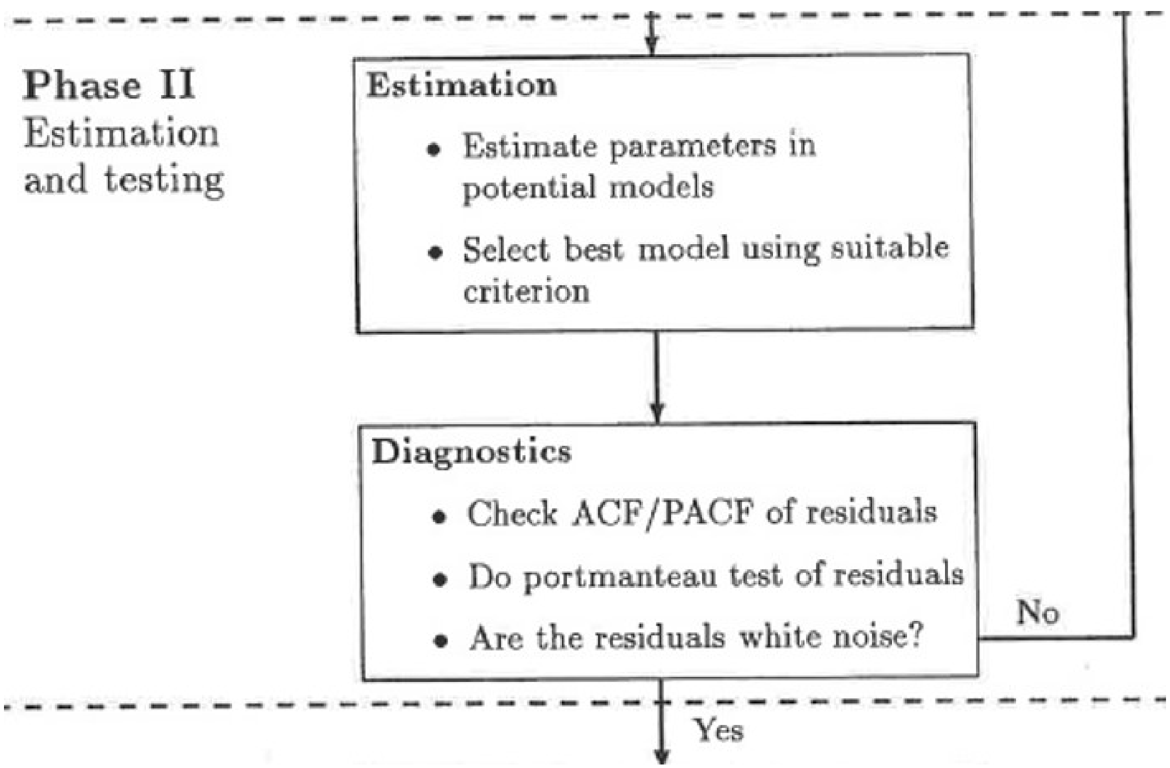
FIN307 MATLAB

FINANICAL TOOLBOX CHAPTER 1 – Financial Data Analysis (PART05)

Box-Jenkins Methodology



Estimation:
Return – ARMA
Volatility - GARCH

Model Selections:
AIC
BIC

Diagnostics:
Return: serial correlation?
Volatility: ARCH effect?

MODEL: what should be included in the models???

- Stylized facts: Sharing **non-trivial** statistical properties
- Such properties, **common** across a wide range of financial instruments

Rama Cont, *Empirical properties of asset returns: Stylized facts and statistical issues*, Quantitative Finance, 1 (2001), pp. 1–14.

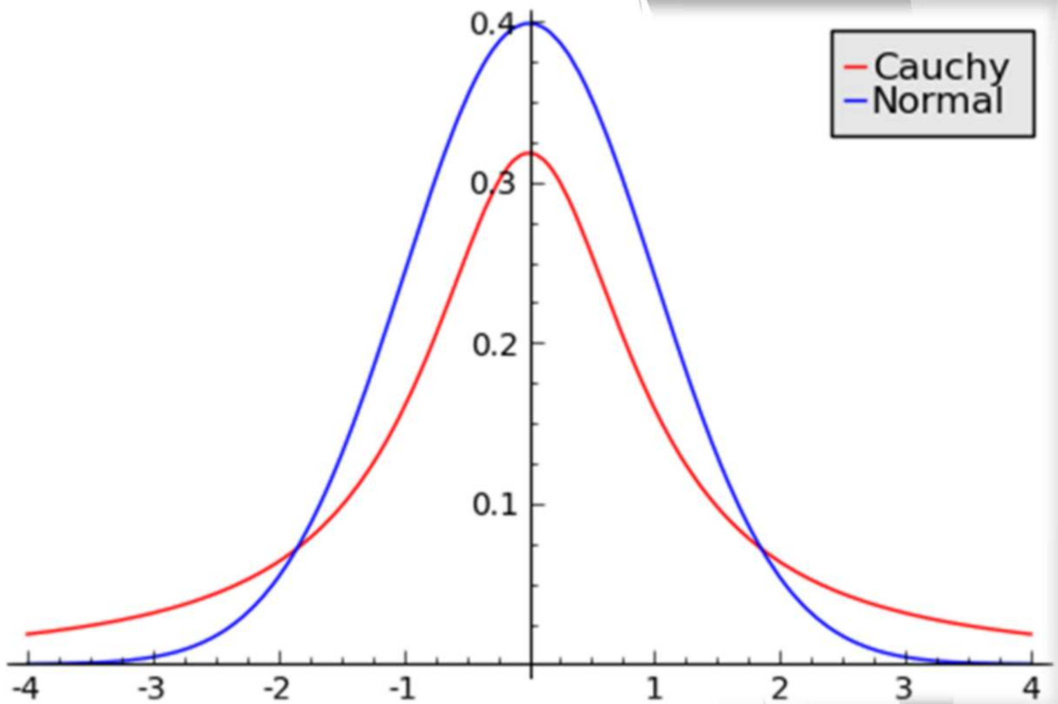
not more than 20 words in a slide!!!

Model should included all Stylized facts

Rama Cont, *Empirical properties of asset returns: Stylized facts and statistical issues*, Quantitative Finance, 1 (2001), pp. 1–14.

1. **Heavy tails:** the (unconditional) distribution of returns seems to display a power-law or Pareto-like tail, with a tail index which is finite, higher than two and less than five for most data sets studied. In particular this excludes stable laws with infinite variance and the normal distribution. However the precise form of the tails is difficult to determine.
2. **Volatility clustering:** different measures of volatility display a positive autocorrelation over several days, which quantifies the fact that high-volatility events tend to cluster in time.
3. **Slow decay of autocorrelation in absolute returns:** the autocorrelation function of absolute returns decays slowly as a function of the time lag, roughly as a power law with an exponent . This is sometimes interpreted as a sign of long-range dependence.
4. **Leverage effect:** most measures of volatility of an asset are negatively correlated with the returns of that asset.

MODELS for Return



- Return assumption: normally distributed

$$R_t = \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t, \qquad z_t \sim iid \quad N(0,1)$$

$$R_t = \sigma_t z_t$$

- Return assumption: student-t distribution

$$R_t = \varepsilon_t$$

$$\varepsilon_t = \sigma_t T_t, \qquad T_t \sim student - T(v), \qquad v\text{=degree of freedom}$$

$$R_t = \sigma_t T_t$$

$$\sqrt{\frac{\nu}{\nu - 2}} e_t \sim t_\nu, \qquad \nu > 2$$

MODELS for Return

Autoregressive (AR) Models

AR(1) $R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$ where $E[\varepsilon_t] = 0$, $Var[\varepsilon_t] = \sigma_\varepsilon^2$

not *independent* white noise.

AR(2) $R_t = \phi_0 + \phi_1 R_{t-1} + \phi_2 R_{t-2} + \varepsilon_t$

Moving Average (MA) Models

MA(1) $R_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$

where ε_t and ε_{t-1} are independent of each other and where $E[\varepsilon_t] = 0$.

MA(q) $R_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$

Combining AR and MA into ARMA Models

Parameter parsimony is key in forecasting, and combining AR and MA models into ARMA models often enables us to model dynamics with fewer parameters.

Consider the ARMA(1,1) model, which includes one lag of R_t and one lag of ε_t:

$$R_t = \phi_0 + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

The general ARMA(p, q) model is

$$R_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

MODELS for Volatility

Return assumption:

$$R_t = \varepsilon_t \qquad \varepsilon_t = \sigma_t z_t, \qquad z_t \sim iid \quad N(0,1)$$

$$R_t = \sigma_t z_t$$

THE ARCH MODEL

$$ARCH(1): \quad \sigma_{t+1}^2 = \omega + \alpha R_t^2$$

$$ARCH(2): \quad \sigma_{t+1}^2 = \omega + \alpha_1 R_t^2 + \alpha_2 R_{t-1}^2$$

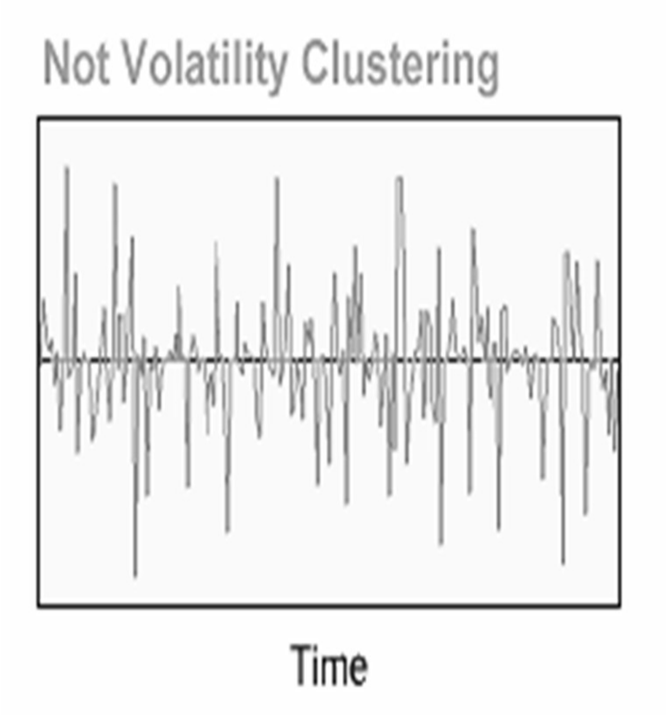
The GARCH Variance Model

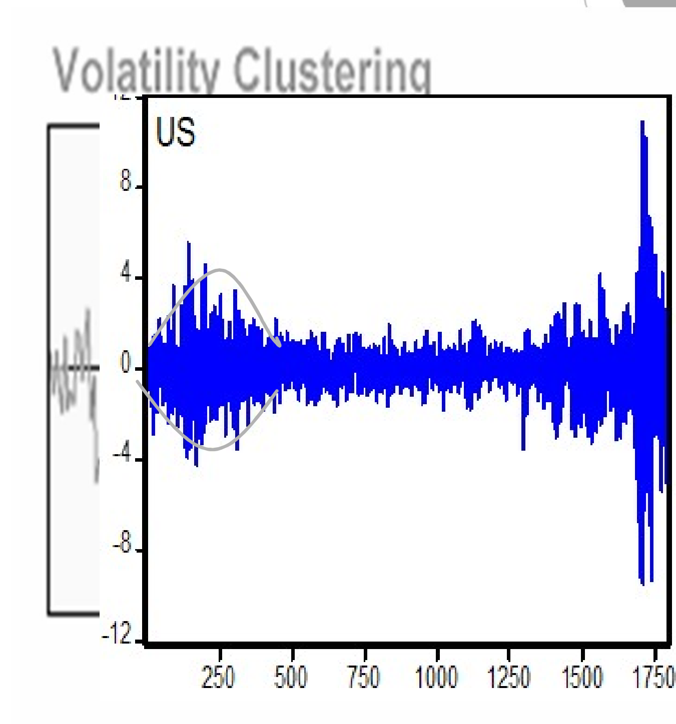
The simplest generalized autoregressive conditional heteroskedasticity (GARCH) model of dynamic variance can be written as

GARCH(1,1)

$$\sigma_{t+1}^2 = \omega + \underbrace{\alpha R_t^2}_{\text{ARCH term}} + \underbrace{\beta \sigma_t^2}_{\text{GARCH term}}, \quad \text{with } \alpha + \beta < 1$$

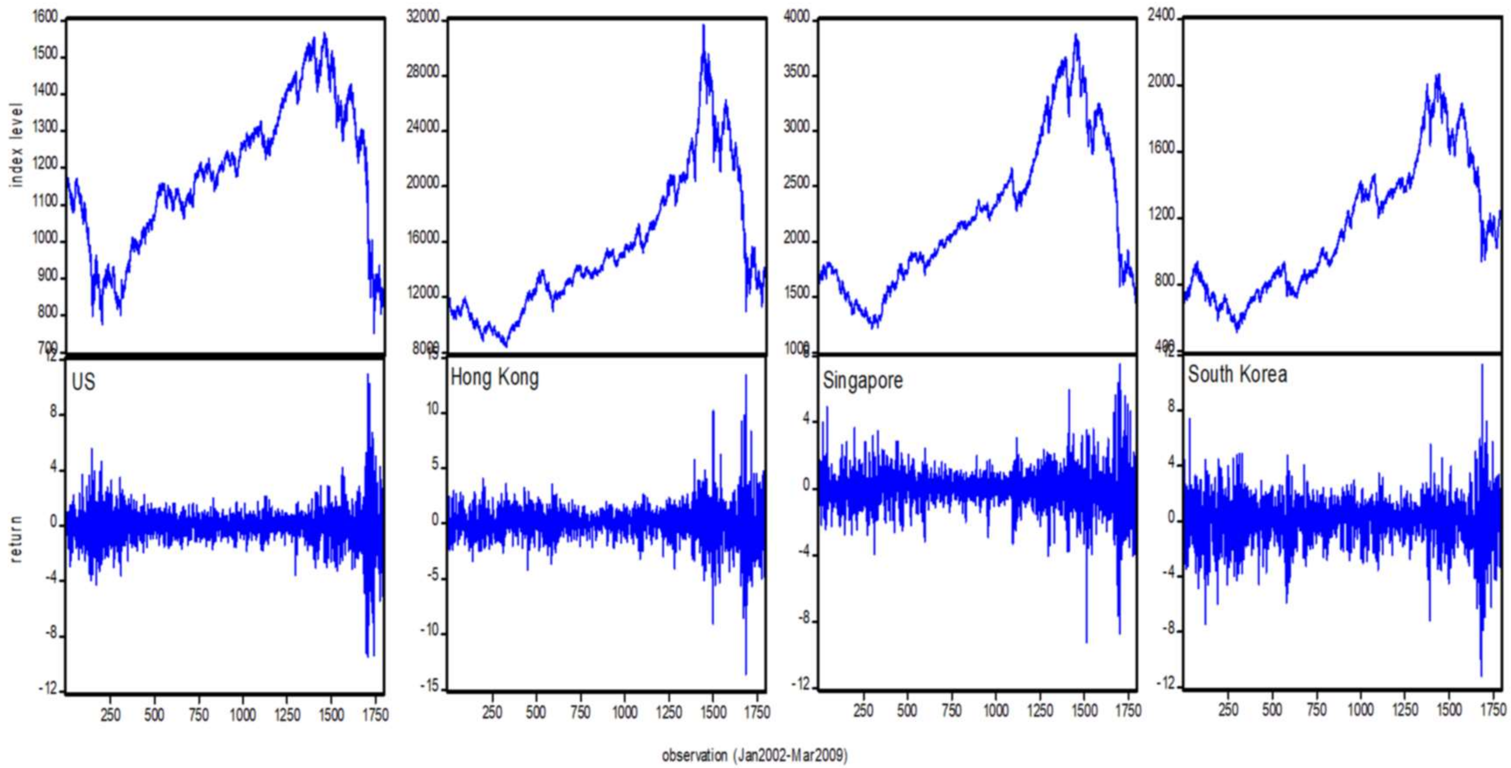
GARCH model with clustering volatility





Details: http://www.riskglossary.com

Global financial markets



Index level and return series

Why does volatility change?

- *News Announcements*: The arrival of unanticipated news (or "news surprises") forces agents to update beliefs. These new beliefs trigger portfolio rebalancing and high periods of volatility correspond to agents dynamically solving for new asset prices. While certain classes of assets have been shown to react to surprises, in particular government bonds and foreign exchange, many appear to be unaffected by even large surprises (see, *inter alia* Engle & Li (1998) and Andersen, Bollerslev, Diebold & Vega (2007)). Additionally, news-induced periods of high volatility are generally short, often on the magnitude of 5 to 30-minutes and the apparent resolution of uncertainty is far too quick to explain the time-variation of volatility seen in asset prices.

Engle, R. F. & Li, L. (1998), Macroeconomic announcements and volatility of treasury futures. UCSD Working Paper No. 97-27.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Vega, C. (2007), 'Real-time price discovery in global stock, bond and foreign exchange markets', *Journal of International Economics* 73(2), 251 – 277.

Source: Kevin Sheppard. Financial Econometrics Notes

clustering volatility

Robert Engle
Why is global financial volatility so high?
Putrajaya DEC 8-12, 2008

BRIDGES
DIALOGUES TOWARDS A CULTURE OF PEACE
Hosted through The International Peace Foundation

..”large shocks ‘tend’ to be followed by large shocks..”

- 2003 Nobel Prize in Economics
- "**A**uto**r**egressive **C**onditional **H**eteroskedasticity With Estimates of the Variance of U.K. Inflation," Econometrica 50 (1982): 987-1008.

half only...

Long memory volatility

Clive W.J. Granger

- Nobel Prize in Economic Sciences 2003;

"An introduction to long-memory time series“, with R. Joyeux, *Journal of Time Series Analysis*, 1, 1980, 15-30.

Clive W.J. Granger
Born: 4 September 1934, Swansea, United Kingdom
Died: 27 May 2009, San Diego, CA, USA
Affiliation at the time of the award: University of California
Prize motivation: "for methods of analyzing economic time series with common trends (cointegration)"
Field: Econometrics

Combination ARMA(r,s)-GARCH(p,q) model

we defined the daily asset log return, R_{t+1}, using the daily closing price, S_{t+1}, as

$$R_{t+1} \equiv \ln(S_{t+1}/S_t)$$

Example: ARMA(1,1)-GARCH(1,1) normal model

Given the assumptions made, we can write the daily return as

$$R_t = \phi_0 + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim i.i.d. \quad N(0,1)$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

Return assumption:

$$R_t = \varepsilon_t \qquad \varepsilon_t = \sigma_t z_t, \qquad z_t \sim iid \quad N(0,1)$$

$$R_t = \sigma_t z_t$$

1. Construct an ARMA-GARCH normal model

2. Construct an ARMA-GARCH student-t model

Matlab function: arima( ), estimate( ), garch( )

Estimations

Recall our assumption that

$$R_t = \sigma_t z_t, \quad \text{with } z_t \sim \text{i.i.d. } N(0,1)$$

The assumption of i.i.d. normality implies that the probability, or the likelihood, l_t, of R_t is

$$l_t = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right)$$

and thus the joint likelihood of our entire sample is

$$L = \prod_{t=1}^{T} l_t = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right)$$

A natural way to choose parameters to fit the data is then to maximize the joint likelihood of our observed sample. Recall that maximizing the logarithm of a function is equivalent to maximizing the function itself since the logarithm is a monotone, increasing function. Maximizing the logarithm is convenient because it replaces products with sums. Thus, we choose parameters $(\alpha, \beta, \ldots)$, which solve

$$Max \ln L = Max \sum_{t=1}^{T} \ln(l_t) = Max \sum_{t=1}^{T} \left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_t^2\right) - \frac{1}{2}\frac{R_t^2}{\sigma_t^2}\right]$$

Extensions to the GARCH Model

The Leverage Effect

We argued that a negative return increases variance by more than a positive return of the same magnitude. This was referred to as the leverage effect, as a negative return on a stock implies a drop in the equity value, which implies that the company becomes more highly levered and thus more risky (assuming the level of debt stays constant).

capturing the leverage effect is to define an indicator variable, I_t, to take on the value 1 if day t's return is negative and zero otherwise.

$$I_t = \begin{cases} 1, & \text{if } R_t < 0 \\ 0, & \text{if } R_t \geq 0 \end{cases}$$

The variance dynamics can now be specified as

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \alpha\theta I_t R_t^2 + \beta\sigma_t^2$$

Thus, a θ larger than zero will capture the leverage effect. This is sometimes referred to as the GJR-GARCH model.

Return assumption:

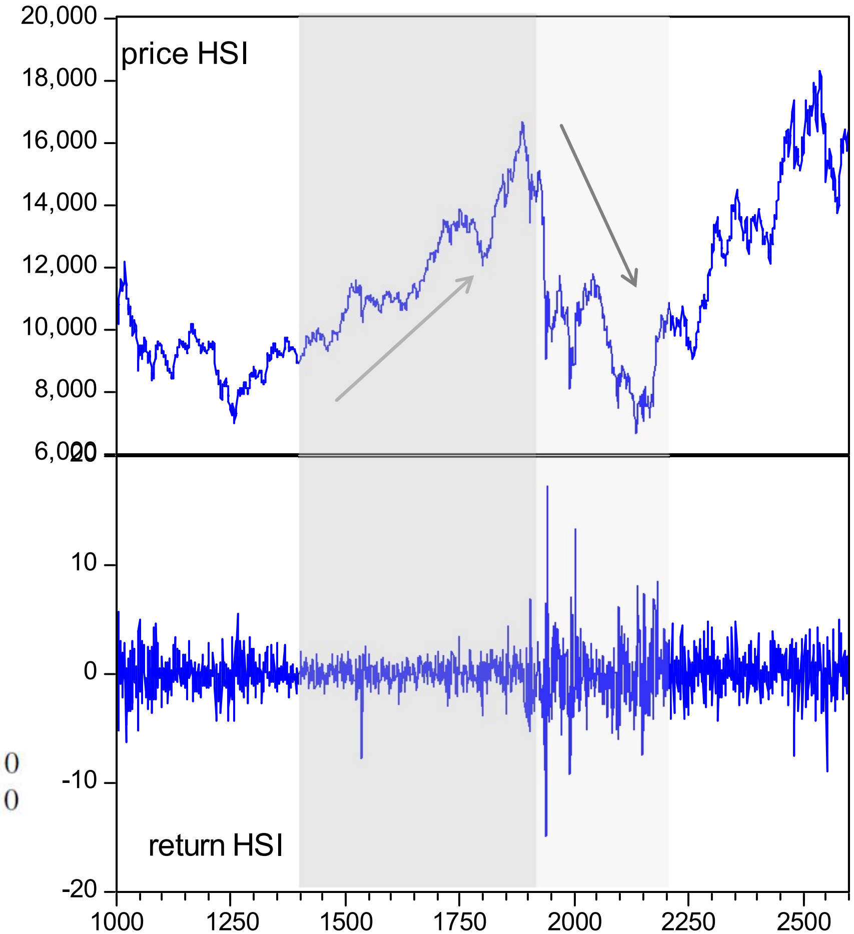
$$R_t = \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t, \qquad z_t \sim iid \quad N(0,1)$$

GARCH model with leverage effect

$$R_t = \varepsilon_t$$

$$I_t = \begin{cases} 1, & \text{if } R_t < 0 \\ 0, & \text{if } R_t \geq 0 \end{cases}$$



- "Bad news" tends to increase future market volatility more than "good news" at the same magnitude;

News ⟷ **Volatility** ⟷ **Risk**

Debt-Equity ratio

ARMA-GARCH model summary:

The following are so-called "stylized features" associated with observed time series of financial returns:

(i) the marginal distributions have heavy tails,
(ii) there is persistence of volatility,
(iii) the returns exhibit aggregational Gaussianity,
(iv) there is asymmetry with respect to negative and positive disturbances and

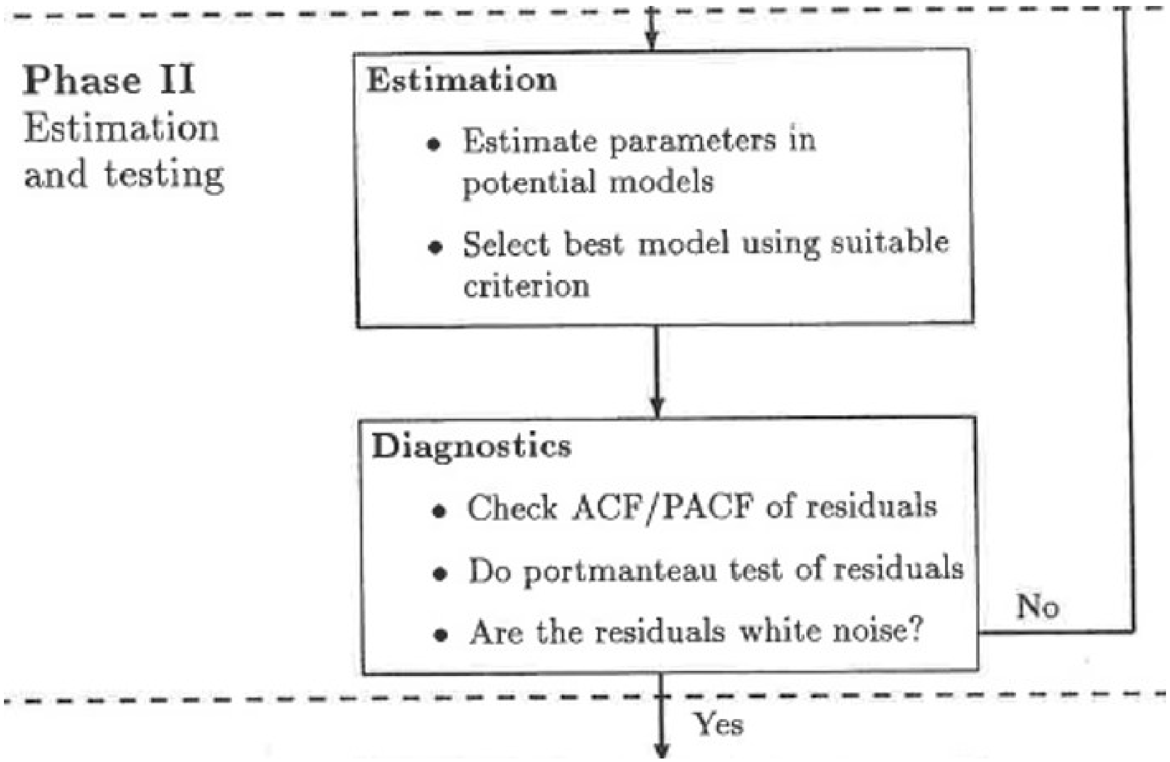
1. Construct an ARMA-GARCH normal model

2. Construct an ARMA-GARCH student-t model

3. Construct an ARMA-GJR model

Matlab function: arima( ), estimate( ), gjr( ),

Box-Jenkins Methodology



Estimation:
Return – ARMA
Volatility - GARCH

Model Selections:
AIC
BIC

Diagnostics:
Return: serial correlation?
Volatility: ARCH effect?

✓ Model selections using information criteria

$$Max\sum_{t=1}^{T}\ln(l_t) = Max\sum_{t=1}^{T}\left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_t^2\right) - \frac{1}{2}\frac{R_t^2}{\sigma_t^2}\right]$$

Akaike Information Criterion

A model fit statistic considers goodness-of-fit and parsimony. Select models that minimize AIC.

When comparing multiple model fits, additional model parameters often yield larger, optimized loglikelihood values. Unlike the optimized loglikelihood value, AIC penalizes for more complex models, i.e., models with additional parameters.

The formula for AIC, which provides insight into its relationship to the optimized loglikelihood and its penalty for complexity, is:

$$aic = -2(logL) + 2(numParam).$$

Bayesian Information Criterion

A model fit statistic considers goodness-of-fit and parsimony. Select models that minimize BIC.

Like AIC, BIC uses the optimal loglikelihood function value and penalizes for more complex models, i.e., models with additional parameters. The penalty of BIC is a function of the sample size, and so is typically more severe than that of AIC.

The formula for BIC is:

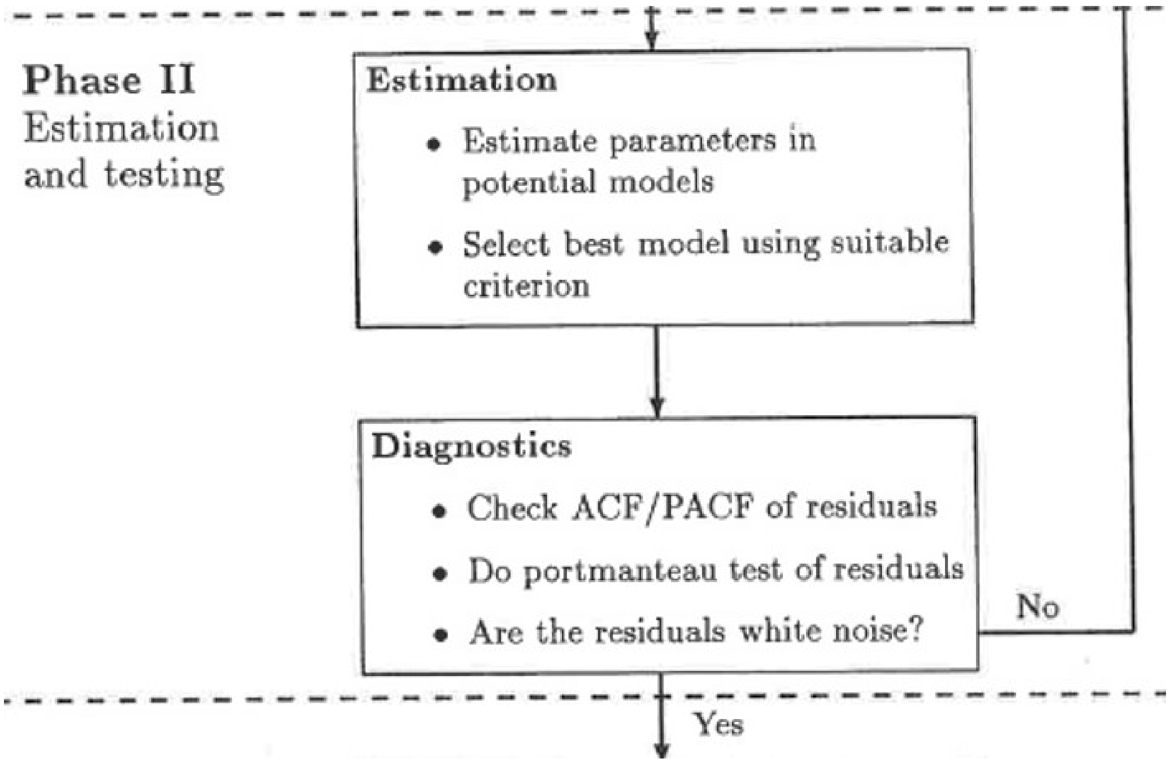
$$bic = -2(logL) + numParam * \log(numObs).$$

MathWorks®

Compare the Model Fits.

```
[resT,vT,logLT] = infer(EstMdlT,r);
[aic,bic] = aicbic([logL,logLT],[5,6],T)
```

Box-Jenkins Methodology



Estimation:
Return – ARMA
Volatility - GARCH

Model Selections:
AIC
BIC

Diagnostics:
Return: serial correlation?
Volatility: ARCH effect?

Example: return with GARCH(1,1) normal model

Given the assumptions made, we can write the daily return as

$$R_t = \varepsilon_t \qquad R_t = \varepsilon_t = residual$$

$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim i.i.d. \quad N(0,1)$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

MathWorks®

Conduct the Ljung-Box Q-test for autocorrelation at lags 5, 10, and 15.

```
[h,p,Qstat,crit] = lbqtest(Y,'Lags',[5,10,15])
```

Diagnostic

```
[h,pValue,stat,cValue] = archtest(res)
[h,pValue,stat,cValue] = archtest(res,param1,val1,param2,val2,...)
```

Residual: $\varepsilon_t = \sigma_t z_t$

Standardized residual: $\tilde{\varepsilon}_t = \dfrac{\varepsilon_t}{\sigma_t} \sim z_t \qquad z_t \sim iid \quad N(0,1)$

Return with no serial correlation?

Standardized residual-squared: $\tilde{\varepsilon}_t^2 = \left(\dfrac{\varepsilon_t}{\sigma_t}\right)^2$

Volatility with no ARCH effect?

MathWorks

Specify Conditional Mean and Variance Models

Create a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Model Selection > Nonspherical Models

Estimate Conditional Mean and Variance Model

Estimate a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Get Started with Econometrics Toolbox

Simulate Conditional Mean and Variance Models

Simulate responses and **conditional variances** from a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Conditional Mean Models

Forecast Conditional Mean and Variance Model

Forecast responses and **conditional variances** from a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Conditional Mean Models

Forecasting

MA(1) model $X_t = Z_t + \theta Z_{t-1},\ where\ Z_t \sim WN(0, \sigma^2)$

information known can thus be written as: $\boldsymbol{I_T} = \{\boldsymbol{X_1}, \boldsymbol{X_2}, \cdots, \boldsymbol{X_T}; \boldsymbol{Z_1}, \boldsymbol{Z_2}, \cdots, \boldsymbol{Z_T}\}$

$$X_{T+1} = Z_{T+1} + \theta Z_T$$

One-period ahead forecast:

$$\begin{aligned} X_{T+1,T} = E(X_{T+1}|I_T) &= E(Z_{T+1} + \theta Z_T|I_T) \\ &= E(Z_{T+1}|I_T) + \theta E(Z_T|I_T) \\ &= E(Z_{T+1}) + \theta E(Z_T|Z_T) \\ &= 0 + \theta Z_T = \theta Z_T \end{aligned}$$

Source: http://www.ams.sunysb.edu

Two-periods ahead forecast:

$$X_{T+2,T} = E(X_{T+2}|I_T)$$

$$= E(Z_{T+2} + \theta Z_{T+1}|I_T)$$

$$= E(Z_{T+2}|I_T) + \theta E(Z_{T+1}|I_T)$$

$$= 0 + \theta * 0 = 0$$

The MA(1) process is not forecastable for more than one period ahead (apart from the unconditional mean).

Source: http://www.ams.sunysb.edu

Forecast in AR(1)

Given $I_T = \{X_1, X_2, \cdots, X_T; Z_1, Z_2, \cdots, Z_T\}$

$\mathrm{X}_t = \emptyset \mathrm{X}_{t-1} + Z_t, where\ Z_t \sim WN(0, \sigma^2)$

$\mathrm{X}_{T+1} = \emptyset \mathrm{X}_T + Z_{T+1}$

given data I_T, the one period ahead forecast is

$$
\begin{aligned}
X_{T+1,T} = E(\mathrm{X}_{T+1}|I_T) &= E(\emptyset \mathrm{X}_T + Z_{T+1}|I_T) \\
&= E(\emptyset \mathrm{X}_T|I_T) + E(Z_{T+1}|I_T) \\
&= \emptyset \mathrm{X}_T + 0 \\
&= \emptyset \mathrm{X}_T
\end{aligned}
$$

Source: http://www.ams.sunysb.edu

Two Periods Ahead Forecast

$$X_{T+2} = \emptyset X_{T+1} + Z_{T+2} \qquad X_{T+1} = \emptyset X_T + Z_{T+1}$$

$$X_{T+2} = \emptyset(\emptyset X_T + Z_{T+1}) + Z_{T+2} \qquad X_{T+2} = \emptyset^2 X_T + \emptyset Z_{T+1} + Z_{T+2}$$

$$\begin{aligned} X_{T+2,T} &= E(X_{T+2}|I_T) \\ &= E(\emptyset^2 X_T + \emptyset Z_{T+1} + Z_{T+2}|I_T) \\ &= \emptyset^2 X_T + \emptyset 0 + 0 = \emptyset^2 X_T \end{aligned}$$

K Periods Ahead Forecast

$$X_{T+k} = \emptyset^k X_T + \emptyset^{k-1} Z_{T+1} + \cdots + \emptyset Z_{T+k-1} + Z_{T+k}$$

$$X_{T+k,T} = E(X_{T+k}|I_T) = \emptyset^k X_T$$

Consider the GARCH(1, 1) model $\sigma_{h+1}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2,$

1-step ahead forecast $\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2.$

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

$t = h + 1,$ $\quad \sigma_{h+2}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2(\epsilon_{h+1}^2 - 1).$

2-step ahead volatility forecast

$E(\epsilon_{h+1}^2 - 1 \mid F_h) = 0,$ $\quad \sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(1).$

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

Forecast an ARMA-GARCH normal model

Matlab function: infer(), forecast()

MathWorks®

Specify Conditional Mean and Variance Models

Create a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Model Selection > Nonspherical Models

Estimate Conditional Mean and Variance Model

Estimate a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Get Started with Econometrics Toolbox

Simulate Conditional Mean and Variance Models

Simulate responses and **conditional variances** from a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Conditional Mean Models

Forecast Conditional Mean and Variance Model

Forecast responses and **conditional variances** from a composite **conditional mean** and **variance model**.

Documentation > Econometrics Toolbox > Conditional Mean Models


THE
END