Lecture 10

Univariate Volatility Modeling

Reading: An introduction to analysis of financial data with R (2013) by Ruey Tsay

MATH 8090 Time Series Analysis
Week 10

Whitney Huang
Clemson University

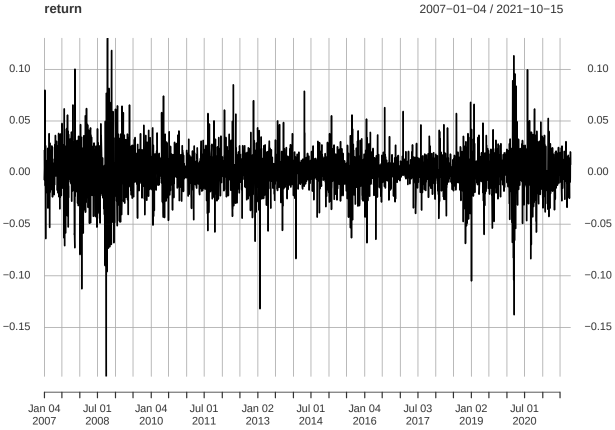
Agenda

Financial Time Series

Source: Google Finance

Log Returns of Apple Stock

$$r_t = \log(y_t/y_{t-1}),\ \text{where } y_t \text{ is the price at time } t$$



Periods of high uncertainty or rapid price changes tend to cluster together ⇒ Volatility Clustering

Modeling Volatility

Volatility is the degree of variation of a trading price series over time, usually measured by the (conditional) standard deviation of (log) returns

Why is volatility important?

- Option pricing, e.g., Black-Scholes formula
- Risk management, e.g., value at risk (VaR)
- Asset allocation, e.g., minimum-variance portfolio
- Interval forecasts

A key challenge: Not directly observable

How to Model Volatility?

We will take a econometric approach by modeling the conditional standard deviation (σ_t) of daily or monthly returns

Basic structure

$$r_t = \mu_t + a_t, \quad \mu_t = \mathbb{E}(r_t|F_{t-1}) = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{j=1}^{q} \theta_j a_{t-j}$$

Volatility models are concerned with time-evolution of

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

the conditional variance of a return

Literature on Univariate Volatility Modeling

- Autoregressive conditional heteroscedastic (ARCH) model [Engle, 1982]
- Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model [Bollerslev, 1986]
- Integrated Generalized Autoregressive Conditional heteroskedasticity (IGARCH) model
- Exponential general autoregressive conditional heteroskedastic (EGARCH) model [Nelson, 1991]
- Asymmetric parametric ARCH models [Ding, Granger, and Engle, 1994]
- Stochastic volatility (SV) models [Melino and Turnbull, 1990; Harvey, Ruiz, and Shephard, 1994; Jacqier, Polson. and Rossi, 1994]

Autoregressive Conditional Heteroscedastic (ARCH) Model

An ARCH(m) model:

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2, \quad \alpha_i \geq 0 \text{ for } 1 \leq i \leq m$$

where $\{\epsilon_t\}$ is a sequence of i.i.d. r.v. with

- $\mathbb{E}(\epsilon_t) = 0$
- $\mathrm{Var}(\epsilon_t) = 1$
- Distribution: standard normal, standardize Student-t, generalized error distribution, or their skewed counterparts

Properties of ARCH Models

Consider an ARCH(1) model

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

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$. We have the following properties:

- $\mathbb{E}(a_t) = \mathbb{E}\left[\mathbb{E}\left(a_t | F_{t-1}\right)\right] = \mathbb{E}\left[\sigma_t \mathbb{E}(\epsilon_t)\right] = 0$
- $\mathrm{Var}(a_t) = \frac{\alpha_0}{1-\alpha_1}$ if $0 < \alpha_1 < 1$
- Under normality,

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

$\Rightarrow 0 < \alpha_1^2 < \frac{1}{3} \Rightarrow$ this implies heavy tails

Building an ARCH Model

1. Modeling the mean effect μ_t and testing for ARCH effects for a_t

   $$H_0: \text{ no ARCH effects versus } H_1: \text{ ARCH effects}$$

   Use Ljung-Box test to $\{a_t^2\}$ [McLeod and Li, 1983] or Lagrange multiplier test [Engle, 1982]

2. Order determination: use PACF of the squared residuals

3. Estimation: conditional MLE

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

We use `R` packages `fGarch` and `rugarch` for volatility modeling

The Advantages And Weaknesses of ARCH Models

Advantages:

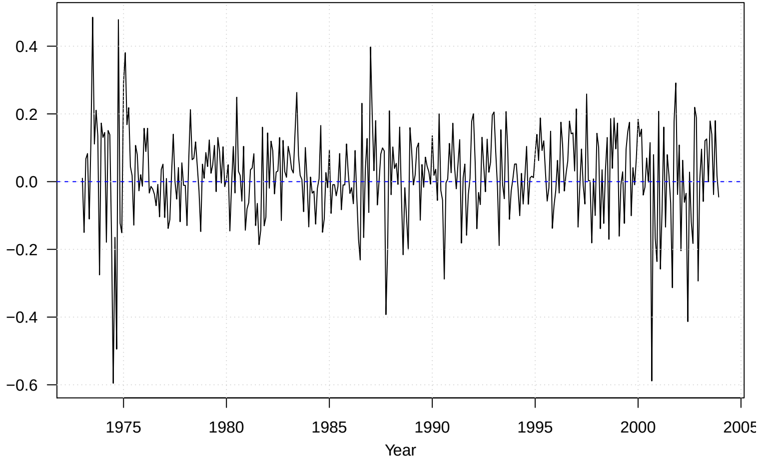
- Simplicity
- Generate volatility clustering
- Heavy tails

Weaknesses:

- Symmetric between positive and negative returns
- Restrictive (e.g., for an ARCH(1) $\alpha_1^2 \in [0, 1/3]$)
- Not sufficiently adaptive in prediction

Example: Monthly Log Returns of Intel Stock

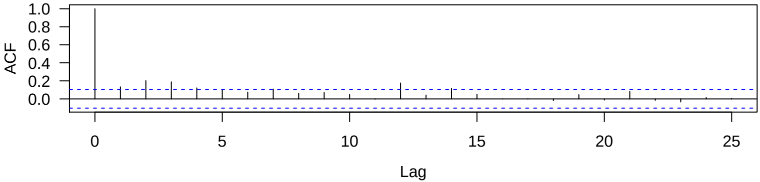
Here we use the monthly log returns of Intel stock to illustrate ARCH modeling

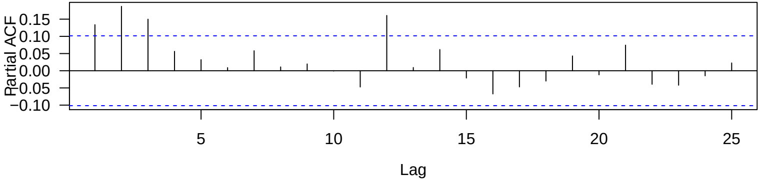


Testing ARCH Effect

Here we test and examine the temporal pattern of the squared residuals

```
> Box.test(y^2, lag = 12, type = 'Ljung')

        Box-Ljung test

data:  y^2
X-squared = 68.67, df = 12, p-value = 5.676e-10
```





ARCH Model Fitting

Here we fit an ARCH(3) for the volatility:

$$r_t = \mu + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{3} \alpha_i a_{t-i}^2,$$

assuming $\epsilon_t \overset{i.i.d.}{\sim} \mathrm{N}(0, 1)$.

```
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
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ARCH Model Fitting

Here we fit an ARCH(3) for the volatility:

$$r_t = \mu + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{3} \alpha_i a_{t-i}^2,$$

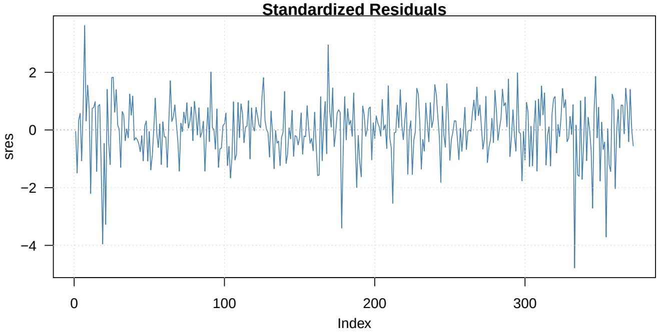
assuming $\epsilon_t \overset{i.i.d.}{\sim} \mathrm{N}(0, 1)$.

```
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
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Let’s fit a simplified ARCH(1) model

ARCH(1) Model Fitting

```
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 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```



ARCH(1) Model Checking


```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  122.404   0
 Shapiro-Wilk Test  R    W      0.9647625 8.273101e-08
 Ljung-Box Test     R    Q(10)  13.72604  0.1858587
 Ljung-Box Test     R    Q(15)  22.31714  0.09975386
 Ljung-Box Test     R    Q(20)  23.88257  0.2475594
 Ljung-Box Test     R^2  Q(10)  12.50025  0.25297
 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
```

- Jarque-Berg & Shapiro-Wilk Tests: Normality
- Lagrange multiplier (LM) Test: ARCH Effects

ARCH(1) Model with Student-t Innovations

```
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 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 242.9678    normalized:  0.6531391

Description:
 Mon Oct 18 15:27:27 2021 by user:


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  130.8931  0
 Shapiro-Wilk Test  R    W      0.9637533 5.744995e-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
```

Generalized AutoregressiveConditional Heteroskedasticity (GARCH) Model

For a log return series r_t, let $a_t = r_t - \mu_t$ be the innovation at time t. Then a_t follows a GARCH(m, s) model if

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

where $\{\epsilon_t\}$ is defined as before, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$

Re-parameterization:
Let $\eta_t = a_t^2 - \sigma_t^2$. 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

GARCH(1, 1) Model

Model:

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

Properties

- Weak stationarity if $0 \leq \alpha_1, \beta_1 \leq 1$, $(\alpha_1 + \beta_1) < 1$
- Volatility clusters
- Heavy tails if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, as

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

- 1-step ahead forecast

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

Multi-Step Ahead Forecasts

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

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

We have 2-step ahead volatility forecast

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

In general, we have

$$\begin{aligned}\sigma_h^2(\ell) &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \quad \ell > 1 \\ &= \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{\ell-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1}\sigma_h^2(1)\end{aligned}$$

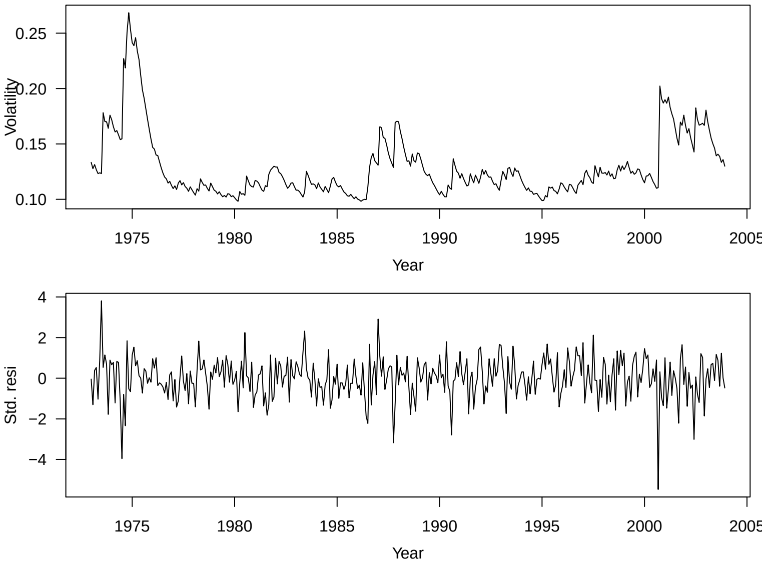
Therefore

$$\sigma_h^2(\ell) \to \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \quad \text{as } \ell \to \infty$$

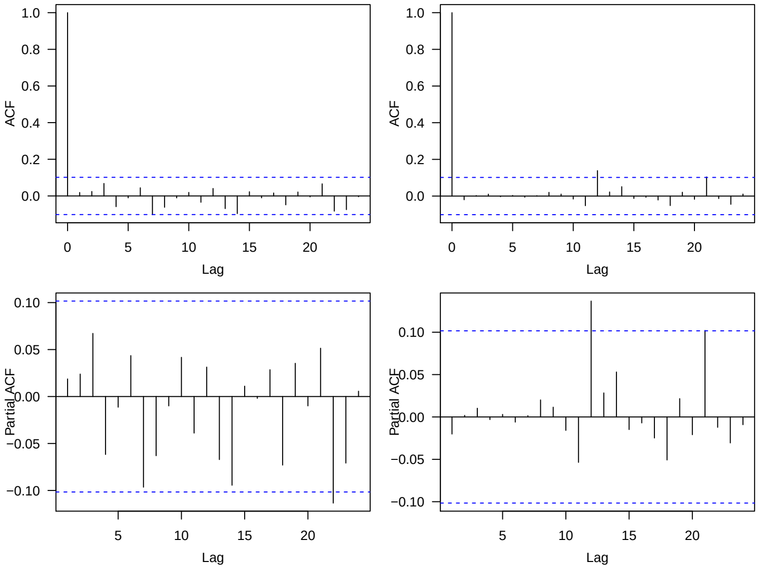
Intel Example Revisited

```
Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.0163276  0.0062624    2.607  0.00913 **
omega   0.0010918  0.0005291    2.063  0.03907 *
alpha1  0.0802716  0.0281162    2.855  0.00430 **
beta1   0.8553014  0.0461374   18.538  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 239.5189    normalized:  0.6438681

Description:
 Mon Oct 18 15:44:32 2021 by user:


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  156.5138  0
 Shapiro-Wilk Test  R    W      0.9676933 2.471139e-07
 Ljung-Box Test     R    Q(10)  9.805485  0.4577215
 Ljung-Box Test     R    Q(15)  16.54435  0.346824
 Ljung-Box Test     R    Q(20)  17.8005   0.6005484
 Ljung-Box Test     R^2  Q(10)  0.5130171 0.9999925
 Ljung-Box Test     R^2  Q(15)  10.24557  0.8040151
 Ljung-Box Test     R^2  Q(20)  11.77988  0.9234441
 LM Arch Test       R    TR^2   9.334459  0.6741288
```

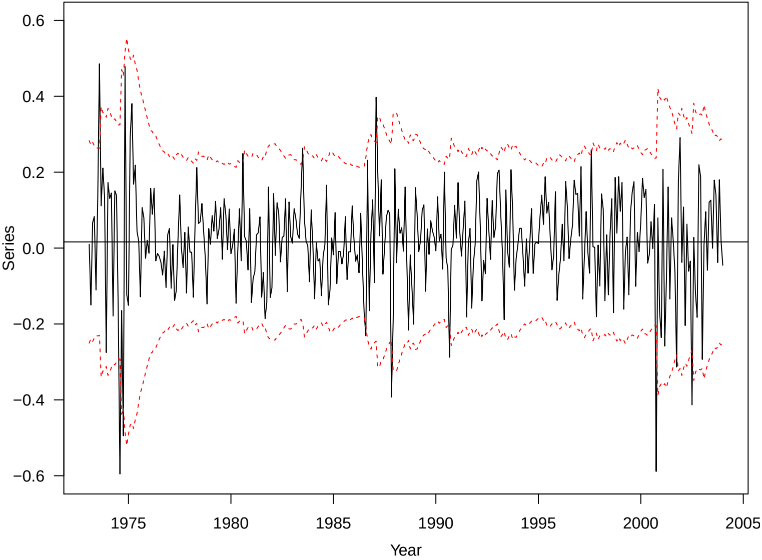
Volatility Series and Standardized Residuals



GARCH Model Checking: ACF and PACF



95% Pointwise Prediction Intervals



The Integrated GARCH Model

If the AR polynomial of the GARCH representation has unit root then we have an IGARCH model

An IGARCH(1, 1) model:

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

ℓ-step ahead forecasts

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

$\Rightarrow$ the effect of $\sigma_h^2(1)$ on future volatilities is persistent, and the volatility forecasts form a straight line with slope α_0

The Exponential GARCH Model [Nelson, 1991]

The EGARCH model is able to capture asymmetric effects between positive and negative asset returns by considering the weight innovation

$$g(\epsilon_t) = \theta\epsilon_t + \gamma\left[|\epsilon_t| - \mathbb{E}(|\epsilon_t|)\right],$$

with $\mathbb{E}[g(\epsilon_t)] = 0$

We can see the asymmetry of $g(\epsilon_t)$ by rewriting it as

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

An EGARCH(m, s) model can be written as

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

EGARCH(1, 1) Model

Model:

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

where the ϵ_t are i.i.d. standard normal. In this case, $\mathbb{E}(|\epsilon_t|) = \sqrt{\frac{2}{\pi}}$ and the model for $\log(\sigma_t^2)$ becomes

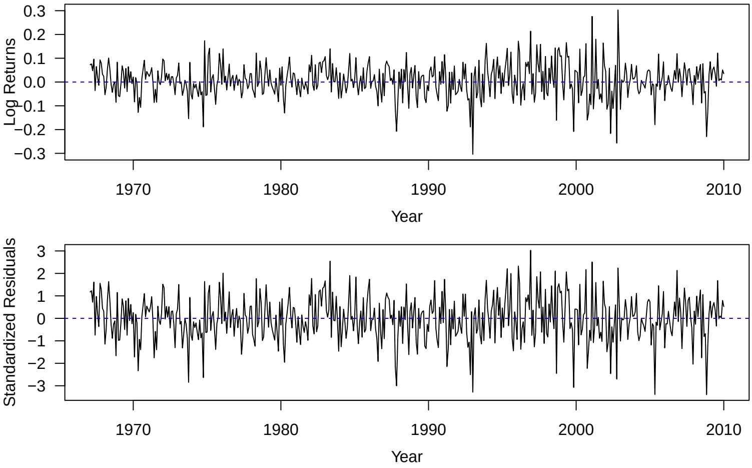
$$(1-\alpha B)\log(\sigma_t^2) = \begin{cases} (1-\alpha)\alpha_0 - \sqrt{\frac{2}{\pi}}\gamma + (\gamma+\theta)\epsilon_{t-1} & \text{if } \epsilon_{t-1} \geq 0, \\ (1-\alpha)\alpha_0 - \sqrt{\frac{2}{\pi}}\gamma + (\gamma-\theta)(-\epsilon_{t-1}) & \text{if } \epsilon_{t-1} < 0 \end{cases}$$

Finally, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\alpha} \exp\left((1-\alpha)\alpha_0 - \sqrt{\frac{2}{\pi}}\gamma\right) \begin{cases} \exp\left[(\gamma+\theta)\frac{a_{t-1}}{\sigma_{t-1}}\right] & \text{if } a_{t-1} \geq 0, \\ \exp\left[(\gamma-\theta)\frac{|a_{t-1}|}{\sigma_{t-1}}\right] & \text{if } a_{t-1} < 0. \end{cases}$$

IBM Stock Example

We consider the monthly log returns of IBM stock from January 1967 to December 2009



EGARCH(1, 1) Model Fit

$$r_t = 0.067 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim \mathrm{N}(0,1)$$

$$\log(\sigma_t^2) = -0.598 + 0.218(|\epsilon_{t-1}| - 0.423\epsilon_{t-1}) + 0.920\log(\sigma_{t-1}^2)$$

Therefore, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\times 0.920}\exp(-0.598) \times \begin{cases} \exp(0.125) & \text{if } \epsilon_{t-1} \geq 0, \\ \exp(-0.310) & \text{if } \epsilon_{t-1} < 0. \end{cases}$$

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

$$\frac{\sigma_t^2(\epsilon_{t-1} = -2)}{\sigma_t^2(\epsilon_{t-1} = 2)} = \frac{\exp\left(-0.31 \times (-2)\right)}{\exp\left(0.125 \times 2\right)} = e^{0.37} = 1.448$$

Therefore, the impact of a negative shock of size two standard deviations is about 44.8% higher than that of a positive shock of the same size

Stochastic Volatility (SV) Model

A (simple) SV model is

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

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

Long-memory SV Model:

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

where $\sigma > 0$, ϵ_t's are i.i.d. $\mathrm{N}(0, 1)$, η_t's are i.i.d. $\mathrm{N}(0, \sigma_\eta^2)$ and independent of ϵ_t, and $0 < d < 0.5$.
In LMSV, we have

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

Thus, the $\log(a_t^2)$ series is a Gaussian long-memory signal plus a non-Gaussian white noise