4 Conditional Heteroscedastic Models

Univariate Volatility Models.

What is asset volatility? Conditional standard deviation of the asset returns.

Why is volatility important?

1. Asset allocation, e.g., minimum-variance portfolio; see pages 184-185 of Campbell, Lo and MacKinlay (1997).
2. Risk management, e.g. value at risk (VaR)
3. Option (derivative) pricing, e.g., Black-Scholes formula
4. Interval forecasts
5. VIX (volatility index of market) is a tradable financial instrument.

A key characteristic: Not directly observable!!

■ **How to calculate volatility?**

There are several versions of volatility, but conditional standard deviation is commonly used.

1. Use high-frequency data: French, Schwert & Stambaugh (1987); see Section 3.15.
   - Realized volatility of daily log returns: use intraday high frequency log returns.
   - Use daily high, low, and closing (log) prices, e.g. range = daily high - daily low.
2. Implied volatility of options data, e.g, VIX of CBOE. Figure 36.
3. Econometric modeling: use daily or monthly returns

We focus on the econometric modeling first. Use of high frequency data will be discussed later.

How to model the evolving volatility? Two general categories

1. "Fixed function" and
2. "Stochastic function"

of the available information.

Basic idea of econometric modeling: Shocks of asset returns are NOT serially correlated, but dependent, implying that the serial dependence in asset returns is nonlinear.

As shown by the ACF of returns and absolute returns of some assets we discussed so far.

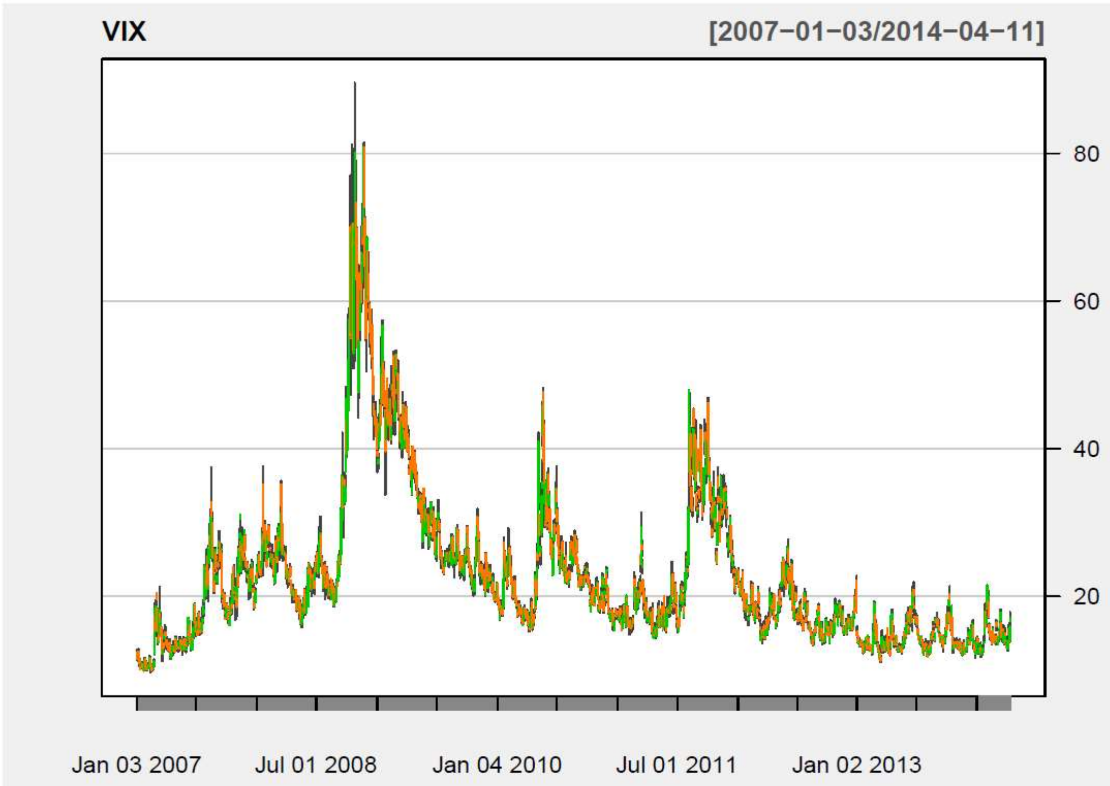

Figure 36: Time plot of the daily VIX index from January 3, 2007 to April 11, 2014.

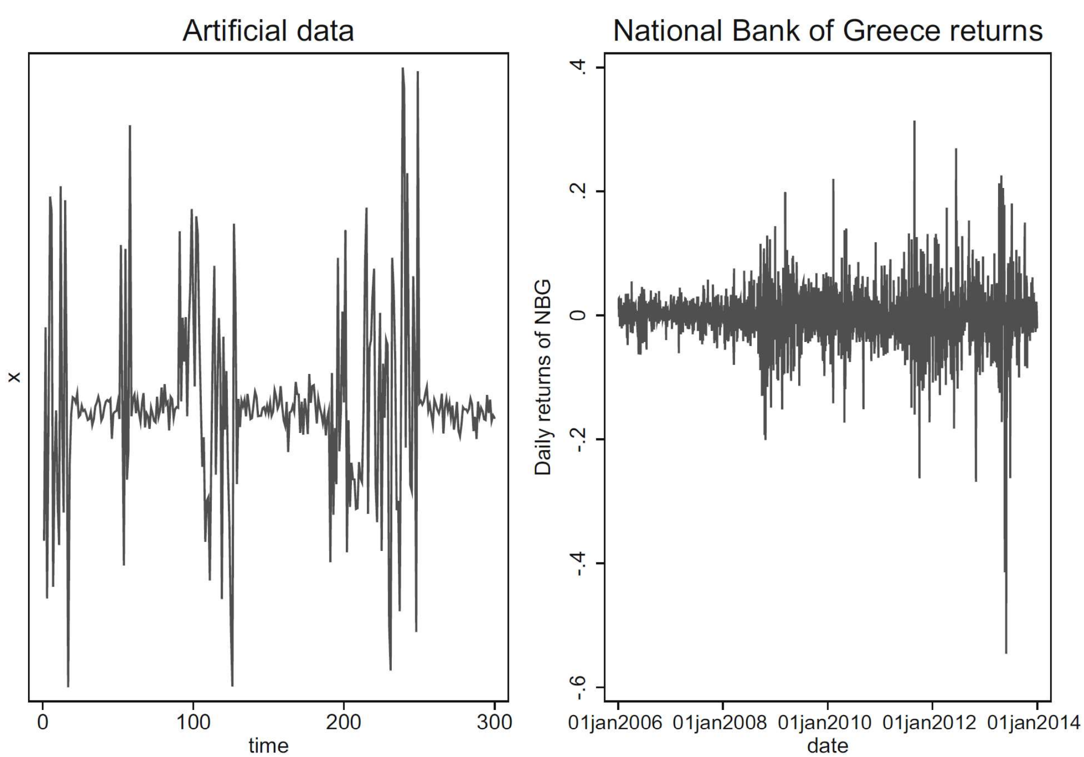

Figure 37: Two cases of volatility clustering.

■ **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 the conditional variance of the return r_t:

$$\sigma_t^2 = Var[r_t|F_{t-1}] = Var[a_t|F_{t-1}]$$

Consider the daily closing index of the S&P500 index from January 03, 2007 to April 13, 2015. The log returns follow approximately an MA(2) model.

$$r_t = a_t - 0.119a_{t-1} - 0.050a_{t-2}, \quad \sigma^2 = 0.00019.$$

R Demonstration:

```
> require(quantmod)
> getSymbols("^GSPC",from="2007-01-03",to="2015-04-13")
[1] "GSPC"
> chartSeries(GSPC,theme="white")
> spc=log(as.numeric(GSPC[,6]))
> rtn=diff(spc)
> acf(rtn)
> m1=arima(rtn,order=c(0,0,2),include.mean=F)
> m1
Call: arima(x = rtn, order = c(0, 0, 2), include.mean = F)
Coefficients:
ma1 ma2
-0.119 -0.0502
s.e. 0.022 0.0228
sigma^2 estimated as 0.0001899: log likelihood = 5966.01, aic = -11926.02
> resi=m1$residuals
> acf(resi)
> acf(resi^2)$
```

Now look at the ACF of squared residuals! Figure 38

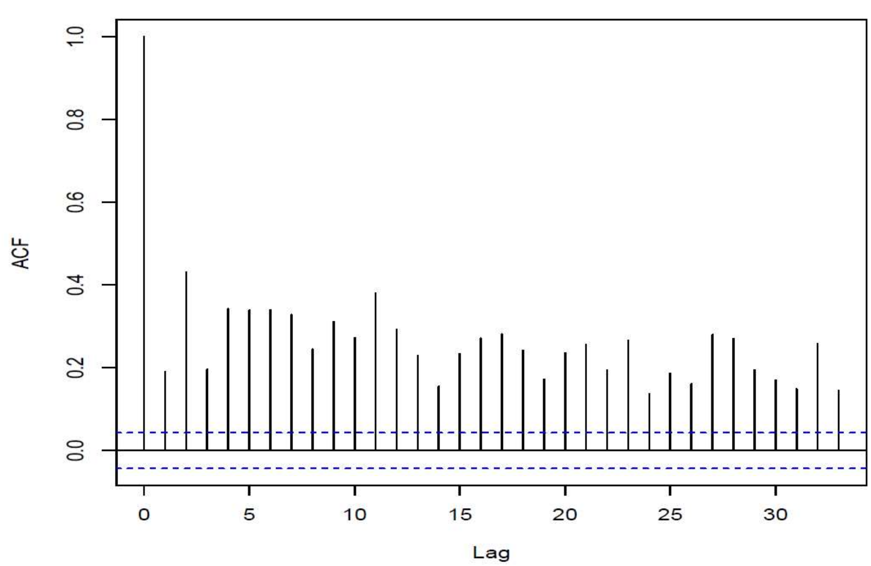

Figure 38: Sample ACF of the squared residuals of an MA(2) model fitted to daily log returns of the S&P 500 index from January 3, 2007 to April 13, 2015.

Note: In most applications, volatility is annualized. This can easily be done by taking care of the data frequency. For instance, if we use daily returns in econometric modeling, then the annualized volatility (in the U.S.) is $\sigma_t^* = \sqrt{252}\sigma_t$, where σ_t is the estimated volatility derived from an employed model. If we use monthly returns, then the annualized volatility is $\sigma_t^* = \sqrt{12}\sigma_t$, where σ_t is the estimated volatility derived from the employed model for the monthly returns. Our discussion, however, continues to use σ_t for simplicity.

In developing volatility models for financial time series, we need to consider some of the main characteristics of the volatility of these time series:

1. Volatility clustering
2. Continuity
3. Leverage effect
4. Volatility changes within a finite range.

Univariate volatility models discussed:

1. Autoregressive conditional heteroscedastic (ARCH) model of Engle (1982),
2. Generalized ARCH (GARCH) model of Bollerslev (1986),
3. GARCH-M models,
4. IGARCH models (used by RiskMetrics),
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. Asymmetric power ARCH (APARCH) models of Ding, Granger and Engle (1994), [TGARCH and GJR models are special cases of APARCH models.]
8. Stochastic volatility (SV) models of Melino and Turnbull (1990), Harvey, Ruiz and Shephard (1994), and Jacquier, Polson and Rossi (1994).

■ ARCH Model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + ... + \alpha_m a_{t-m}^2, \quad a_t = \sigma_t \epsilon_t,$$

where $\{\epsilon_t\}$ is a sequence of iid r.v. with mean 0 and variance 1, $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$. Distribution of ϵ_t: Standard normal, standardized Student-t, generalized error dist (ged), or their skewed counterparts.

Properties of ARCH models

Consider an ARCH(1) model

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

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

1. $E[a_t] = 0$
2. $Var[a_t] = \frac{\alpha_0}{1-\alpha_1}$ if $0 < \alpha_1 < 1$
3. Under normality,

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

   provided $0 < \alpha_1^2 < \frac{1}{3}$. The 3rd property implies heavy tails. The tail distribution of a_t is heavier than that of a normal distribution. Therefore, a_t is more likely to produce "outliers" compared to a white noise.

Advantages

– Simplicity
– Generates volatility clustering
– Heavy tails (high kurtosis)

Weaknesses

– Symmetric between positive and negative prior returns
– Restrictive
– Provides no explanation
– Not sufficiently adaptive in prediction (over-specification of volatility)

Building an ARCH Model

1. Modeling the mean effect and testing for ARCH effects (dependence in the squared residuals)

   H_o: no ARCH effects versus H_a: ARCH effects

   Use Q-statistics of squared residuals; Box test.
2. Order determination: Use PACF of the squared residuals. (In practice, simply try some reasonable order).
3. Estimation: Conditional MLE or Quasi MLE

4. Model checking: Q-stat of standardized residuals and squared standardized residuals. Skewness and Kurtosis of standardized residuals.

   R provides many plots for model checking and for presenting the results.

5. Software: We use R with the package fGarch.

Note: In this course, we estimate volatility models using the R package *fGarch* with *garchFit* command. The program is easy to use and allows for several types of innovational distributions: Gaussian (*norm*) : the default, standardized Student-t distribution (*std*), generalized error distribution (*ged*), skew normal distribution (*snorm*), skew Student-t (*sstd*), skew generalized error distribution (*sged*), and standardized inverse normal distribution (*snig*). Except for the inverse normal distribution, other distribution functions are discussed in the textbook. Readers should check the book for details about the density functions and their parameters.

Example: Monthly log returns of Intel stock

R Demonstration:

```
>library(fGarch)
> da=read.table("m-intc7303.txt",header=T)
> head(da)
date rtn
1 19730131 0.01005
.....
6 19730629 0.13333
> intc=log(da$rtn+1) <== log returns
$
> acf(intc)
> acf(intc^2)
> pacf(intc^2)
> Box.test(intc^2,lag=10,type='Ljung')
Box-Ljung test
data: intc^2
X-squared = 59.7216, df = 10, p-value = 4.091e-09
> m1=garchFit(~garch(3,0),data=intc,trace=F) <== trace=F reduces the amount of output.
> summary(m1)
Title: GARCH Modelling
Call: garchFit(formula = ~garch(3, 0), data = intc, trace = F)
Mean and Variance Equation:
data ~ garch(3, 0)
[data = intc]
Conditional Distribution: norm
Coefficient(s):
mu omega alpha1 alpha2 alpha3
0.016572 0.012043 0.208649 0.071837 0.049045
Std. Errors:
based on Hessian
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 0.016572 0.006423 2.580 0.00988 **
omega 0.012043 0.001579 7.627 2.4e-14 ***
alpha1 0.208649 0.129177 1.615 0.10626
alpha2 0.071837 0.048551 1.480 0.13897
alpha3 0.049045 0.048847 1.004 0.31536
---
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 169.7731 0
Shapiro-Wilk Test R W 0.9606957 1.970413e-08
Ljung-Box Test R Q(10) 10.97025 0.3598405
Ljung-Box Test R Q(15) 19.59024 0.1882211
Ljung-Box Test R Q(20) 20.82192 0.40768
Ljung-Box Test R^2 Q(10) 5.376602 0.864644
Ljung-Box Test R^2 Q(15) 22.73460 0.08993976
Ljung-Box Test R^2 Q(20) 23.70577 0.255481
```

```
LM Arch Test R TR^2 20.48506 0.05844884
Information Criterion Statistics:
AIC BIC SIC HQIC
-1.228111 -1.175437 -1.228466 -1.207193
> m1=garchFit(~garch(1,0),data=intc,trace=F)
> summary(m1)
Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 0), data = intc, trace = F)
Mean and Variance Equation:
data ~ garch(1, 0)
[data = intc]
Conditional Distribution: norm
Coefficient(s):
mu omega alpha1
0.016570 0.012490 0.363447
Std. Errors:
based on Hessian
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 0.016570 0.006161 2.689 0.00716 **
omega 0.012490 0.001549 8.061 6.66e-16 ***
alpha1 0.363447 0.131598 2.762 0.00575 **
---
Log Likelihood:
230.2423 normalized: 0.6189309
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 122.4040 0
Shapiro-Wilk Test R W 0.9647629 8.274158e-08
Ljung-Box Test R Q(10) 13.72604 0.1858587 <=== Meaning?
Ljung-Box Test R Q(15) 22.31714 0.09975386 <==== implication?
Ljung-Box Test R Q(20) 23.88257 0.2475594
Ljung-Box Test R^2 Q(10) 12.50025 0.2529700
Ljung-Box Test R^2 Q(15) 30.11276 0.01152131
Ljung-Box Test R^2 Q(20) 31.46404 0.04935483
LM Arch Test R TR^2 22.036 0.0371183
Information Criterion Statistics:
AIC BIC SIC HQIC
-1.221733 -1.190129 -1.221861 -1.209182
> plot(m1)
Make a plot selection (or 0 to exit):
1: Time Series
2: Conditional SD
3: Series with 2 Conditional SD Superimposed
4: ACF of Observations
5: ACF of Squared Observations
6: Cross Correlation
7: Residuals
8: Conditional SDs
9: Standardized Residuals
10: ACF of Standardized Residuals
11: ACF of Squared Standardized Residuals
12: Cross Correlation between r^2 and r
13: QQ-Plot of Standardized Residuals
Selection: 13
Make a plot selection (or 0 to exit):
1: Time Series
2: Conditional SD
3: Series with 2 Conditional SD Superimposed
4: ACF of Observations
5: ACF of Squared Observations
6: Cross Correlation
7: Residuals
8: Conditional SDs
9: Standardized Residuals
10: ACF of Standardized Residuals
11: ACF of Squared Standardized Residuals
12: Cross Correlation between r^2 and r
13: QQ-Plot of Standardized Residuals
Selection: 0
```

The fitted ARCH(1) model is

$$r_t = 0.0176 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1)$$
$$\sigma_t^2 = 0.0125 + 0.363 a_{t-1}^2.$$

Model checking statistics indicate that there are some higher order dependence in the volatility, e.g., see Q(15) for the squared standardized residuals. It turns out that a GARCH(1,1) model fares better for the data.

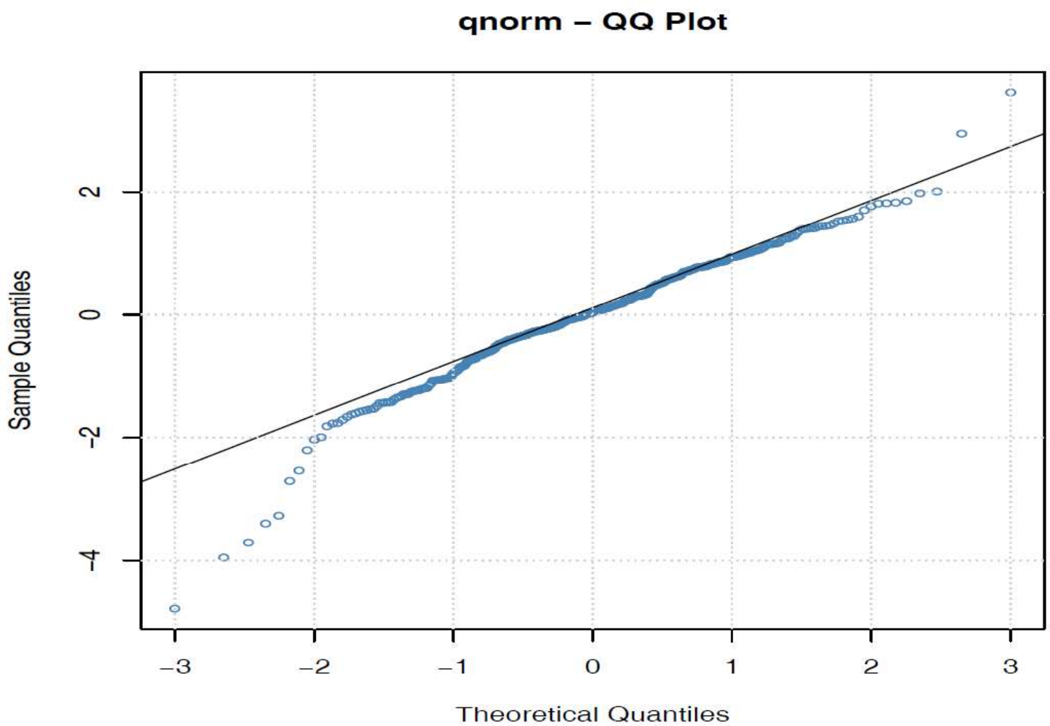

Figure 39: QQ-plot for standardized residuals of an ARCH(1) model with Gaussian innovations for monthly log returns of INTC stock: 1973 to 2003.

Next, consider Student-t innovations.

R Demonstration:

```
> m2=garchFit(~garch(1,0),data=intc,cond.dist="std",trace=F)
> summary(m2)
Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 0), data = intc, cond.dist = "std", trace = F)
Mean and Variance Equation:
data ~ garch(1, 0)
[data = intc]
Conditional Distribution: std <====== Standardized Student-t.
Coefficient(s):
mu omega alpha1 shape
0.021571 0.013424 0.259867 5.985979
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 0.021571 0.006054 3.563 0.000366 ***
omega 0.013424 0.001968 6.820 9.09e-12 ***
alpha1 0.259867 0.119901 2.167 0.030209 *
shape 5.985979 1.660030 3.606 0.000311 *** <== Estimate of degrees of freedom
---
Log Likelihood:
```

```
242.9678 normalized: 0.6531391
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 130.8931 0
Shapiro-Wilk Test R W 0.9637529 5.744026e-08
Ljung-Box Test R Q(10) 14.31288 0.1591926
Ljung-Box Test R Q(15) 23.34043 0.07717449
Ljung-Box Test R Q(20) 24.87286 0.2063387
Ljung-Box Test R^2 Q(10) 15.35917 0.1195054
Ljung-Box Test R^2 Q(15) 33.96318 0.003446127
Ljung-Box Test R^2 Q(20) 35.46828 0.01774746
LM Arch Test R TR^2 24.11517 0.01961957
Information Criterion Statistics:
AIC BIC SIC HQIC
-1.284773 -1.242634 -1.285001 -1.268039
> plot(m2)
Make a plot selection (or 0 to exit):
1: Time Series
2: Conditional SD
3: Series with 2 Conditional SD Superimposed
4: ACF of Observations
5: ACF of Squared Observations
6: Cross Correlation
7: Residuals
8: Conditional SDs
9: Standardized Residuals
10: ACF of Standardized Residuals
11: ACF of Squared Standardized Residuals
12: Cross Correlation between r^2 and r
13: QQ-Plot of Standardized Residuals
Selection: 13 <== The plot shows that the model needs further improvements.
> predict(m2,5) <===== Prediction
meanForecast meanError standardDeviation
1 0.02157100 0.1207911 0.1207911
2 0.02157100 0.1312069 0.1312069
3 0.02157100 0.1337810 0.1337810
4 0.02157100 0.1344418 0.1344418
5 0.02157100 0.1346130 0.1346130
```

■

The fitted model with Student-t innovations is

$$\begin{aligned} r_t &= 0.0216 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{5.99} \\ \sigma_t^2 &= 0.0134 + 0.260 a_{t-1}^2. \end{aligned}$$

We use $t_{5.99}$ to denote the standardized Student-t distribution with 5.99 d.f.

Comparison with normal innovations:

- Using a heavy-tailed dist for ϵ_t reduces the ARCH effect.
- The difference between the models is small for this particular instance.

You may try other distributions for ϵ_t.

■ GARCH Model

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

where $\{\epsilon_t\}$ is a sequence of iid r.v. with mean 0 and variance 1, $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$, $\beta_j \geq 0$, $\forall j > 0$, and $\sum_{i=1}^{\max\{m,s\}} (\alpha_i + \beta_i) < 1$.

to understand properties of GARCH models, e.g. moment equations, forecasting, etc, use this re-parameterization: Let $\eta_t = a_t^2 - \sigma_t^2$. $\{\eta_t\}$ un-correlated series.

The GARCH model becomes

$$a_t^2 = \alpha_0 + \sum_{i=1}^{\max\{m,s\}} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}$$

This is an ARMA form for the squared series a_t^2.

From this, we can show that $Var[a_t] = E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{\max\{m,s\}} (\alpha_i + \beta_i)}$

As you see, the constraint $\sum_{i=1}^{\max\{m,s\}} (\alpha_i + \beta_i) < 1$ implies a finite unconditional variance while the conditional variance evolves over time.

Now, focus on a GARCH(1,1) model

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

- Weak stationarity: $\alpha_1 \geq 0$, $\beta_1 \leq 1$, and $(\alpha_1 + \beta_1) < 1$.
- Volatility clusters

– Heavy tails: if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then

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

Again, the GARCH structure is more likely to generate extreme values for r_t compared to a normal white noise, even when the distribution of ϵ_t is Gaussian.

– For 1-step ahead forecast,

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

For multi-step ahead forecasts, use $a_t^2 = \sigma_t^2 \epsilon_t^2$ and rewrite the model as

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

2-step ahead volatility forecast

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

In general, we have

$$\sigma_h^2(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \quad \ell > 1.$$

$$\lim_{\ell \to \infty} \sigma_h^2(\ell) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)} = \mathrm{var}(a_t).$$

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

Example: Monthly excess returns of S&P 500 index starting from 1926 for 792 observations.

The fitted of a Gaussian AR(3) model

$$\begin{aligned}
\bar{r}_t &= r_t - 0.0044 \\
\bar{r}_t &= 0.082\bar{r}_{t-1} - 0.017\bar{r}_{t-2} - 0.11\bar{r}_{t-3} + a_t
\end{aligned}$$

$\hat{\sigma}_a^2 = 0.00333$. For the GARCH effects, use a GARCH(1,1) model, we have a joint estimation:

$$\begin{aligned}
\bar{r}_t &= r_t - 0.0062 \\
\bar{r}_t &= 0.042\bar{r}_{t-1} - 0.034\bar{r}_{t-2} - 0.007\bar{r}_{t-3} + a_t \\
\sigma_t^2 &= 7.37 \times 10^{-5} + 0.119a_{t-1}^2 + 0.86\sigma_{t-1}^2
\end{aligned}$$

Implied unconditional variance of a_t is

$$\frac{0.0000737}{1 - 0.119 - 0.86} = 0.00351.$$

close to the expected value. All AR coefficients are statistically insignificant.

A simplified model:

$$r_t = 0.00625 + a_t, \quad \sigma_t^2 = 7.45 \times 10^{-5} + 0.116 a_{t-1}^2 + 0.86 \sigma_{t-1}^2$$

Model checking:

For $a_t : Q(10) = 11.63(0.310)$ and $Q(20) = 24.68(0.213)$.

For $a_t^2 : Q(10) = 5.69(0.84)$ and $Q(20) = 10.92(0.948)$.

Forecast: 1-step ahead forecast:

$$\sigma_h^2(1) = 0.00008 + 0.116 a_h^2 + 0.86 \sigma_h^2$$

Horizon	1	2	3	4	5	∞
Return	0.00625	0.00625	0.00625	0.00625	0.00625	0.00625
Volatility	0.0525	0.0527	0.0529	0.0531	0.0532	$\sqrt{0.00310} = 0.0556$

R Demonstration:

```
> sp5=scan("sp500.txt")
Read 792 items
> pacf(sp5)
> 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
> m2=garchFit(~arma(3,0)+garch(1,1),data=sp5,trace=F)
> summary(m2)
Title: GARCH Modelling
Call:
garchFit(formula = ~arma(3,0)+garch(1, 1), data = sp5, trace = F)
Mean and Variance Equation:
data ~ arma(3, 0) + garch(1, 1)
[data = sp5]
Conditional Distribution: norm
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 7.708e-03 1.607e-03 4.798 1.61e-06 ***
ar1 3.197e-02 3.837e-02 0.833 0.40473
ar2 -3.026e-02 3.841e-02 -0.788 0.43076
ar3 -1.065e-02 3.756e-02 -0.284 0.77677
omega 7.975e-05 2.810e-05 2.838 0.00454 **
alpha1 1.242e-01 2.247e-02 5.529 3.22e-08 ***
beta1 8.530e-01 2.183e-02 39.075 < 2e-16 ***
---
Log Likelihood:
1272.179 normalized: 1.606287
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 73.04842 1.110223e-16
Shapiro-Wilk Test R W 0.985797 5.961994e-07
Ljung-Box Test R Q(10) 11.56744 0.315048
Ljung-Box Test R Q(15) 17.78747 0.2740039
Ljung-Box Test R Q(20) 24.11916 0.2372256
Ljung-Box Test R^2 Q(10) 10.31614 0.4132089
Ljung-Box Test R^2 Q(15) 14.22819 0.5082978
Ljung-Box Test R^2 Q(20) 16.79404 0.6663038
LM Arch Test R TR^2 13.34305 0.3446075
Information Criterion Statistics:
AIC BIC SIC HQIC
-3.194897 -3.153581 -3.195051 -3.179018
> m2=garchFit(~garch(1,1),data=sp5,trace=F)
```

```
> summary(m2)
Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 1), data = sp5, trace = F)
Mean and Variance Equation:
data ~ garch(1, 1)
[data = sp5]
Conditional Distribution: norm
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 7.450e-03 1.538e-03 4.845 1.27e-06 ***
omega 8.061e-05 2.833e-05 2.845 0.00444 **
alpha1 1.220e-01 2.202e-02 5.540 3.02e-08 ***
beta1 8.544e-01 2.175e-02 39.276 < 2e-16 ***
---
Log Likelihood:
1269.455 normalized: 1.602848
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 80.32111 0
Shapiro-Wilk Test R W 0.9850517 3.141228e-07
Ljung-Box Test R Q(10) 11.22050 0.340599
Ljung-Box Test R Q(15) 17.99703 0.262822
Ljung-Box Test R Q(20) 24.29896 0.2295768
Ljung-Box Test R^2 Q(10) 9.920157 0.4475259
Ljung-Box Test R^2 Q(15) 14.21124 0.509572
Ljung-Box Test R^2 Q(20) 16.75081 0.6690903
LM Arch Test R TR^2 13.04872 0.3655092
Information Criterion Statistics:
AIC BIC SIC HQIC
-3.195594 -3.171985 -3.195645 -3.186520
> plot(m2)
Make a plot selection (or 0 to exit):
1: Time Series
2: Conditional SD
3: Series with 2 Conditional SD Superimposed
4: ACF of Observations
5: ACF of Squared Observations
6: Cross Correlation
7: Residuals
8: Conditional SDs
9: Standardized Residuals
10: ACF of Standardized Residuals
11: ACF of Squared Standardized Residuals
12: Cross Correlation between r^2 and r
13: QQ-Plot of Standardized Residuals
Selection: 3
> predict(m2,6)
meanForecast meanError standardDeviation
1 0.007449721 0.05377242 0.05377242
2 0.007449721 0.05388567 0.05388567
3 0.007449721 0.05399601 0.05399601
4 0.007449721 0.05410353 0.05410353
5 0.007449721 0.05420829 0.05420829
6 0.007449721 0.05431038 0.05431038
```

Turn to Student-t innovation. (R output omitted.)
Estimation of degrees of freedom:

$$r_t = 0.0085 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_7$$
$$\sigma_t^2 = 0.000125 + 0.113 a_{t-1}^2 + 0.842 \sigma_{t-1}^2$$

where the estimated degrees of freedom is 7.00.
Forecasting evaluation: Not easy to do; see Andersen and Bollerslev (1998).

■ IGARCH model

An IGARCH(1,1) model:

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

For the monthly excess returns of the S&P 500 index, we have

$$r_t = 0.007 + a_t, \quad \sigma_t^2 = 0.0001 + 0.806\sigma_{t-1}^2 + 0.194a_{t-1}^2.$$

For an IGARCH(1,1) model,

$$\sigma_h^2(\ell) = \sigma_h^2(1) + (\ell - 1)\alpha_0, \quad \ell \geq 1$$

where h is the forecast origin. Note that

- Effect of $\sigma_h^2(1)$ on future volatilities is persistent.
- The volatility forecasts form a straight line with slope α_0. See Nelson (1990) for more info.
- Special case: $\alpha_0 = 0$. Volatility forecasts become a constant. This property is used in RiskMetrics to VaR calculation.

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

$$r_t = 0.0069 + a_t, \quad a_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = 0.099a_{t-1}^2 + 0.901\sigma_{t-1}^2$$

R Demonstration:

```
> source("Igarch.R")
> sp5=scan(file="sp500.txt")
> Igarch(sp5,include.mean=T)
Estimates: 0.006874402 0.9007153
Maximized log-likehood: -1258.219
Coefficient(s):
Estimate Std. Error t value Pr(>|t|)
mu 0.0068744 0.0015402 4.46332 8.07e-06 ***
beta 0.9007153 0.0158018 57.00082 < 2e-16 ***
```

■ The GARCH-M model

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

where c is referred to as risk premium, which is expected to be positive.

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

$$r_t = 4.22 \times 10^{-3} + 0.561\sigma_t^2 + a_t, \quad \sigma_t^2 = 0.814 \times 10^{-5} + 0.122a_{t-1}^2 + 0.854\sigma_{t-1}^2$$

Standard error of risk premium is 0.896 so that the estimate is not statistically significant at the usual 5% level.

R Demonstration:

```
> source("garchM.R")
> sp5=scan(file="sp500.txt")
> m1=garchM(sp5)
Maximized log-likehood: 1269.053
Coefficient(s):
Estimate Std. Error t value Pr(>|t|)
mu 4.22469e-03 2.40670e-03 1.75539 0.0791929 .
gamma 5.61297e-01 8.96194e-01 0.62631 0.5311105
omega 8.13623e-05 2.92094e-05 2.78548 0.0053449 **
alpha 1.21976e-01 2.21373e-02 5.50995 3.5893e-08 ***
beta 8.54361e-01 2.22261e-02 38.43945 < 2.22e-16 ***
```

Remarks: The R script garchM is relatively slow. It is computing intensive due to its use of a recursive loop in evaluating likelihood function.

- The Threshold GARCH (TGARCH) or GJR Model

  A TGARCH(s,m) or GJR(s,m) model is defined as

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

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

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

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

- The EGARCH model

  The idea (concept) of EGARCH model is useful. In practice, it is easier to use the TGARCH model.

  Asymmetry in responses to past positive and negative returns:

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

  with $E[g(\epsilon_t)] = 0$.

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

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

  An EGARCH(m, s) model:

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

  Some features of EGARCH models:

  - uses log transformation to relax the positiveness constraint
  - asymmetric responses

  Consider an EGARCH(1,1) model

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

  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 $\alpha_* = (1-\alpha)\alpha_0 - \sqrt{2/\pi}\gamma$.

This is a nonlinear function similar to that of the threshold AR model of Tong (1978, 1990). Specifically, we have

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

The coefficients $(\theta+\gamma)$ and $(\theta-\gamma)$ show the asymmetry in response to positive and negative a_{t-1}. The model is, therefore, nonlinear if $\theta \neq 0$. Thus, θ is referred to as the *leverage* parameter.

Focus on the function $g(\epsilon_{t-1})$. The leverage parameter θ shows the effect of the sign of a_{t-1} whereas γ denotes the magnitude effect.

See Nelson (1991) for an example of EGARCH model.

Forecasting: some recursive formula available.

Another parameterization of EGARCH models

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1\frac{|a_{t-1}| + \gamma_1 a_{t-1}}{\sigma_{t-1}} + \beta_1\ln(\sigma_{t-1}^2),$$

where γ_1 denotes the leverage effect.

Below, I re-analyze the IBM log returns by extending the data to December 2009. The sample size is 1008.

The fitted model is

$$r_t = 0.012 + a_t, \quad a_t = \epsilon_t\sigma_t,$$
$$\ln(\sigma_t^2) = -0.611 + 0.231\frac{|a_{t-1}| - 0.250a_{t-1}}{\sigma_{t-1}} + 0.92\ln(\sigma_{t-1}^2),$$

Since EGARCH and TGARCH (above) share similar objective and the latter is easier to estimate, we shall use TGARCH model.

- The Asymmetric Power ARCH (APARCH) Model

This model was introduced by Ding, Engle and Granger (1993) as a general class of volatility models. The basic form is

$$r_t = \mu_t + a_t, \quad a_t = \epsilon_t\sigma_t, \quad \epsilon_t \sim D(0,1)$$
$$\sigma_t^\delta = \omega + \sum_{i=1}^{s}\alpha_i(|a_{t-i}| + \gamma_i a_{t-i})^\delta + \sum_{j=1}^{m}\beta_j\sigma_{t-j}^\delta$$

where δ is a non-negative real number. In particular, $\delta = 2$ gives rise to the TGARCH model and $\delta = 0$ corresponds to using $\log(\sigma_t)$.

Theoretically, one can use any power δ to obtain a model. In practice, two things deserve further consideration. First, δ will also affect the specification of the mean equation, i.e., model for μ_t. Second, it is hard to interpret δ, except for some special values such as 0, 1, 2.

In R, one can fix the value of δ a priori using the subcommand *include.delta=F*, $\delta = 2$. below, I pre-fix $\delta = 2$. Thus, we can use APARCH model to estimate TGARCH model. Consider the percentage log returns of monthly IBM stock from 1926 to 2009.

R Demonstration:

```
> da=read.table("m-ibm2609.txt",header=T)
> head(da)
date ibm
1 19260130 -0.010381
.....
6 19260630 0.068493
> ibm=log(da$ibm+1)*100
> m1=garchFit(~aparch(1,1),data=ibm,trace=F,delta=2,include.delta=F)
> summary(m1)
Title:
GARCH Modelling
Call:
garchFit(formula = ~aparch(1, 1), data = ibm, delta = 2, include.delta = F,
trace = F)
Mean and Variance Equation:
data ~ aparch(1, 1)
[data = ibm]
Conditional Distribution: norm
Coefficient(s):
mu omega alpha1 gamma1 beta1
1.18659 4.33663 0.10767 0.22732 0.79468
Std. Errors: based on Hessian
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 1.18659 0.20019 5.927 3.08e-09 ***
omega 4.33663 1.34161 3.232 0.00123 **
alpha1 0.10767 0.02548 4.225 2.39e-05 ***
gamma1 0.22732 0.10018 2.269 0.02326 *
beta1 0.79468 0.04554 17.449 < 2e-16 ***
---
Log Likelihood:
-3329.177 normalized: -3.302755
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 67.07416 2.775558e-15
Shapiro-Wilk Test R W 0.9870142 8.597234e-08
Ljung-Box Test R Q(10) 16.90603 0.07646942
Ljung-Box Test R Q(15) 24.19033 0.06193099
Ljung-Box Test R Q(20) 31.89097 0.04447407
Ljung-Box Test R^2 Q(10) 4.591691 0.9167342
Ljung-Box Test R^2 Q(15) 11.98464 0.6801912
Ljung-Box Test R^2 Q(20) 14.79531 0.7879979
LM Arch Test R TR^2 7.162971 0.8466584
Information Criterion Statistics:
AIC BIC SIC HQIC
6.615430 6.639814 6.615381 6.624694
> plot(m1) <= shows normal distribution is not a good fit.
>
> m1=garchFit(~aparch(1,1),data=ibm,trace=F,delta=2,include.delta=F,cond.dist="std")
> summary(m1)
Title:
GARCH Modelling
Call:
garchFit(formula = ~aparch(1, 1), data = ibm, delta = 2, cond.dist = "std",
include.delta = F, trace = F)
Mean and Variance Equation:
data ~ aparch(1, 1)
[data = ibm]
Conditional Distribution: std
Coefficient(s):
mu omega alpha1 gamma1 beta1 shape
1.20476 3.98975 0.10468 0.22366 0.80711 6.67329
Std. Errors: based on Hessian
```

```
Error Analysis:
Estimate Std. Error t value Pr(>|t|)
mu 1.20476 0.18715 6.437 1.22e-10 ***
omega 3.98975 1.45331 2.745 0.006046 **
alpha1 0.10468 0.02793 3.747 0.000179 ***
gamma1 0.22366 0.11595 1.929 0.053738 .
beta1 0.80711 0.04825 16.727 < 2e-16 ***
shape 6.67329 1.32779 5.026 5.01e-07 ***
---
Log Likelihood:
-3310.21 normalized: -3.283938
Standardised Residuals Tests:
Statistic p-Value
Jarque-Bera Test R Chi^2 67.82336 1.887379e-15
Shapiro-Wilk Test R W 0.9869698 8.212564e-08
Ljung-Box Test R Q(10) 16.91352 0.07629962
Ljung-Box Test R Q(15) 24.08691 0.06363224
Ljung-Box Test R Q(20) 31.75305 0.04600187
Ljung-Box Test R^2 Q(10) 4.553248 0.9189583
Ljung-Box Test R^2 Q(15) 11.66891 0.7038973
Ljung-Box Test R^2 Q(20) 14.18533 0.8209764
LM Arch Test R TR^2 6.771675 0.872326
Information Criterion Statistics:
AIC BIC SIC HQIC
6.579782 6.609042 6.579711 6.590898
> plot(m1)
Make a plot selection (or 0 to exit):
1: Time Series
2: Conditional SD
3: Series with 2 Conditional SD Superimposed
4: ACF of Observations
5: ACF of Squared Observations
6: Cross Correlation
7: Residuals
8: Conditional SDs
9: Standardized Residuals
10: ACF of Standardized Residuals
11: ACF of Squared Standardized Residuals
12: Cross Correlation between r^2 and r
13: QQ-Plot of Standardized Residuals
Selection: 13
$
```

For the percentage log returns of IBM stock from 1926 to 2009, the fitted GJR model is

$$r_t = 1.20 + a_t, \quad a_t = \epsilon_t \sigma_t, \quad \epsilon_t \sim t^*_{6.67}$$
$$\sigma_t^2 = 3.99 + 0.105(|a_{t-1}| - 0.224 a_{t-1})^2 + 0.807\sigma_{t-1}^2$$

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

Note that, we can obtain the model for the log returns as

$$r_t = 0.012 + a_t, \quad a_t = \epsilon_t \sigma_t, \quad \epsilon_t \sim t^*_{6.67}$$

$$\sigma_t^2 = 3.99 \times 10^{-4} + 0.105(|a_{t-1}| - 0.224 a_{t-1})^2 + 0.807\sigma_{t-1}^2$$

The sample variance of the IBM log returns is about 0.005 and the empirical 2.5% percentile of the data is about -0.130. If we use these two quantities for σ_{t-1}^2 and a_{t-1}, respectively, then we have

$$\frac{\sigma_t^2(-)}{\sigma_t^2(+)} = \frac{0.0004 + 0.105(0.130 + 0.224 \times 0.130)^2 + +0.807 \times 0.005}{0.0004 + 0.105(0.130 - 0.224 \times 0.130)^2 + +0.807 \times 0.005} = 1.849$$

In this particular case, the negative prior return has about 85% higher impact on the conditional variance.

Another R package: *rugarch* can be used to fit volatility models too.

R Demonstration:

```
> sp5=scan("sp500.txt")
> require(rugarch)
>spec1=ugarchspec(variance.model=list(model="iGARCH",garchOrder=c(1,1)),
mean.model=list(armaOrder=c(0,0)))
> mm=ugarchfit(data=sp5,spec=spec1)
> mm
*---------------------------------*
* 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
.....
```

Package Note: We shall use *fGarch* to estimate most volatility models, but you can also work with the package *rugarch*, which can be used to estimate GRACH-M, IGARCH, and EGARCH models.

■ **Stochastic Volatility model**

A (simple) SV model is

$$a_t = \epsilon_t \sigma_t, \quad (1 - \alpha_1 B - ... - \alpha_m B^m) \ln(\sigma_t^2) = \alpha_0 + \nu_t$$

where ϵ_t's are iid N(0, 1), ν_t's are iid $N(0, \sigma_\nu^2)$, $\{\epsilon_t\}$ and $\{\nu_t\}$ are independent.

Estimation of SV model. We shall use R script *svfit.R* or the R package *stochvol* to estimate SV models.

Long-memory SV model (Not covered)

A simple LMSV is

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

where $\sigma > 0$, ϵ_t's are iid $N(0, 1)$, η_t's are iid $N(0, \sigma_\eta^2)$, and independent of ϵ_t, and $0 < d < 0.5$.

The model says

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

Thus, the $\ln(a_t^2)$ series is a Gaussian long-memory signal plus a non-Gaussian white noise; see Breidt, Crato and de Lima (1998).

Application: See Examples 3.4 and 3.5 of the textbook.

■ Alternative Approaches to Estimating Volatility

Some alternative methods:

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

Moving window:

A simple approach to capture time-varying feature of the volatility.

Hard to determine the size of the window.

Demonstration: Use the *quantmod* package to download the daily trading information of SPDR S&P 500 from January 3, 2002 to April 30,2015. The tick symbol is SPY. Use the adjusted index value to compute daily log returns of SPY. An R script, *mvwindow.R*, is available on the course web.

Instructions:

1. Download the data and save it in your R working directory.

2. Compile the program using the command: source(*mvwindow.R*)
3. To run the program: *mvol=mvwindow(rt,size)*, where *rt* denotes the return series and *size* is the size of the moving window.
4. The output is the volatility, i.e., σ_t, stored in *sigma.t*.

Demonstration shown in class.

Use of High-Frequency Data:

Suppose we like to estimate the monthly volatility of a stock return.

Data: Daily returns

Let r_t^m be the t-th month log return.

Let $\{r_{t,i}\}_{i=1}^n$ be the daily log returns within the t-th month.

Using properties of log returns, we have

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

Assuming that the conditional variance and covariance exist, we have

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

where F_{t-1} is the information available at month $t-1$ (inclusive).

Further simplification is possible under additional assumptions.

If $r_{t,i}$ is a white noise series, then

$$\mathrm{Var}[r_t^m|F_{t-1}] = n\mathrm{Var}[r_{t,1}]$$

where $\mathrm{Var}[r_{t,1}]$ can be estimated from the daily returns $\{r_{t,i}\}_{i=1}^n$ by

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

where $\bar{r}_t$ is the sample mean of the daily log returns in month t (i.e., $\bar{r}_t = \frac{\sum_{i=1}^n r_{t,i}}{n}$).

The estimated monthly volatility is then

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

If $r_{t,i}$ follows an MA(1) model, then

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2}{n-1} + 2(n-1)\mathrm{Cov}[(r_{t,1}, r_{t,2})]$$

which can be estimated by

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

Advantage: Simple

Weaknesses:

- Models for daily returns $r_{t,i}$ are unknown.
- Typically, 21 or 22 trading days in a month, resulting in a small sample size.

See Figure 40 for an illustration. Ex 3.6 of the text.

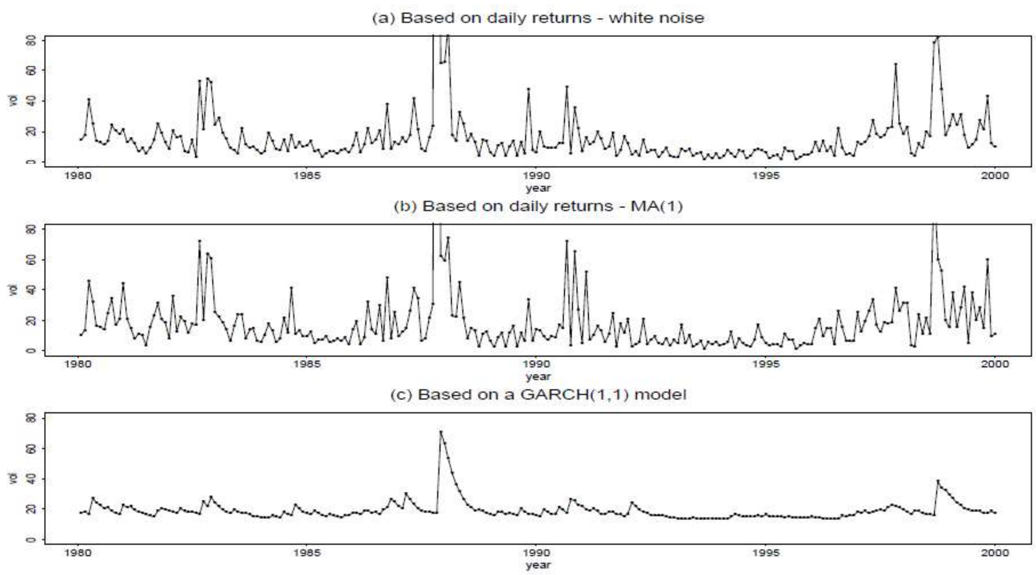

Figure 40: Time plots of estimated monthly volatility for the log returns of S&P 500 index from January 1980 to December 1999: (a) assumes that the daily log returns form a white noise series, (b) assumes that the daily log returns follow an MA(1) model, and (c) uses monthly returns from January 1962 to December 1999 and a GARCH(1,1) model.

Realized integrated volatility:

If the sample mean $\bar{r}_t$ is zero, then

$$\hat{\sigma}_m^2 \approx \sum_{i=1}^{n} r_{t,i}^2$$

Therefore, you can use cumulative sum of squares of daily log returns within a month as an estimate of monthly volatility.

Consider tick-by-tick data: Apply the idea to *intraday 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 intraday information

Weaknesses:

- Effects of market micro-structure noises
- Overlook overnight volatilities.

Further discussion

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

What next? Two approaches have been proposed:

- Optimal sampling interval: Bandi and Russell (2006). Find an optimal Δ. Or equivalently, the optimal sample size $n^* = 6.5hours/\Delta$ can be chosen as

$$n^* \approx [\frac{Q}{(\hat{\sigma}_{noise}^2)^2}]^{1/3},$$

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

- Sub-sampling: Zhang et al. (2006). Choose Δ between 10 to 20 minutes. Compute integrated volatility for each of the possible Δ-interval return series. Then, compute the average. In fact, the authors propose a so-called two scales realized volatility (TSRV) estimate. The form is

$$RV = a_n ARV_K - b_n ARV_J$$

where ARV_i denotes the average realized volatility of time interval i, a_n is a real number approaching 1 and $b_n = a_n \times n_K/n_J$, and $n_K = (n-K+1)/K$ with n is the number of transactions within the day. J can be 1 or $J \ll K$. When $J = 1$, the second term can be regarded as estimate of the noise. When K is much larger than J, the second term is typically small.

Use of Daily Open, High, Low and Close Prices

Figure 41 shows a time plot of price versus time for the tth trading day. Define

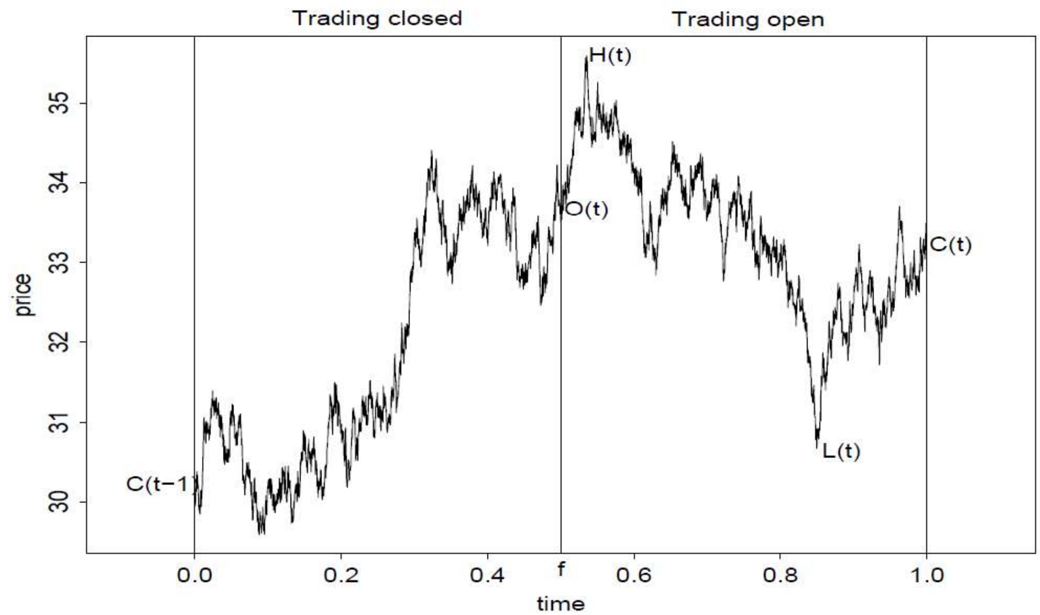

Figure 41: price versus time for the tth trading day.

- C_t: the closing price of the tth trading day;
- O_t: he opening price of the tth trading day;
- f: fraction of the day (in interval [0,1]) that trading is closed;
- H_t: the highest price of the tth trading period;
- L_t: the lowest price of the tth trading period;
- F_{t-1}: public information available at time $t-1$.

The conventional variance (or volatility) is $\sigma_t^2 = E[(C_t - C_{t-1})^2 | F_{t-1}]$.

Some alternatives:

- $\hat{\sigma}_{0,t}^2 = (C_t - C_{t-1})^2$.
- $\hat{\sigma}_{1,t}^2 = \frac{(O_t - C_{t-1})^2}{2f} + \frac{(C_t - O_t)^2}{2(1-f)}, \quad 0 < f < 1$
- $\hat{\sigma}_{2,t}^2 = \frac{(H_t - L_t)^2}{4\ln(2)} \approx 0.3607(H_t - L_t)^2$
- $\hat{\sigma}_{3,t}^2 = 0.17\frac{(O_t - C_{t-1})^2}{f} + \frac{(H_t - L_t)^2}{(1-f)4\ln(2)}, \quad 0 < f < 1$
- $\hat{\sigma}_{5,t}^2 = 0.5(H_t - L_t)^2 - [2\ln(2) - 1](C_t - O_t)^2 \approx 0.5(H_t - L_t)^2 - 0.386(C_t - O_t)^2$
- $\hat{\sigma}_{6,t}^2 = 0.12\frac{(O_t - C_{t-1})^2}{f} + 0.88\frac{\hat{\sigma}_{5,t}^2}{(1-f)}, \quad 0 < f < 1.$

A more precise, but complicated, estimator $\hat{\sigma}_{4,t}^2$ was also considered. But it is close to $\hat{\sigma}_{5,t}^2$.

Defining the efficiency factor of a volatility estimator as

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

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.

For log-return volatility, one takes the logarithms of the Open, High, Low and Close prices.

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

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

as a robust estimator of the volatility, where

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

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

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

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

This estimate seems to perform reasonably well.

Remark: One must consider the stock split in the above calculation.

Takeaway

Some alternative approaches to volatility estimation are 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.