0001437) = 0.00014$

$\sigma_2^2 = \omega + \alpha(-0.00317)^2 + \beta(0.0000649) = 0.000072$

$\sigma_3^2 = \omega + \alpha(0.01908)^2 + \beta(0.0003782) = 0.000376$

(See the entries of u_i and $Var(u)$ on P.12)

$$\text{cov}_{12} = \omega + \alpha(0.00294)(-0.00317) + \beta(-0.00000229) = 0.0000112$$
$$\text{cov}_{13} = \omega + \alpha(0.00294)(0.01908) + \beta(0.000145) = 0.000135$$
$$\text{cov}_{23} = \omega + \alpha(-0.00317)(0.01908) + \beta(-0.00000222) = 0.000012$$
$$\rho_{12} = \text{cov}_{12}/(\sigma_1\sigma_2) = 0.1117$$
$$\rho_{13} = \text{cov}_{13}/(\sigma_1\sigma_3) = 0.5886$$
$$\rho_{23} = \text{cov}_{23}/(\sigma_2\sigma_3) = 0.07353$$

A more sophisticated approach is to use **multivariate GARCH** model where the parameters are estimated simultaneously and the entire variance-covariance matrix is updated using the vector *u*.

2.8 EWMA and GARCH using EXCEL

The EXCEL we are using does not have the built-in GARCH or EWMA function. However, we can use the *solver* function to find the MLE of GARCH and EWMA. Let us illustrate this by fitting a GARCH(1,1) model to estimate the volatility of HSBC.

1. Set up the stock price of HSBC in B2:B1044 and the corresponding percentage return ui in C3:C1044.
2. Set up the cells J2, K2 and L2 for the parameters ω, α, and β. Enter some initial values to start with. For example, set ω=0.00001, α=0.05, β=0.9.
3. In cell M2, enter the formula =H2/(1-I2-J2) for computing the long run variance rate using the current values of ω, α and β.
4. In cell D3, enter the formula =C3^2. This serves as the initial value for *v*.
5. In cell D4, enter the formula =$J$2+$K$2*C3^2+$L$2*D3. This is the GARCH(1,1) model according to (2.7).
6. Copy the formula in cell D4 to D5:D1044.
7. In cell E4, enter the formula E3 =-LN(D3)-C3^2/D3. This is the first term in the summation of the log-likelihood function.
8. Copy the formula in E3 to E4:E1044.
9. In cell H6, enter the formula =SUM(E3:E1044). This is the value of the log-likelihood function we want to maximize.
10. Now use the *solver* function in the Tools menu. Specify J6 as the target cell and J2:L2 as the variable cells. Choose the max option and solve it. The parameter values in J2:L2 as well as columns C, D and E will change such that J6 is maximized.
11. The column D will be the final series of the estimated variance rate. If we want to plot the volatilities series, we can create the column F which is the square root of D and plot the volatilities in column F.

Variance targeting technique

Sometime, this estimation procedure is very unstable and sensitive to initial values of ω, α, and β. The idea of variance targeting technique is to reduce the free parameters to α and β only. The long-run variance rate V_L can be estimated by the sample variance of u_i and hence $\omega = V_L(1-\alpha-\beta)$.

1. Set up the cells K11 and L11 for α and β with some initial values as before.
2. Enter =VAR(C3:C1044) in M11 for the long-run variance rate.
3. Enter =M11*(1-K11-L11) in J11 for ω.
4. Set up columns G and H for v and the terms in the log-likelihood function as before but using new α and β.
5. Enter =SUM(H3:H1044) in J14 for the cell to be maximized.
6. Use the solver function as before to maximize J14.

Note: The EWMA model can be fitted using a similar method.

2.9 Combining GARCH with other models

An obvious application is to combine the GARCH model with models in Chapter 1. Recall that $u_t = (S_t - S_{t-1})/S_{t-1} \sim N(\mu, \sigma^2)$ or $u_t/\sigma \sim N(0,1)$, again we just assume $u = 0$ for simplicity. However, we can replace σ by σ_t, where σ_t^2 follows a GARCH(1,1) model, i.e., $u_t/\sigma_t \sim N(0,1)$ and $\sigma_{t+1}^2 = \omega + \alpha u_t^2 + \beta\sigma_t^2$. This is known as Normal-GARCH(1,1) model. Given ω, α, β and σ_{t+1}^2, we can simulate M sample paths for $u_{i,t+1}, \ldots, u_{i,t+K}$ for $i=1,\ldots,M$ accordingly to the following scheme:

1. Generate M normal random numbers: $z_1, \ldots z_M$ iid $N(0,1)$.
2. For $i=1,\ldots,M$, compute $u_{i,t+1} = \sigma_{t+1} z_i$ and $\sigma_{t+2}^2 = \omega + \alpha\, u_{i,t+1}^2 + \beta\, \sigma_{t+1}^2$.
3. Repeat steps 1 and 2 with $t=t+1$ until $t=t+K$.

Note that we have to generate a new set of M normal random numbers in step 1 for each t. Once we have M sample paths of $u_{i,t+K}$, we can valuate the option price and other derivative from them.

Other models can be simulated similarly. For example, we can also use the standardized $t(v)$ in Section 1.5 with the GARCH(1,1) model: $u_i/\sigma_i \sim t(v)$ and $\sigma_i^2 = \omega + \alpha u_{i-1}^2 + \beta\sigma_{i-1}^2$. This is known as t-GARCH(1,1) model as well as using $EWMA(\lambda)$ model instead of GARCH(1,1) model.

When using regression model for financial time series data, the errors are usually autocorrelated. We can handle this problem by fitting an AR(m) model to the error term in the regression model. For example,

$$y_t = x_t'\beta + v_t$$

$$v_t = -\varphi_1 v_{t-1} - \varphi_2 v_{t-2} - \ldots - \varphi_m v_{t-m} + \varepsilon_t \quad \text{where } \varepsilon_t \text{ iid } N(0, \sigma^2)$$

However, if the error variance is not constant but varies with *t*, say, σ_t^2, we can go one step further to fit a GARCH(p,q) model to the error term ε_t.

$$y_t = x_t'\beta + v_t$$
$$v_t = -\varphi_1 v_{t-1} - \varphi_2 v_{t-2} - \ldots - \varphi_m v_{t-m} + \varepsilon_t$$
$$\varepsilon_t = \sqrt{h_t}\, e_t \text{ where } e_t \text{ iid } N(0,1) \quad \Rightarrow \quad \varepsilon_t \sim N(0,h_t)$$
$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \gamma_j h_{t-j}$$

This model sometimes denoted as AR(m)-GARCH(p,q) model. In SAS, there is an AUTOREG procedure to fit this model. We can simply use this procedure to fit a GARCH(1,1) model to the percentage change of HSBC stock price.

2.10 Extension of GARCH model

1. Leverage GARCH (L-GARCH); also called Threshold GARCH (T-GARCH) or GJR-GARCH, (Glosten, Jagnathan and Runkle, 1993).

To account for the asymmetric effect of stock price on the volatility, we extent the GARCH(1,1) model to the following GJR-GARCH(1,1) model:

$$\sigma_n^2 = \omega + \beta\,\sigma_{n-1}^2 + \alpha\, u_{n-1}^2 + \theta\, I_{n-1}\, u_{n-1}^2 \tag{2.13}$$

where $I_{n-1} = 1$ if $u_{n-1} < 0$ and $I_{n-1} = 0$ otherwise.

Note that if θ >0, bad news ($u_{n-1} < 0$) has an effect $(\alpha + \theta)u_{n-1}^2$ on the variance; while good news ($u_{n-1} \geq 0$) has an effect $\alpha\, u_{n-1}^2$.

2. Exponential GARCH (E-GARCH)

Another commonly used model is the following E-GARCH model:

$$\ln \sigma_n^2 = \omega + \beta \ln \sigma_{n-1}^2 + \alpha \frac{|u_{n-1}| + \theta\, u_{n-1}}{\sigma_{n-1}} \tag{2.14}$$

Note that σ_n^2 is the exponent of RHS of (2.14) and is always > 0. Usually θ <0 so that σ_n^2 increases when $u_{n-1} < 0$, and vice versa.

These models can be implemented in EXCEL similar to Section 2.8.

3. Multivariate GARCH

The GARCH can also be extended to a multivariate setting:

$$\Sigma_n = \Omega + A * (u_{n-1}\, u_{n-1}') + B * \Sigma_{n-1}$$

where u_{n-1} is a $p \times 1$ column vector, Σ_n, Ω, A, B are $p \times p$ symmetric matrices, * stand for Hadamard product (entry by entry).

Appendix

1. Proof of Box-Ljung test

To prove that for time series a_i with $\mathrm{E}(a_i)=0$, if there is no serial autocorrelation in a_i, then for

$r_k = \frac{\sum_{i=1}^{n-k} a_i a_{i+k}}{\sum_{i=1}^{n} a_i^2}$, we get $Q(K) := \sum_{k=1}^{K} \frac{n(n+2)}{n-k} r_k^2 \sim \chi_K^2$.

We need to show that $k = 1, 2, .., K$, $\sqrt{\frac{n(n+2)}{n-k}} r_k \overset{i.i.d.}{\sim} N(0,1)$ approximately.

First, we assume the following characteristics for the sequence of a_i

1. For no serial correlation in a_i such that $\mathrm{E}(a_i{}^k a_j{}^l) = \mathrm{E}(a_i{}^k) \cdot \mathrm{E}(a_j{}^l)$ for $i \neq j$
2. $\mathrm{E}(a_i) = 0$
3. $\mathrm{E}(a_i{}^2)$ is constant with respect to i
4. $\mathrm{E}(a_i^4) = 3[\mathrm{E}(a_i^2)]^2$

Now, let $r_{k,n} := \frac{x(k,n)}{y(n)}$,

(i) For denominator of $r_k(n)$: $y(n) := \sum_{i=1}^{n} a_i^2$

$$E[y(n)] = E(\sum_{i=1}^{n} a_i{}^2) = \sum_{i=1}^{n} E(a_i{}^2) = nE(a_i{}^2)$$

$$E\{[y(n)]^2\} = E(\sum_{i=1}^{n} a_i{}^2 \sum_{j=1}^{n} a_j{}^2) = \sum_{i=1}^{n} E(a_i{}^4 + \sum_{i=1}^{n} \sum_{j=1.j\neq i}^{n} a_i{}^2 a_j{}^2) = \sum_{i=1}^{n} E(a_i{}^4) + \sum_{i=1}^{n} \sum_{j=1.j\neq i}^{n} E(a_i{}^2)E(a_j{}^2)$$

$$= \sum_{i=1}^{n} 3[E(a_i{}^2)]^2 + \sum_{i=1}^{n} \sum_{j=1.j\neq i}^{n} E(a_i{}^2)E(a_j{}^2), \mathit{from}(4)$$

$$= 3n[E(a_i{}^2)]^2 + n(n-1)[E(a_i{}^2)]^2, \mathit{from}(1)$$

$$= n(n+2)[E(a_i{}^2)]^2$$

(ii) For the numerator $x(k,n) := \sum_{i=1}^{n-k} a_i a_{i+k}$

$$E[x(k,n)] = E(\sum_{i=1}^{n-k} a_i a_{i+k}) = \sum_{i=1}^{n-k} E(a_i)E(a_{i+k}) = 0$$

$$E[x(k,n)^2] = E(\sum_{i=1}^{n-k} a_i a_{i+k} \cdot \sum_{j=1}^{n-k} a_j a_{j+k}) = E(\sum_{i=1}^{n-k}\sum_{j=1}^{n-k} a_i a_{i+k} a_j a_{j+k})$$

$$= E[\sum_{i=1}^{n-k}(a_i a_{i+k} a_i a_{i+k} + a_i a_{i+k} a_{i-k} a_i + \sum_{j=1, j\neq i, j\neq i-k}^{n-k} a_i \; a_{i+k} a_j a_{j+k})]$$

$$= \sum_{i=1}^{n-k}(E(a_i{}^2)E(a_{i+k}{}^2) + E(a_i{}^2)E(a_{i+k})E(a_{i-k}) + \sum_{j=1, j\neq i, j\neq i-k}^{n-k} E(a_i)E(a_{i+k})E(a_j)E(a_{j+k}))$$

$$= (n-k)[E(a_i{}^2)]^2$$

For $k \neq l$,

$$Cov[x(k,n), x(l,n)] = E(\sum_{i=1}^{n-k} a_i a_{i+k} \sum_{j=1}^{n-l} a_j a_{j+l})$$

$$= E(\sum_{i=1}^{n-k}\sum_{j=1}^{n-l} a_i a_{i+k} a_j a_{j+l})$$

$$= E[\sum_{j=1}^{n-l}(a_i a_{i+k} a_i a_{i+l} + a_i a_{i+k} a_{i+k} a_{i+k+l} + a_i a_{i+k} a_{i-l} a_i + a_i a_{i+k} a_{i+k-l} a_{i+k}$$

$$+ \sum_{j=1, j\neq i, j\neq i+k, j\neq i-l, j\neq i+k-l}^{n-l} a_i a_{i+k} a_j a_{j+l})]$$

$$= \sum_{j=1}^{n-l}[E(a_i{}^2)E(a_{i+k})E(a_{i+l}) + E(a_i)E(a_{i+k}{}^2)E(a_{i+k+l}) + E(a_i{}^2)E(a_{i+k})E(a_{i-l})$$

$$+ E(a_i)E(a_{i+k}{}^2)E(a_{i+k-l}) + \sum_{j=1, j\neq i, j\neq i+k, j\neq i-l, j\neq i+k-l}^{n-l} E(a_i)E(a_{i+k})E(a_j)E(a_{j+l})]$$

$$= 0$$

Using Slutsky's theorem,

$$E[r_k] \to \frac{E(\frac{1}{n}x(k,n))}{\sigma^2} = 0;$$

$$E[r_k \cdot r_l,] \to \frac{E(\frac{1}{n}x(k,n)\cdot\frac{1}{n}x(l,n))}{\sigma^2} = \frac{\frac{1}{n^2}Cov(x(k,n), x(l,n))}{\sigma^2} = 0$$

*Recall Slutsky's theorem $x_n \xrightarrow{d} c, y_n \xrightarrow{d} d > 0,$ for some constants c and d.

Then $\frac{x_n}{y_n} \xrightarrow{d} \frac{c}{d}$, so $E(\frac{x_n}{y_n}) = \frac{c}{d}$.

$$E({r_k}^2) \to \frac{E(x(k,n)^2)}{E(y(n)^2)} = \frac{(n-k)[E(a_i^2)]^2}{n(n+2)[E(a_i^2)]^2}$$

Hence, $\mathrm{E}\left[\frac{n(n+2)}{n-k}{r_k}^2\right] = 1$

Using martingale central limit theorem, $\sqrt{\frac{n(n+2)}{(n-k)}}\frac{\sum_{i=1}^{n-k} a_i a_{i+k}}{\sum_{i=1}^{n} a_i^2}$ are asymptotically normally distributed and independent of *k*.

Hence for $k = 1, 2, .., K$, $\sqrt{\frac{n(n+2)}{(n-k)}} r_k \overset{i.i.d.}{\sim} N(0,1)$, and $Q(K) = \sum_{k=1}^{K} \frac{n(n+2)}{(n-k)} {r_k}^2 \sim \chi_K^2$

2. Differential equation analogue of GARCH(1,1)

The GARCH(1,1) process is

$\sigma_{n+1}^2 = \gamma V_L + \alpha u_n^2 + \beta \sigma_n^2$, where $\alpha + \beta + \gamma = 1$, hence

$$\sigma_{n+1}^2 - \sigma_n^2 = \gamma (V_L - \sigma_n^2) + \alpha ({u_n}^2 - \sigma_n^2)$$

Using $z_i = u_i / \sigma_i \sim N(0,1)$ for all i, we get

$$\sigma_{n+1}^2 - \sigma_n^2 = \gamma (V_L - \sigma_n^2) + \alpha \sigma_n^2 ({z_n}^2 - 1)$$

Let $V(t) = \sigma_n^2$, and Δt be the time duration for time step from i to $i+1$, and realizing z follow Wienner process, we get

$$V(t + \Delta t) - V(t) = \{\gamma [V_L - V(t)]\Delta t + \alpha V(t)\sqrt{2} dz\}$$

Taking the limit of small Δt,

$$dV = \gamma [V_L - V(t)] dt + \alpha V(t) \sqrt{2} dz$$

Taking the expectation over the above expression and using $\mathrm{E}(dz) = 0$

$$dV = \gamma [V_L - V(t)] dt$$

Solving the above ODE , we get

$$V(t) = V(0) + [V(0) - V_L] e^{-\gamma t}$$

Since $\ln(1 + x) \approx x$ for small x , we get

$\ln(\alpha + \beta) = \alpha + \beta - 1 = -\gamma$ for small γ, hence

$$V(t) = V(0) + [V_L - V(0)]e^{\ln(\alpha+\beta)t}$$

Reference:

Chapter 5 of Risk Management and Financial Institutions, by John Hull, Wiley.