Time Series Analysis

Lecture 7

Review

1. Review of AR(1) model
2. Brownian motion and Functional central limit theorem
3. Asymptotic properties of unit root processes and tests for unit root
4. Generalization to processes with serial correlation

Today's Topics

1. ARCH
2. Extensions
3. Data Examples

Today's Topics

1. ARCH
    - ARCH
    - Maximum Likelihood Estimation with Gaussian v_t
    - Maximum Likelihood Estimation with Non-Gaussian v_t
    - Quasi-Maximum Likelihood Estimation
    - Testing for *ARCH*
    - Model Building for ARCH(m)
2. Extensions
3. Data Examples

Autoregression Conditional Heteroskedasticity(ARCH)

Consider an $AR(p)$ process,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t, \qquad (1)$$

where u_t is white noise:

$$\begin{aligned} E(u_t) &= 0, \\ E(u_t u_\tau) &= \begin{cases} \sigma^2, & \text{for } t = \tau; \\ 0, & \text{otherwise}. \end{cases} \end{aligned}$$

The process is covariance-stationary provided that the roots of

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0,$$

are outside the unit circle.The optimal linear forecast of the level of y_t for an $AR(p)$ process is

$$\hat{E}(y_t | y_{t-1}, y_{t-2}, \cdots) = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p},$$

where $\hat{E}(y_t | y_{t-1}, y_{t-2}, \cdots)$ denotes the linear projection of y_t on a constant and $(y_1, y_2, \cdots)$.

Sometimes we might be interested in forecasting not only the level of the series y_t but also its variance.

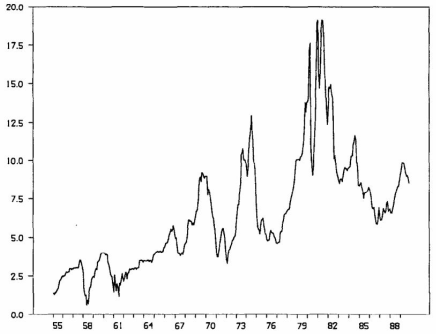

FIGURE 21.1 U.S. federal funds rate (monthly averages quoted at an annual rate), 1955–89.

Changes in the variance are quite important for understanding financial markets, since investors require higher expected returns as compensation for holding riskier assets.

Although the unconditional variance of u_t is the constant σ^2, the conditional variance of u_t could change over time. One approach is to describe the square of u_t following an $AR(m)$ process:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2 + w_t, \qquad (2)$$

where w_t is a new white noise process:

$$\begin{aligned} E(w_t) &= 0, \\ E(w_t w_\tau) &= \begin{cases} \lambda^2, & \textit{for } t = \tau; \\ 0, & \textit{otherwise}. \end{cases} \end{aligned}$$

Expression (2) implies that

$$E(u_t^2 | u_{t-1}^2, u_{t-2}^2, \cdots) = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2. \quad (3)$$

A white noise process u_t satisfying (2) is described as an autoregressive conditional heteroskedastic process of order m (Engle, 1982), denoted $u_t \sim ARCH(m)$.

Conditions for u_t

Since u_t is random and u_t^2 cannot be negative, this can be a sensible representation only if (2) is positive and (3) is nonnegative for all realizations of $\{u_t\}$. This can be ensured if w_t is bounded from below by $-\alpha_0$ with $\alpha_0 > 0$ and if $\alpha_j \geq 0$ for $j = 1, 2, \cdots, m$. In order for u_t^2 to be covariance-stationary, we further require that the roots of

$$1 - \alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_m z^m = 0,$$

lie outside the unit circle. If the α_j are all nonnegative, this is equivalent to the requirement that

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m < 1. \tag{4}$$

When these conditions are satisfied, the unconditional variance of u_t is given by

$$\sigma^2 = E(u_t^2) = \alpha_0/(1 - \alpha_1 - \alpha_2 - \cdots - \alpha_m).$$

Let $\hat{u}^2_{t+s|t}$ denote an s-period-ahead linear forecast:

$$\hat{u}^2_{t+s|t} = \hat{E}(u^2_{t+s}|u_t^2, u^2_{t-1}, \cdots).$$

This can be calculated by iterating on

$$\begin{aligned}(\hat{u}^2_{t+j|t} - \sigma^2) &= \alpha_1(\hat{u}^2_{t+j-1|t} - \sigma^2) + \alpha_2(\hat{u}^2_{t+j-2|t} - \sigma^2) \\ &\quad + \cdots + \alpha_m(\hat{u}^2_{t+j-m|t} - \sigma^2),\end{aligned}$$

for $j = 1, 2, \cdots, s$ where

$$\hat{u}^2_{\tau|t} = u^2_\tau, for\ \tau \leq t.$$

The s-period-ahead forecast $\hat{u}^2_{t+s|t}$ converges in probability to σ^2 as $s \to \infty$ assuming that w_t has finite variance and that (4) is satisfied.

$ARCH(m)$: an alternative representation

Suppose that

$$u_t = \sqrt{h_t} v_t, \tag{5}$$

where $\{v_t\}$ is an $i.i.d$ sequence with zero mean and unit variance:

$$E(v_t) = 0, \quad Var(v_t) = 1.$$

If h_t evolves according to

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2, \tag{6}$$

then (5) implies that

$$E(u_t^2 | u_{t-1}^2, u_{t-2}^2, \cdots) = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2.$$

Hence, if u_t is generated by (5) and (6), then u_t follows an $ARCH(m)$ process in which the linear projection is also the conditional expectation.

Notice further that when (5) and (6) are substituted into (2), the result is

$$\begin{aligned} h_t \cdot v_t^2 &= h_t + w_t, \\ w_t &= h_t \cdot (v_t^2 - 1). \end{aligned}$$

Note from the above equation that although the unconditional variance of w_t was assumed to be constant,

$$E(w_t^2) = \lambda^2 = E(h_t^2) \cdot E(v_t^2 - 1)^2 = E[Var(u_t^2 | I_t)],$$

the conditional variance of w_t changes over time,

$$E(w_t^2 | I_t) = h_t^2 \cdot E(v_t^2 - 1)^2 = Var(u_t^2 | I_t).$$

Taking the $ARCH(1)$ specification for illustration, we find

$$\begin{aligned}
E(h_t^2) &= E(\alpha_0 + \alpha_1 u_{t-1}^2)^2 \\
&= E\{(\alpha_1^2 \cdot u_{t-1}^4) + (2\alpha_1\alpha_0 \cdot u_{t-1}^2) + \alpha_0^2\} \\
&= \alpha_1^2[Var(u_{t-1}^2) + [E(u_{t-1}^2)]^2] + 2\alpha_1\alpha_0 \cdot E(u_{t-1}^2) + \alpha_0^2 \\
&= \alpha_1^2[\frac{\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2}] + \frac{2\alpha_1\alpha_0^2}{1-\alpha_1} + \alpha_0^2 \\
&= \frac{\alpha_1^2\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2}
\end{aligned}$$

Thus, we conclude that λ^2 must satisfy

$$\lambda^2 = [\frac{\alpha_1^2\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2}] \times E(v_t^2 - 1)^2.$$

Even when $|\alpha_1| < 1$, the above equation may not have any real solution for λ.

Implications under Gaussian v_t

For example, if $v_t \sim N(0,1)$, then $E(v_t^2-1)^2 = 2$ and requires that

$$\frac{(1-3\alpha_1^2)\lambda^2}{1-\alpha_1^2} = \frac{2\alpha_0^2}{(1-\alpha_1)^2}.$$

The equation has no real solution for λ whenever $\alpha_1^2 \geq \frac{1}{3}$. When $\alpha_1^2 < \frac{1}{3}$, we get

$$\begin{aligned}
E(u_t^4) &= Var(u_t^2) + (Eu_t^2)^2 = E[Var(u_t^2|I_t)] + Var[E(u_t^2|I_t)] + (Eu_t^2)^2 \\
&= \lambda^2 + \alpha_1^2\frac{\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2} \\
&= \frac{3\alpha_0^2(1-\alpha_1^2)}{(1-\alpha_1)^2(1-3\alpha_1^2)}.
\end{aligned}$$

As a result, u_t is leptokurtic as its kurtosis

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

Today's Topics

1. ARCH
    - ARCH
    - Maximum Likelihood Estimation with Gaussian v_t
    - Maximum Likelihood Estimation with Non-Gaussian v_t
    - Quasi-Maximum Likelihood Estimation
    - Testing for *ARCH*
    - Model Building for ARCH(m)
2. Extensions
3. Data Examples

Maximum Likelihood Estimation with Gaussian v_t

Suppose that we are interested in estimating the parameters of a regression model with *ARCH* disturbances. Let the regression equation be

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t. \tag{7}$$

Here $\mathbf{x}_t$ denotes a vector of predetermined explanatory variables, which could include lagged values of y. The disturbance term u_t is assumed to satisfy (5) and (6).

Let $\mathbf{y}_t$ denote the vector of observations obtained through date t:

$$\mathbf{y}_t = (y_t, y_{t-1}, \cdots, y_1, y_0, \cdots, y_{-m+1}, \mathbf{x}_t', \mathbf{x}_{t-1}', \cdots, \mathbf{x}_1', \mathbf{x}_0', \cdots, \mathbf{x}_{-m+1}').$$

If $v_t \sim i.i.d.N(0,1)$ with v_t independent of both $\mathbf{x}_t$ and $\mathbf{y}_{t-1}$, then the conditional distribution of y_t is Gaussian with mean $\mathbf{x}_t'\boldsymbol{\beta}$ and variance h_t:

$$f(y_t|\mathbf{x}_t, \mathbf{y}_{t-1}) = \frac{1}{\sqrt{(2\pi h_t)}} \exp\{\frac{-(y_t - \mathbf{x}_t'\boldsymbol{\beta})^2}{2h_t}\},$$

where

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1(y_{t-1} - \mathbf{x}_{t-1}'\boldsymbol{\beta})^2 + \alpha_2(y_{t-2} - \mathbf{x}_{t-2}'\boldsymbol{\beta})^2 + \cdots \\ &+ \alpha_m(y_{t-m} - \mathbf{x}_{t-m}'\boldsymbol{\beta})^2 \qquad (8) \\ &\equiv [\mathbf{z}_t(\boldsymbol{\beta})]'\boldsymbol{\delta}, \end{aligned}$$

for

$$\begin{aligned} \boldsymbol{\delta} &\equiv (\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_m)' \\ \mathbf{z}_t(\boldsymbol{\beta})' &\equiv [1, (y_{t-1} - \mathbf{x}_{t-1}'\boldsymbol{\beta})^2, (y_{t-2} - \mathbf{x}_{t-2}'\boldsymbol{\beta})^2, \cdots, (y_{t-m} - \mathbf{x}_{t-m}'\boldsymbol{\beta})^2]. \end{aligned}$$

Collect the unknown parameters to be estimated in an $(a \times 1)$ vector $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} \equiv (\boldsymbol{\beta}', \boldsymbol{\delta}')'.$$

The sample log likelihood conditional on the first m observations is then

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \sum_{t=1}^{T} \log f(y_t | \mathbf{x}_t, \mathbf{y}_{t-1}; \boldsymbol{\theta}) \\ &= -(T/2)\log(2\pi) - (1/2)\sum_{t=1}^{T} \log(h_t) \\ &\quad -(1/2)\sum_{t=1}^{T}(y_t - \mathbf{x}_t'\boldsymbol{\beta})^2/h_t. \end{aligned} \tag{9}$$

The derivative of the log of the conditional likelihood of the tth observation with respect to the parameter vector $\boldsymbol{\theta}$, known as the tth score, is given by

$$\underset{(a\times 1)}{\mathbf{s}_{\mathrm{t}}(\boldsymbol{\theta})} = \frac{\partial \log f(y_t|\mathbf{x}_t, \mathbf{y}_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

$$= \{(u_t^2 - h_t)/(2h_t^2)\} \begin{bmatrix} \sum_{j=1}^{m} -2\alpha_j u_{t-j}\mathbf{x}_{t-j} \\ \mathbf{z}_t(\boldsymbol{\beta}) \end{bmatrix} + \begin{bmatrix} (\mathbf{x}_t u_t)/h_t \\ \mathbf{0} \end{bmatrix}$$

The gradient of the log likelihood function can be calculated analytically from the sum of the scores,

$$\nabla \mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} s_t(\boldsymbol{\theta}),$$

or numerically by numerical differentiation of the log likelihood.

Today's Topics

Maximum Likelihood Estimation with Non-Gaussian v_t

The preceding formulation of the likelihood function assumed that v_t has a Gaussian distribution. However, the unconditional distribution of many financial time series seems to have fatter tails than allowed by the Gaussian family. Some of this can be explained by the presence of $ARCH$. Even so, there is a fair amount of evidence that the conditional distribution of u_t is often non-Gaussian as well.

$v_t \sim t_\nu$

Bollerslev(1987) proposed that v_t might be drawn from a t distribution with ν degrees of freedom, where ν is regarded as a parameter to be estimated by maximum likelihood. If u_t has a t distribution with ν degrees of freedom and scale parameter M_t, then its density is given by

$$f(u_t) = \frac{\Gamma[(\nu+1)/2]}{(\pi\nu)^{1/2}\Gamma(\nu/2)} M_t^{-1/2} \left[1 + \frac{u_t^2}{M_t \nu}\right]^{-(\nu+1)/2}, \qquad (10)$$

where $\Gamma(\cdot)$ is the gamma function. If $\nu > 2$, then ν_t has mean zero and variance

$$E(u_t^2) = M_t \nu(\nu - 2).$$

Hence, a t variable with ν degrees of freedom and variance h_t is obtained by taking the scale parameter M_t to be

$$M_t = h_t(\nu - 1)/\nu,$$

for which the density (10) becomes

$$f(u_t) = \frac{\Gamma[(\nu+1)/2]}{\pi^{1/2}\Gamma(\nu/2)}(\nu-2)^{-1/2}h_t^{-1/2}\left[1 + \frac{u_t^2}{h_t(\nu-2)}\right]^{-(\nu+1)/2}.$$

The sample log likelihood conditional on the first m observations then becomes

$$
\begin{aligned}
\sum_{t=1}^{T} & \log f(y_t|\mathbf{x}_t, \mathbf{y}_{t-1}; \boldsymbol{\theta}) \\
= & \; T \log \left\{ \frac{\Gamma[(\nu+1)/2]}{\pi^{1/2}\Gamma(\nu/2)} (\nu-2)^{-1/2} \right\} - (1/2) \sum_{t=1}^{T} \log(h_t) \\
& -[(\nu+1)/2] \sum_{t=1}^{T} \log \left[1 + \frac{(y_t - \mathbf{x}_t'\boldsymbol{\beta})^2}{h_t(\nu-2)} \right],
\end{aligned}
$$

where

$$
\begin{aligned}
h_t &= \alpha_0 + \alpha_1 (y_{t-1} - \mathbf{x}_{t-1}'\boldsymbol{\beta})^2 + \alpha_2 (y_{t-2} - \mathbf{x}_{t-2}'\boldsymbol{\beta})^2 + \cdots \\
&+ \alpha_m (y_{t-m} - \mathbf{x}_{t-m}'\boldsymbol{\beta})^2 \\
&= [\mathbf{z}_t(\boldsymbol{\beta})]'\boldsymbol{\alpha}_0.
\end{aligned}
$$

Other distributions for v_t

(a) Normal-Poisson mixture distribution (Jorion, 1988, RFS)

(b) Power exponential distribution (Baillie and Bollerslev, 1989, JBES)

(c) Normal-log normal mixture (Hsieh, 1989, JBES)

(d) Generalized exponential distribution (Nelson, 1991, ET)

Today's Topics

1. ARCH
   - ARCH
   - Maximum Likelihood Estimation with Gaussian v_t
   - Maximum Likelihood Estimation with Non-Gaussian v_t
   - Quasi-Maximum Likelihood Estimation
   - Testing for *ARCH*
   - Model Building for ARCH(m)
2. Extensions
3. Data Examples

Quasi-Maximum Likelihood Estimation

Maximization of the Gaussian log likelihood function(9) can provide consistent estimates of the parameters $\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_m$ of this linear representation even when the distribution of u_t is non-Gaussian, provided that v_t satisfies

$$E(v_t|\mathbf{x}_t, \mathbf{y}_{t-1}) = 0,$$

and

$$E(v_t^2|\mathbf{x}_t, \mathbf{y}_{t-1}) = 1.$$

However, the standard error have to be adjusted. Let $\hat{\boldsymbol{\theta}}_T$ be the estimate that maximizes the Gaussian log likelihood(9), and let $\boldsymbol{\theta}$ be the true value that characterizes the linear representation(5), (7) and (8). Then even when v_t is actually non-Gaussian, under certain regularity conditions,

$$\sqrt{T}(\boldsymbol{\theta}_T - \boldsymbol{\theta}) \xrightarrow{L} N(\mathbf{0}, \mathbf{D}^{-1}\mathbf{S}\mathbf{D}^{-1}),$$

where

$$\mathbf{S} = p \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} [\mathbf{s}_t(\boldsymbol{\theta})] \cdot [\mathbf{s}_t(\boldsymbol{\theta})]',$$

and

$$\begin{aligned}
\mathbf{D} &= p \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} -E\{\frac{\partial \mathbf{s}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} | \mathbf{x}_t, \mathbf{y}_{t-1}\} \\
&= p \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \left\{ [1/(2h_t^2)] \begin{bmatrix} \sum_{j=1}^{m} -2\alpha_j u_{t-j} \mathbf{x}_{t-j} \\ \mathbf{z}_t(\boldsymbol{\beta}) \end{bmatrix} \right. \\
&\times \left. \left[\sum_{j=1}^{m} -2\alpha_j u_{t-j} \mathbf{x}'_{t-j} [\mathbf{z}_j(\boldsymbol{\beta})] \right]' + (1/h_t) \begin{bmatrix} \mathbf{x}_t \mathbf{x}'_t & 0 \\ 0 & 0 \end{bmatrix} \right\}
\end{aligned}$$

with

$$\mathbf{y}_t = (y_t, y_{t-1}, \cdots, y_1, y_0, \cdots, y_{-m+1}, \mathbf{x}'_t, \mathbf{x}'_{t-1}, \cdots, \mathbf{x}'_1, \mathbf{x}'_0, \cdots, \mathbf{x}'_{-m+1})'.$$

The matrix $\mathbf{S}$ can be consistently estimated by

$$\hat{\mathbf{S}}_T = T^{-1}\sum_{t=1}^{T}[\mathbf{s}_t(\hat{\boldsymbol{\theta}}_T)]\cdot[\mathbf{s}_t(\hat{\boldsymbol{\theta}}_T)]',$$

Similarly, the matrix $\mathbf{D}$ can be consistently estimated by

$$\begin{aligned}\hat{\mathbf{D}}_T &= T^{-1}\sum_{t=1}^{T}\left\{[1/(2\hat{h}_t^2)]\begin{bmatrix}\sum_{j=1}^{m}-2\hat{\alpha}_j\hat{u}_{t-j}\mathbf{x}_{t-j}\\ \mathbf{z}_t(\hat{\boldsymbol{\beta}})\end{bmatrix}\right.\\ &\quad\times\left.\left[\sum_{j=1}^{m}-2\hat{\alpha}_j\hat{u}_{t-j}\mathbf{x}_{t-j}'[\mathbf{z}_j(\hat{\boldsymbol{\beta}})]\right]' + (1/\hat{h}_t)\begin{bmatrix}\mathbf{x}_t\mathbf{x}_t' & 0\\ 0 & 0\end{bmatrix}\right\}.\end{aligned}$$

Standard errors for $\hat{\boldsymbol{\theta}}_T$ that are robust to misspecification of the family of the densities can thus be obtained from the square root of diagonal elements of

$$T^{-1}\hat{\mathbf{D}}_T^{-1}\hat{\mathbf{S}}_T\hat{\mathbf{D}}_T^{-1}.$$

Recall that if the model is correctly specified so that the data were really generated by a Gaussian model, then $\mathbf{S} = \mathbf{D}$, and this simplifies to the usual asymptotic variance matrix for maximum likelihood estimation.

See Weiss (1984, 1986), Bollerslev and Wooldridge (1992) for detailed discussions.

Today's Topics

Testing for *ARCH*

Engle(1982) derived the following test for ARCH effect based on the Lagrange multiplier principle.

1. Estimate the regression (7) by *OLS* for observations $t = -m+1, -m+2, \cdots, T$ and obtain the residuals $\hat{u}_t$.
2. Regress $\hat{u}_t^2$ on a constant and m of its own lagged values:
$$\hat{u}_t^2 = \alpha_0 + \alpha_1 \hat{u}_{t-1}^2 + \alpha_2 \hat{u}_{t-2}^2 + \cdots + +\alpha_m \hat{u}_{t-m}^2 + e_t,$$
for $t = 1, 2, \cdots, T$.
3. Compute the statistic
$$T \cdot R_u^2 = T \cdot \left(1 - \sum \hat{e}_t^2 / \sum \hat{u}_t^4\right),$$
or
$$F = \frac{(SSR_0 - SRR_1)/m}{SRR_1/(T-m-1)}$$
which are approximately $\chi^2(m)$, under the null hypothesis that u_t is actually i.i.d $N(0, \sigma^2)$.

Today's Topics

1. ARCH
   - ARCH
   - Maximum Likelihood Estimation with Gaussian v_t
   - Maximum Likelihood Estimation with Non-Gaussian v_t
   - Quasi-Maximum Likelihood Estimation
   - Testing for *ARCH*
   - Model Building for ARCH(m)
2. Extensions
3. Data Examples

Model Building for ARCH(m)

The follow procedure may be followed for practical data analysis, if ARCH effect is of importance concern:

1. Specify a mean model by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.
2. Use the residuals of the mean equation to specify an ARCH(m) (volatility) model and test if ARCH effects are statistically significant, and determine the order m via PACF, or Ljung-Box test.
3. Perform a joint estimation of the mean and volatility equations.
4. Perform diagnostic check for the fitted model carefully and refine it if necessary.

Today's Topics

1. ARCH
2. Extensions
3. Data Examples

Today's Topics

1. ARCH

2. Extensions
   - Generalized Autoregressive Conditional Heteroskedasticity(GARCH)
   - Integrated GARCH
   - The ARCH-M Model
   - Exponential GARCH

3. Data Examples

An $ARCH(m)$ process: $u_t = \sqrt{h_t} \cdot v_t$, where v_t is i.i.d with zero mean and unit variance and where h_t evolves according to

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots \alpha_m u_{t-m}^2$$

We can imagine a process for which the conditional variance depends on an infinite number of lags of u_{t-j}^2,

$$h_t = \alpha_0 + \pi(L) u_t^2, \tag{11}$$

where

$$\pi(L) = \sum_{j=1}^{\infty} \pi_j L^j.$$

A natural idea is to parameterize $\pi(L)$ as the ratio of two finite-order polynomials:

$$\pi(L) = \frac{\alpha(L)}{1 - \delta(L)} = \frac{\alpha_1 L^1 + \alpha_2 L^2 + \cdots + \alpha_m L^m}{1 - \delta_1 L^1 - \delta_2 L^2 - \cdots - \delta_r L^r}, \tag{12}$$

where for now we assume that the roots of $1 - \delta(z) = 0$ are outside the unit circle.

If (11) is multiplied by $1 - \delta(L)$, the result is

$$[1 - \delta(L)]h_t = [1 - \delta(1)]\alpha_0 + \alpha(L)u_t^2,$$

or

$$\begin{aligned} h_t &= \kappa + \delta_1 h_{t-1} + \delta_2 h_{t-2+\cdots+\delta_r h_{t-r}} \\ &+\alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2, \end{aligned} \tag{13}$$

for $\kappa \equiv [1 - \delta_1 - \delta_2 - \cdots - \delta_r]\alpha_0$. Expression (13) is the *generalized autoregressive conditional heteroskedasticity* model, denoted $u_t \sim GARCH(r, m)$, proposed by Bollerslev(1986).

One's first guess from expression (12) and (13) might be that $\delta(L)$ describes the "autoregressive" terms for the variance while $\alpha(L)$ captures the "moving average" terms. However, this is not the case.

$u_t^2 \sim ARMA(p, r)$

To see why, add u_t^2 to both sides of (13) and rewrite the resulting expression as

$$\begin{aligned} h_t + u_t^2 &= \kappa - \delta_1(u_{t-1}^2 - h_{t-1}) - \delta_2(u_{t-2}^2 - h_{t-2}) - \cdots \\ &\quad -\delta_r(u_{t-r}^2 - h_{t-r}) + \delta_1 u_{t-1}^2 + \delta_2 u_{t-2}^2 + \cdots \\ &\quad +\delta_r u_{t-r}^2 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2 + u_t^2, \end{aligned}$$

or

$$\begin{aligned} u_t^2 &= \kappa + (\delta_1 + \alpha_1)u_{t-1}^2 + (\delta_2 + \alpha_2)u_{t-2}^2 + \cdots \\ &\quad +(\delta_p + \alpha_p)u_{t-p}^2 + w_t - \delta_1 w_{t-1} - \delta_2 w_{t-2} - \cdots - \delta_r w_{t-r}, \end{aligned}$$

where $w_t \equiv u_t^2 - h_t$ and $p \equiv max\{m, r\}$. We have further defined $\delta_j \equiv 0$ for $j > r$ and $\alpha \equiv 0$ for $j > m$.

If u_t is described by a $GARCH(r, m)$ process, the u_t^2 follows an $ARMA(p, r)$ process, where p the larger of r and m.

The nonnegative requirement is satisfied if $\kappa > 0$ and $\alpha_j \geq 0,\ \delta_j \geq 0$ for $j = 1, 2, \cdots, p$. From our analysis of *ARMA* processes, it then follows that u_t^2 is covariance-stationary provided that w_t has finite variance and that the roots of

$$1 - (\delta_1 + \alpha_1)z - (\delta_2 + \alpha_2)z^2 - \cdots - (\delta_p + \alpha_p)z^p = 0,$$

are outside the unit circle. Given the nonnegativity restriction, this means that u_t^2 is covariance-stationary if

$$(\delta_1 + \alpha_1) + (\delta_2 + \alpha_2) + \cdots + (\delta_p + \alpha_p) < 1.$$

Assume that this condition holds, the unconditional mean of u_t^2 is

$$E(u_t^2) = \sigma^2 = \kappa/[1 - (\delta_1 + \alpha_1) - (\delta_2 + \alpha_2) - \cdots - (\delta_p + \alpha_p)].$$

The forecast of u_{t+s}^2 based on $u_t^2, u_{t-1}^2, \cdots$, denoted $\hat{u}_{t+s|t}^2$ can be calculated by iterating on

$$\hat{u}_{t+s|t}^2 - \sigma^2 = \begin{cases} (\delta_1 + \alpha_1)(\hat{u}_{t+s-1|t}^2 - \sigma^2) + (\delta_2 + \alpha_2)(\hat{u}_{t+s-2|t}^2 - \sigma^2) \\ + \cdots + (\delta_p + \alpha_p)(\hat{u}_{t+s-p|t}^2 - \sigma^2) - \delta_s \hat{w}_t - \delta_{s+1} \hat{w}_{t-1} \\ - \cdots - \delta_r \hat{w}_{t+s-r} \qquad \qquad \textit{for } s = 1, 2, \cdots, r \\ (\delta_1 + \alpha_1)(\hat{u}_{t+s-1|t}^2 - \sigma^2) + (\delta_2 + \alpha_2)(\hat{u}_{t+s-2|t}^2 - \sigma^2) \\ + \cdots + (\delta_p + \alpha_p)(\hat{u}_{t+s-p|t}^2 - \sigma^2) \qquad \textit{for } s = r+1, r+2, \cdots, \end{cases}$$

$$\begin{aligned} \hat{u}_{\tau|t}^2 &= u_\tau^2 && \textit{for } \tau \leq t, \\ \hat{w}_\tau &= u_\tau^2 - \hat{u}_{\tau|\tau-1}^2 && \textit{for } \tau = t, t-1, \cdots, t-r+1. \end{aligned}$$

Calculation of the sequence of conditional variances $\{h_t\}_{t=1}^{T}$ from (13) requires presample values for $h_{-p+1}, \cdots, h_0$ and $u^2_{-p+1}, \cdots, u^2_0$. If we have observations on y_t and x_t for $t = 1, 2, \cdots, T$, Bollerslev(1986,p.316) suggested setting

$$h_j = u_j^2 = \hat{\sigma}^2 \qquad \textit{for } j = -p+1, \cdots, 0,$$

where

$$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} (y_t - \mathbf{x}_t' \boldsymbol{\beta})^2.$$

The sequence $\{h_t\}_{t=1}^{T}$ can be used to evaluate the log likelihood from the expression given in (9). This can be maximized numerically with respect to $\boldsymbol{\beta}$ and the parameters $\kappa, \delta_1, \cdots, \alpha_1, \cdots, \alpha_m$ of the *GARCH* process.

Today's Topics

Integrated GARCH

Suppose that $u_t = \sqrt{h_t} \cdot v_t$, where v_t is i.i.d with zero mean and unit variance and where h_t obeys the $GARCH(r, m)$ specification

$$\begin{aligned} h_t &= \kappa + \delta_1 h_{t-1} + \delta_2 h_{t-2} + \cdots + \delta_r h_{t-r} \\ &\quad + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2. \end{aligned}$$

The *ARMA* process for u_t^2 would have a unit root if

$$\sum_{j=1}^{r} \delta_j + \sum_{j=1}^{m} \alpha_j = 1.$$

Engle and Bollerslev (1986) referred to a model satisfying as an *integrated GARCH* process, denoted *IGARCH*.

Today's Topics

1. ARCH
2. Extensions
   - Generalized Autoregressive Conditional Heteroskedasticity(GARCH)
   - Integrated GARCH
   - The ARCH-M Model
   - Exponential GARCH
3. Data Examples

The ARCH-M Model

Finance theory suggests that an asset with a higher perceived risk would pay a higher return on average. For example, let r_t denote the ex post rate of return on some asset minus the return on a safe alternative asset. Suppose that r_t is decomposed into a component anticipated by investors as data $t-1$ (denoted μ_t) and a component that was unanticipated (denoted u_t):

$$r_t = \mu_t + u_t.$$

Then the theory suggests that the mean return(μ) would be related to the variance of the return (h_t). In general, the ARCH-in-mean, or ARCH-M, regression model introduced by Engle, Lilien, and Robins (1987) is characterized by

$$\begin{aligned} y_t &= \mathbf{x}_t' \boldsymbol{\beta} + \delta h_t + u_t, \\ u_t &= \sqrt{h_t} \cdot v_t, \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2. \end{aligned}$$

for v_t i.i.d with zero mean and unit variance.

Today's Topics

1. ARCH

2. Extensions
   - Generalized Autoregressive Conditional Heteroskedasticity(GARCH)
   - Integrated GARCH
   - The ARCH-M Model
   - Exponential GARCH

3. Data Examples

Exponential GARCH

As before, let $u_t = \sqrt{h_t} \cdot v_t$ where v_t is i.i.d with zero mean and unit variance. Nelson (1991) proposed the following model for the evolution of the conditional variance of u_t:

$$\log h_t = \alpha_0 + \sum_{j=1}^{\infty} \pi_j \cdot \{|v_{t-j}| - E|v_{t-j}| + \aleph v_{t-j}\}. \tag{14}$$

Nelson's model is sometimes referred to as exponential GARCH, or EGARCH.

If $\pi_j > 0$, Nelson's model implies that a deviation of $|v_{t-j}|$ from its expected value causes the variance of u_t to be larger than otherwise.

The Leverage Effect

The $\aleph$ parameter allows this effect to be asymmetric. If $\aleph = 0$, then a positive surprise($v_{t-j} > 0$) has the same effect on volatility as a negative surprise of the same magnitude. If $-1 < \aleph < 0$, a positive surprise increases volatility less than a negative surprise. If $\aleph < -1$, a positive surprise actually reduces volatility while a negative surprise increases volatility.

$\aleph < 0$ is sometimes described as the leverage effect.

One of the Key Advantages of *EGARCH*

One of the key advantages of Nelson's specification is that since (14) describes the log of h_t, the variance itself (h_t) will be positive regardless of whether the π_j coefficients are positive. Thus, in contrast to the *GARCH* model, no restriction need to be imposed on (14) for estimation. Nelson(1991,p.351) showed that (14) implies that $\log h_t, h_t$, and u_t are all strictly stationary provided that $\sum_{j=1}^{\infty} \pi_j^2 < \infty$.

Today's Topics

1. ARCH
2. Extensions
3. Data Examples

R codes and real data examples

- The estimation, model diagnostics and forecasting for the class of ARCH models can be performed via the R package fGARCH and rugarch. See the r-project website for the illustrations.
- An online tutorial on the rugarch is available at `http://www.dataguru.cn/article-794-1.html`
- Real data examples together with R codes can be found in Chapter 3 of Tsay (2010).