Forecasting

in Economics, Business, Finance and Beyond

CH 8: Noise: Conditional Variance Dynamics

Nonlinear Models

The stationary models so far covered in this course are *linear* In nature, that is they can be expressed as,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \{Z_t\} \sim \text{IID}(0, \sigma^2),$$

usually with $\{Z_t\}$ Gaussian. (X_t is then a **Gaussian linear process**). Such processes have a number of properties that are often found to be violated by observed time series:

Time-irreversibility. In a Gaussian linear process, $(X_t, \ldots, X_{t+h})$ has the same distribution as $(X_{t+h}, \ldots, X_t)$, for any h>0 (obs not necessarily equally spaced). Deviations from the time-reversibility property in observed time series are suggested by sample paths that rise to their maxima and fall away at different rates.

Example: SUNSPOTS.TSM.

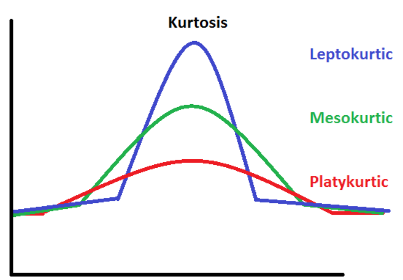
2

Kurtosis

Normal distributions are mesokurtic: kurtosis = 0

There are other mesokurtic distributions.

Leptokurtic: kurtosis > 0 Platykurtic: kurtosis < 0



3

Nonlinear Models

Bursts of outlying values are frequently observed in practical time series, and are seen also in the sample paths of nonlinear models. They are rarely seen in the sample paths of Gaussian linear processes.

Example: E1032.TSM. Daily % returns of Dow Jones Industrial Index from 7/1/97 to 4/9/99.

Changing volatility. Many observed time series, particularly financial ones, exhibit periods during which they are less predictable or more variable (volatile), depending on their past history. This dependence of predictability on past history cannot be modeled with a linear time series, since the minimum h-step MSE is independent of the past.

4

Basic Structure and Properties

Standard models (e.g., ARMA)

- Unconditional mean: constant
- Unconditional variance: constant
- Conditional mean: varies
- Conditional variance: constant (unfortunately)
- k-step-ahead forecast error variance: depends only on k, not on past observations and errors (again unfortunately)

We will look at error models that allow for variability in both the conditional and unconditional variances.

5

ARCH/GARCH

The **ARCH** and **GARCH** nonlinear models we are about to consider, do take into account the possibility that certain past histories may permit more accurate forecasting than others, and can identify the circumstances under which this can be expected to occur.

6

Distinguishing Between WN and IID Series

To distinguish between linear and nonlinear processes, we will need to be able to decide in particular when a WN sequence is also IID. (This is only an issue for non-Gaussian processes, since the two concepts coincide otherwise.)

Evidence for dependence in a WN sequence, can be obtained by looking at the ACF of the *absolute values* and/or *squares* of the process. For instance, if $\{X_t\} \sim WN(0,\sigma^2)$ with finite 4th moment, we can look at $\rho_{X^2}(h)$, the ACF of $\{X_t^2\}$ at lag h.

7

Distinguishing Between WN and IID Series

If $\rho_{X^2}(h) \neq 0$ **for some nonzero lags h**, we can conclude $\{X_t\}$ is **not IID**. This is the basis of the McLeod and Li test of residuals.

If $\rho_{X^2}(h) = 0$ for all nonzero lags h, there is insufficient evidence to conclude $\{X_t\}$ is not IID. (An IID WN sequence would have exactly this behavior.)

Similarly for $\rho_{|X|}(h) = 0$.

Example: (CHAOS.TSM). Sample ACF/PACF suggests WN.

ACF of squares & abs values suggests dependence.

Actually: $X_n = 4X_{n-1}(1 - X_{n-1})$, a deterministic (albeit chaotic) sequence!

8

The ARCH(p) Process

If P_t denotes the price of a financial series at time t, the *return at time t*, Z_t, is the relative gain, defined variously as,

$$Z_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad \text{or,} \quad Z_t = \frac{P_t}{P_{t-1}},$$

or the logs thereof. For modeling the changing volatility frequently observed in such series, Engle (1982) introduced the (now popular) ***A****uto****R****egressive* ***C****onditional* ***H****eteroscedastic process of order p*, ARCH(p), as a stationary solution, $\{Z_t\}$, of the equations, $Z_t = e_t\sqrt{h_t}, \quad \{e_t\} \sim \text{IID N}(0,1),$

with h_t, the variance of Z_t conditional on the past, given by,

$$h_t = \text{Var}[Z_t | Z_s, s < t] = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2,$$

and $\alpha_0 > 0$, and $\alpha_j \geq 0$, j=1,…,p.

9

The ARCH(p) Process

<u>Remarks</u>

- **Conditional variance, h_t**, is sometimes denoted $\boldsymbol{\sigma_t^2}$.
- If we square the first equation and subtract this equation from it, we see that an ARCH(p) satisfies,

$$Z_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + v_t,$$

where $v_t = h_t(e_t^2 - 1)$, is a WN sequence. Thus, if $E(Z_t^4) < \infty$, **the squared ARCH(p) process, $\{Z_t^2\}$, follows an AR(p).**

This fact can be used for ARCH model identification, by inspecting the sample PACF of $\{Z_t^2\}$.

10

It can be shown that $\{Z_t\}$, has mean zero, constant variance, and is uncorrelated. It is therefore **WN**, but is **not IID**, since

$$E[Z_t^2 | Z_{t-1},\ldots,Z_{t-p}] = \left(\alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2\right) E[e_t^2 | Z_{t-1},\ldots,Z_{t-p}] = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2.$$

The marginal distribution of Z_t is **symmetric**, **non-Gaussian**, and **leptokurtic** (heavy-tailed).

The ARCH(p) is **conditionally Gaussian** though, in the sense that Z_t given $Z_{t-1},\ldots, Z_{t-p}$, is Gaussian with known distribution,

$$Z_t | Z_{t-1},\ldots,Z_{t-p} \sim N(0, h_t).$$

This enables us to easily write down the likelihood of $\{Z_{p+1},\ldots, Z_n\}$, conditional on $\{Z_1,\ldots, Z_p\}$, and hence compute (conditional) ML estimates of the model parameters.

11

The conditional normality of $\{Z_t\}$ means that the **best k-step predictor** of Z_{n+k} given $Z_n,\ldots,Z_1$, is $\hat{Z}_n(k) = 0,$ with

$$\text{Var}(\hat{Z}_n(k)) = \hat{h}_n(k) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \hat{h}_n(k-i),$$

where $\hat{h}_n(k-i) = Z_{n+k-i}^2, \text{ if } k - i \leq 0.$

(This formula is to be used recursively starting with k=1.) 95% confidence bounds for the forecast are therefore

$$0 \pm 1.96\sqrt{\hat{h}_n(k)}$$

Note that using the ARCH model gives the **same** point forecasts as if it had been modeled as IID noise. The refinement occurs only for the variance of said forecasts.

For model checking, the residuals $e_t = Z_t / \sqrt{h_t} \sim$ IID N(0,1).

A **weakness** of the ARCH(p) is the fact that positive and negative shocks Z_t, have the same effect on the volatility h_t (h_t is a function of past values of Z_t^2).

12

Example: (ARCH.TSM)

Shows a realization of an ARCH(1) with α_0=1 and α_1=0.5, i.e.

$$Z_t = e_t\sqrt{1+0.5Z_{t-1}^2}, \quad \{e_t\} \sim \text{IID N}(0,1).$$

Sample ACF/PACF suggests WN, but ACF of squares and absolute values reveals dependence. In a residual analysis, only the McLeod-Li test picks up the dependence. (The normality test also rejects ; leptokurtotic.)

13

The GARCH(p,q) Process

The ***G**eneralized **ARCH**(p) process of order q*, GARCH(p,q), was introduced by Bollerslev (1986). This model is identical to ARCH(p), except that the conditional variance formula is replaced by,

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j},$$

with α_0>0, α_j≥0, β_j≥0, for j=1,2,….

Similarly to the ARCH(p), we can show that,

$$Z_t^2 = \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \beta_i) Z_{t-i}^2 + v_t - \sum_{j=1}^{q} \beta_j v_{t-j},$$

where m=max(p,q), and $v_t = h_t(e_t^2-1)$, is a WN sequence. Thus, if $\alpha_1+\ldots+\alpha_p+\beta_1+\ldots+\beta_q<1$, the squared GARCH(p,q) process, **$\{Z_t^2\}$, follows an ARMA(m,q)** with mean

$$E(Z_t^2) = \alpha_0 \Big/ 1-\sum_{i=1}^{m}(\alpha_i+\beta_i).$$

14

Although GARCH models suffer from the same weaknesses as ARCH models, they do a good job of capturing the **persistence of volatility** or **volatility clustering**, typical in stock returns, whereby small (large) values tend to be followed by small (large) values.

It is usually found that using **heavier-tailed distributions** (such as Student's t) for the process $\{e_t\}$, **provides a better fit** to financial data. (This applies equally to ARCH.) Thus more generally, and with h_t as above, we define a GARCH(p,q) process, $\{Z_t\}$, as a stationary solution of

$$Z_t = e_t\sqrt{h_t}, \quad \{e_t\} \sim \text{IID}(0,1),$$

with the distribution on $\{e_t\}$ either normal or scaled t_ν, ν>2. (The scale factor is necessary to make $\{e_t\}$ have unit variance.)

Order selection, like the ARMA case, is difficult, but should be based on AICC. Usually a GARCH(1,1) is used.

15

Apart from GARCH, several different extensions of the basic ARCH model have been proposed, each designed to accommodate a specific feature observed in practice:

Exponential GARCH (EGARCH). Allows for asymmetry in the effect of the shocks. Positive and negative returns can impact the volatility in different ways.

Integrated GARCH (IGARCH). Unit-root GARCH models similar to ARIMA models. The key feature is the long memory or persistence of shocks on the volatility.

A **plethora** of others: T-GARCH, GARCH-M, FI-GARCH; as well as ARMA models driven by GARCH noise, and regression models with GARCH errors. (*Analysis of Financial Time Series,* R.S. Tsay, 2002, Wiley.)

16

Why Do We Care?

- GARCH error models DO NOT change point estimates in forecasts, but they DO provide improved estimates of confidence bounds in forecasts.
- Allowing for time-varying volatility is crucially important in certain economic and financial contexts. The volatility of financial asset returns, for example, is often time-varying.
- The important result pre-GARCH is not the particular formulae for the unconditional mean and variance, but the fact that they are fixed, as required for covariance stationarity.
- One often sees volatility clustering, such that large changes tend to be followed by large changes, and small by small, of either sign.

17

Why Do We Care?

- The GARCH process was designed to allow for a time-varying conditional variance, but it is certainly worth emphasizing: the conditional variance is itself a serially correlated time series process.
- Real-world nancial asset returns, which are often modeled as GARCH processes, are typically unconditionally symmetric but leptokurtic.
- It turns out that the implied unconditional distribution of the conditionally Gaussian GARCH process introduced above is also symmetric and leptokurtic.

18

Why Do We Care?

- For finite h-step ahead prediction, the dependence of the prediction error variance on the current information set t can be exploited to improve interval and density forecasts.
- GARCH error processes are covariance stationary if their roots are outside the unit circle, just like ARMA processes.
- GARCH error processes express current volatility as an exponentially weighted moving average of past squared returns.

19

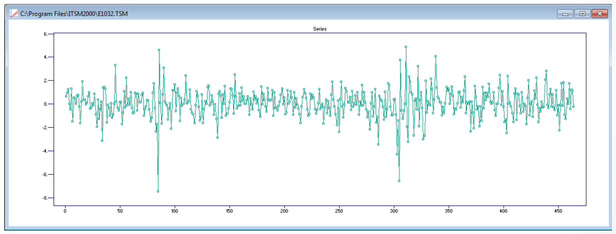
Why Do We Care?

- The temporal aggregation (aggregation over time) of GARCH processes, such as when we convert a series of daily returns to weekly returns, and then to monthly returns, then quarterly, and so on, converges toward normality. It turns out that convergence toward normality under temporal aggregation is a feature of real-world financial asset returns.
- The symmetry and leptokurtosis of the unconditional distribution of the GARCH process, as well as the disappearance of the leptokurtosis under temporal aggregation, provide nice independent confirmation of the accuracy of GARCH approximations to asset return volatility dynamics.

20

Example: GARCH Modeling (E1032.TSM)

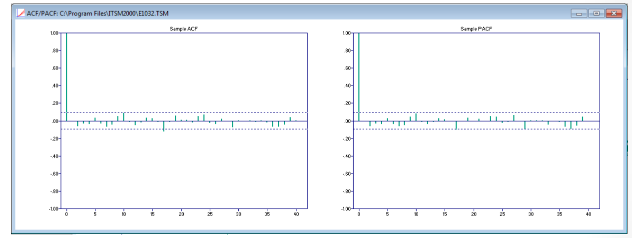
Series $\{Y_t\}$ is the percent daily returns of Dow Jones, 7/1/97 - 4/9/99. Clear periods of high (10/97, 8/98) and low volatility.



21

Example: GARCH Modeling (E1032.TSM)

Lack of autocorrelation evident in sample ACF/PACF.

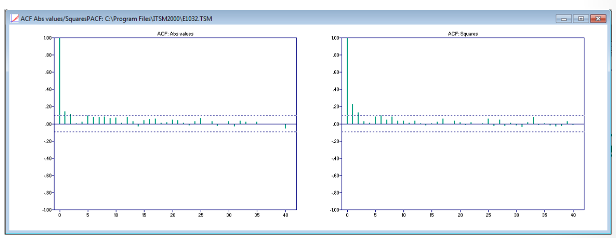


22

Example: GARCH Modeling (E1032.TSM)

Sample ACF of squares and abs values suggest dependence. This suggests fitting a model of the form

$$Y_t = a + Z_t, \quad \{Z_t\} \sim \text{GARCH(p,q)}.$$



23

Example: GARCH Modeling (E1032.TSM)

Let us fit a GARCH(1,1) to $\{Z_t\}$.

Steps in ITSM:

Specify (1,1) for model order by clicking red **GAR** button. Can choose initial values for coefficients, or use defaults. Make sure "`use normal noise`" is selected.

Red **MLE** button -> subtract mean.

Red **MLE** button several more times until estimates stabilize. Should repeat modeling with different initial estimates of coefficients to increase chances of finding the true MLEs.

24

Example: GARCH Modeling (E1032.TSM)

Comparison of models of different orders for p & q, can be made with the aid of AICC. A small search shows that the GARCH(1,1) is indeed the minimum AICC GARCH model.

Final estimates: $\hat{a} = .061,\ \hat{\alpha}_0 = .130,\ \hat{\alpha}_1 = .127,\ \hat{\beta}_0 = .792,$ with AICC=1469.0.

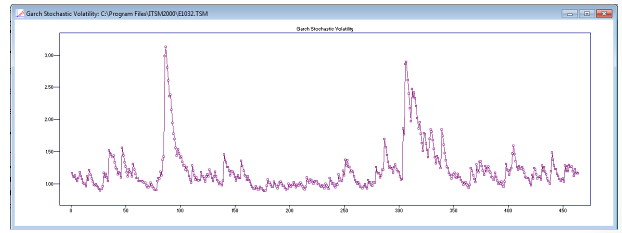
```
Subtracted Mean =  .0608
```

Garch ML estimates: C:\Program Files\ITSM2000\E1032.TSM

```
Garch Model for Z(t):

Z(t) = sqrt(h(t)) e(t)
h(t) = .1302263 + .1266649 Z^2(t-1)
     + .7919715 h(t-1)

Alpha Coefficients
       .130226        .126665

Standard Error of Alpha Coefficients
       .048636        .019031

Beta Coefficients
       .791972

Standard Error of Beta Coefficients
       .040387

AICC(Garch) = .146901E+04
```


Example: GARCH Modeling (E1032.TSM)

Red **SV** (stochastic volatility) button shows the corresponding estimates of the conditional standard deviations, $\sigma_t = \sqrt{h_t}$, confirming the changing volatility of $\{Y_t\}$.

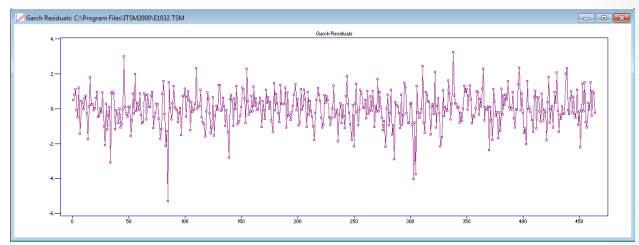


Example: GARCH Modeling (E1032.TSM)

Under the fitted model, the residuals (red **RES** button) should be approx IID N(0,1). Examine ACF of squares and abs values of residuals (**5th red** button) to check independence (OK, confirmed by McLeod-Li test). Select `Garch > Garch residuals > QQ-Plot(normal)` to check normality (expect line through origin with slope 1). Deviations from line are too large; try a heavier-tailed distribution for $\{e_t\}$.

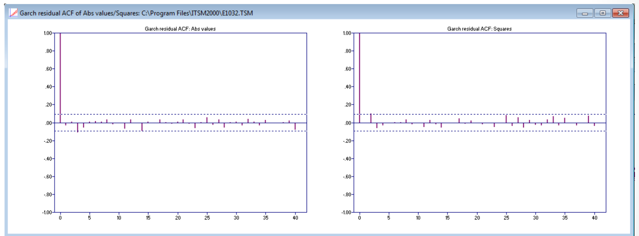
Example: GARCH Modeling (E1032.TSM)

Under the fitted model, the residuals (red **RES** button) should be approx IID N(0,1).



Example: GARCH Modeling (E1032.TSM)

Examine ACF of squares and abs values of residuals (**5th red** button) to check independence (OK)



Example: GARCH Modeling (E1032.TSM)

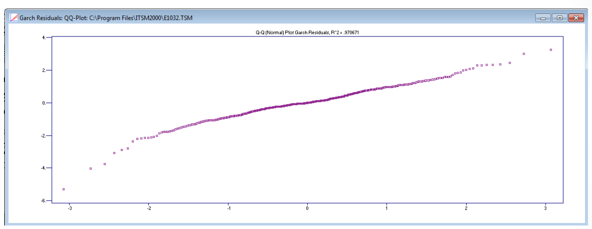
Click **GARCH -> GARCH residuals -> Tests of randomness** to confirm iid (McLeod – Li test now OK). Still fails normality.

Tests of randomness: C:\Program Files\ITSM2000\E1032.TSM

```
============================================
ITSM::(Tests of randomness on Garch residuals)
============================================

Ljung - Box statistic = 23.350 Chi-Square ( 22 ), p-value = .27194

McLeod - Li statistic = 14.799 Chi-Square ( 22 ), p-value = .87072

# Turning points = .29300E+03~AN(.30800E+03,sd = 9.0646), p-value = .09797

# Diff sign points = .22800E+03~AN(.23150E+03,sd = 6.2249), p-value = .57394

Rank test statistic = .54074E+05~AN(.53708E+05,sd = .16685E+04), p-value = .82637

Jarque-Bera test statistic (for normality) = .14516E+03 Chi-Square (2), p-value = .00000

Order of Min AICC YW Model for Garch Residuals = 1
```


Example: GARCH Modeling (E1032.TSM)

Select `Garch > Garch residuals > QQ-Plot(normal)` to check normality (expect line through origin with slope 1).



Deviations from line are too large; try a heavier-tailed distribution for $\{e_t\}$ (t-distribution).

Example: GARCH Modeling (E1032.TSM)

Repeat the modeling steps from scratch, but this time checking "`use t-distribution for noise`" in every dialog box where it appears.

Resulting min-AICC model is also GARCH(1,1), with same mean,

$$\hat{\nu} = 5.71,\ \hat{\alpha}_0 = .132,\ \hat{\alpha}_1 = .067,\ \hat{\beta}_0 = .840,$$

and AICC=1437.9 (better than previous model).

Passes residual checks, the QQ-Plot (**6th red** button) is closer to ideal line than before.

Note that even if fitting a model with t noise is what is initially desired, one should first fit a model with Gaussian noise as in this example. This will generally improve the fit.

Forecasting of volatility not yet implemented in ITSM.

Example: ARMA models with GARCH noise (SUNSPOTS.TSM)

Searching for ML ARMA model with `Autofit` gives ARMA(3,4). ACF/PACF of residuals is compatible with WN, but ACF of squares and abs values indicates they are not IID. We can fit a Gaussian GARCH(1,1) to the residuals as follows:

- Red **GAR** button > specify (1,1) for model order.
- Red **MLE** button > subtract mean.
- Red **MLE** button several more times until estimates stabilize.
- AICC for GARCH fit (805.1): use for comparing alternative GARCH models for the ARMA residuals.
- AICC adjusted for ARMA fit (821.7): use for comparing alternative ARMA models for the original data (with or without GARCH noise).

33

Example: ARMA models with GARCH noise (SUNSPOTS.TSM)

$$X_t = 2.463Z_{t-1} - 2.248Z_{t-2} + .757Z_{t-3} + Z_t - .948Z_{t-1} - .296Z_{t-2} + .313Z_{t-3} + .136Z_{t-4},$$

where

$$Z_t = \sqrt{h_t}e_t$$

and

$$h_t = 31.152 + .223Z_{t-1}^2 + .596h_{t-1}.$$

34