STAT 626: Outline Lecture 20

ARCH-GARCH Models (§5.4)

1. **Taking Care of Time-Varying Variances:** σ_t^2

2. **Time Series Decomposition:** $x_t = \mu_t + \sigma_t \varepsilon_t, \quad \mathbf{Var}(\sigma_t \varepsilon_t) = \sigma_t^2.$

3. **How to Model Time-Varying Variances?**
   Recall that Squared Residuals y_t^2 are Reasonable "Estimates" of σ_t^2:

$$y_t^2 \approx \sigma_t^2.$$

4. **Often y_t^2's appear more correlated than y_t's (Granger, 1970's).**

5. **AutoRegressive Conditionally Heteroscedastic (ARCH) Models:**(Engle, 1982)

$$y_t = \sigma_t \varepsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2.$$

   AR Models for Squared Residuals y_t^2.

   This point of view is helpful in using the ACF and PACF of the series y_t^2 to identify the orders of the ARCH(p) models.

6. **Generalized ARCH (GARCH) Models**

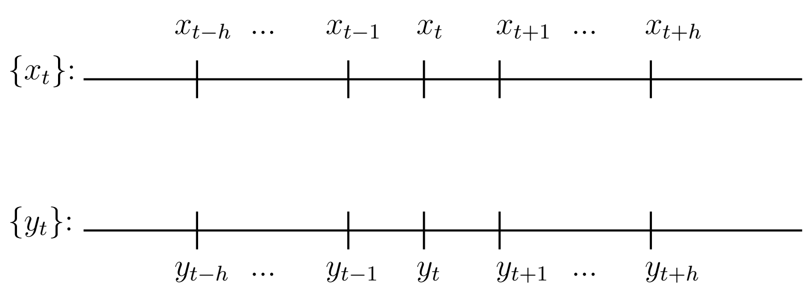
   ARMA Models for Squared Residuals y_t^2.

Coming Attractions

7. **Regression with Autocorrelated Errors (§5.6):**
   Taking Care of Correlations in the Residuals.

8. **Multivariate Time Series (§5.8):**
   Why ACF is symmetric and CCF is not?
   Proof without words!

$$\gamma_{xx}(h) = \text{Cov}(x_{t+h}, x_t) = \text{Cov}(x_t, x_{t+h}) = \gamma_{xx}(-h), \quad h = 1, 2, \ldots.$$



$$\gamma_{xy}(h) = \text{Cov}(x_{t+h}, y_t)$$

$$\gamma_{xy}(-h) = \text{Cov}(x_{t-h}, y_t) \neq \gamma_{xy}(h), \;\; h = 1, 2, \ldots$$

interest here is testing the null that $(\beta_1, \phi) = (0, 1)$, simultaneously, versus the alternative that $\beta_1 \neq 0$ and $|\phi| < 1$. In this case, the null hypothesis is that the process is a random walk with drift, versus the alternative hypothesis that the process is stationary around a global trend (consider the global temperature series examined in Example 2.1).

Example 5.3 Testing Unit Roots in the Glacial Varve Series

In this example we use the R package `tseries` to test the null hypothesis that the log of the glacial varve series has a unit root, versus the alternate hypothesis that the process is stationary. We test the null hypothesis using the available DF, ADF and PP tests; note that in each case, the general regression equation incorporates a constant and a linear trend. In the ADF test, the default number of AR components included in the model, say k, is $[\![(n-1)^{\frac{1}{3}}]\!]$, which corresponds to the suggested upper bound on the rate at which the number of lags, k, should be made to grow with the sample size for the general ARMA(p, q) setup. For the PP test, the default value of k is $[\![.04n^{\frac{1}{4}}]\!]$.

```
library(tseries)
adf.test(log(varve), k=0)                    # DF test
```

```
Dickey-Fuller = -12.8572, Lag order = 0, p-value < 0.01
alternative hypothesis: stationary
```

```
adf.test(log(varve))                         # ADF test
```

```
Dickey-Fuller = -3.5166, Lag order = 8, p-value = 0.04071
alternative hypothesis: stationary
```

```
pp.test(log(varve))                          # PP test
```

```
Dickey-Fuller Z(alpha) = -304.5376,
Truncation lag parameter = 6, p-value < 0.01
alternative hypothesis: stationary
```

In each test, we reject the null hypothesis that the logged varve series has a unit root. The conclusion of these tests supports the conclusion of the previous section that the logged varve series is long memory rather than integrated.

5.4 GARCH Models

Recent problems in finance have motivated the study of the volatility, or variability, of a time series. Although ARMA models assume a constant variance, models such as the autoregressive conditionally heteroscedastic or ARCH model, first introduced by Engle (1982), were developed to model changes in volatility. These models were later extended to generalized ARCH, or GARCH models by Bollerslev (1986).

In §3.8, we discussed the return or growth rate of a series. For example, if x_t is the value of a stock at time t, then the return or relative gain, y_t, of the stock at time t is

$$y_t = \frac{x_t - x_{t-1}}{x_{t-1}}. \tag{5.34}$$

Definition (5.34) implies that $x_t = (1+y_t)x_{t-1}$. Thus, based on the discussion in §3.8, if the return represents a small (in magnitude) percentage change then

$$\nabla[\log(x_t)] \approx y_t. \tag{5.35}$$

Either value, $\nabla[\log(x_t)]$ or $(x_t - x_{t-1})/x_{t-1}$, will be called the return, and will be denoted by y_t. It is the study of y_t that is the focus of ARCH, GARCH, and other volatility models. Recently there has been interest in stochastic volatility models and we will discuss these models in Chapter 6 because they are state-space models.

Typically, for financial series, the return y_t, does not have a constant conditional variance, and highly volatile periods tend to be clustered together. In other words, there is a strong dependence of sudden bursts of variability in a return on the series own past. For example, Figure 1.4 shows the daily returns of the New York Stock Exchange (NYSE) from February 2, 1984 to December 31, 1991. In this case, as is typical, the return y_t is fairly stable, except for short-term bursts of high volatility.

The simplest ARCH model, the ARCH(1), models the return as

$$y_t = \sigma_t \epsilon_t \tag{5.36}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2, \tag{5.37}$$

where ϵ_t is standard Gaussian white noise; that is, $\epsilon_t \sim \text{iid N}(0,1)$. As with ARMA models, we must impose some constraints on the model parameters to obtain desirable properties. One obvious constraint is that α_1 must not be negative, or else σ_t^2 may be negative.

As we shall see, the ARCH(1) models return as a white noise process with nonconstant conditional variance, and that conditional variance depends on the previous return. First, notice that the conditional distribution of y_t given y_{t-1} is Gaussian:

$$y_t \mid y_{t-1} \sim \text{N}(0, \alpha_0 + \alpha_1 y_{t-1}^2). \tag{5.38}$$

In addition, it is possible to write the ARCH(1) model as a non-Gaussian AR(1) model in the square of the returns y_t^2. First, rewrite (5.36)-(5.37) as

$$y_t^2 = \sigma_t^2 \epsilon_t^2$$

$$\alpha_0 + \alpha_1 y_{t-1}^2 = \sigma_t^2,$$

and subtract the two equations to obtain

$$y_t^2 - (\alpha_0 + \alpha_1 y_{t-1}^2) = \sigma_t^2 \epsilon_t^2 - \sigma_t^2.$$

Now, write this equation as

$$y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + v_t, \tag{5.39}$$

This the connection with AR(1) for sq. res.

where $v_t = \sigma_t^2(\epsilon_t^2 - 1)$. Because ϵ_t^2 is the square of a N$(0,1)$ random variable, $\epsilon_t^2 - 1$ is a shifted (to have mean-zero), χ_1^2 random variable.

To explore the properties of ARCH, we define $Y_s = \{y_s, y_{s-1}, ...\}$. Then, using (5.38), we immediately see that y_t has a zero mean:

$$E(y_t) = EE(y_t \bigm| Y_{t-1}) = EE(y_t \bigm| y_{t-1}) = 0. \tag{5.40}$$

Because $E(y_t \bigm| Y_{t-1}) = 0$, the process y_t is said to be a *martingale difference*.

Because y_t is a martingale difference, it is also an uncorrelated sequence. For example, with $h > 0$,

$$\begin{aligned} \text{cov}(y_{t+h}, y_t) &= E(y_t y_{t+h}) = EE(y_t y_{t+h} \,|\, Y_{t+h-1}) \\ &= E\left\{y_t E(y_{t+h} \,|\, Y_{t+h-1})\right\} = 0. \end{aligned} \tag{5.41}$$

The last line of (5.41) follows because y_t belongs to the information set Y_{t+h-1} for $h > 0$, and, $E(y_{t+h} \,|\, Y_{t+h-1}) = 0$, as determined in (5.40).

An argument similar to (5.40) and (5.41) will establish the fact that the error process v_t in (5.39) is also a martingale difference and, consequently, an uncorrelated sequence. If the variance of v_t is finite and constant with respect to time, and $0 \le \alpha_1 < 1$, then based on Property 3.1, (5.39) specifies a causal AR(1) process for y_t^2. Therefore, $E(y_t^2)$ and $\text{var}(y_t^2)$ must be constant with respect to time t. This, implies that

$$E(y_t^2) = \text{var}(y_t) = \frac{\alpha_0}{1 - \alpha_1} \tag{5.42}$$

and, after some manipulations,

$$E(y_t^4) = \frac{3\alpha_0^2}{(1-\alpha_1)^2} \frac{1-\alpha_1^2}{1-3\alpha_1^2}, \tag{5.43}$$

provided $3\alpha_1^2 < 1$. These results imply that the kurtosis, κ, of y_t is

$$\kappa = \frac{E(y_t^4)}{[E(y_t^2)]^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2}, \tag{5.44}$$

which is always larger than 3 (unless $\alpha_1 = 0$), the kurtosis of the normal distribution. Thus, the marginal distribution of the returns, y_t, is leptokurtic, or has "fat tails."

In summary, an ARCH(1) process, y_t, as given by (5.36)-(5.37), or equivalently (5.38), is characterized by the following properties.

- If $0 \le \alpha_1 < 1$, the process y_t itself is white noise and its unconditional distribution is symmetrically distributed around zero; this distribution is leptokurtic.
- If, in addition, $3\alpha_1^2 < 1$, the square of the process, y_t^2, follows a causal AR(1) model with ACF given by $\rho_{y^2}(h) = \alpha_1^h \ge 0$, for all $h > 0$. If $3\alpha_1 \ge 1$, but $\alpha_1 < 1$, then y_t^2 is strictly stationary with infinite variance.

Estimation of the parameters α_0 and α_1 of the ARCH(1) model is typically accomplished by conditional MLE. The conditional likelihood of the data $y_2, ..., y_n$ given y_1, is given by

$$L(\alpha_0, \alpha_1 \mid y_1) = \prod_{t=2}^{n} f_{\alpha_0,\alpha_1}(y_t \mid y_{t-1}), \tag{5.45}$$

where the density $f_{\alpha_0,\alpha_1}(y_t \mid y_{t-1})$ is the normal density specified in (5.38). Hence, the criterion function to be minimized, $l(\alpha_0, \alpha_1) \propto -\ln L(\alpha_0, \alpha_1 \mid y_1)$ is given by

$$l(\alpha_0, \alpha_1) = \frac{1}{2}\sum_{t=2}^{n} \ln(\alpha_0 + \alpha_1 y_{t-1}^2) + \frac{1}{2}\sum_{t=2}^{n}\left(\frac{y_t^2}{\alpha_0 + \alpha_1 y_{t-1}^2}\right). \tag{5.46}$$

Regression with ARCH Errors.

Estimation is accomplis[illegible]ds, as described in §3.6. In this case, analytic expres[illegible]tor, $l^{(1)}(\alpha_0, \alpha_1)$, and Hessian matrix, $l^{(2)}(\alpha_0, \alpha_1)$, as described in Example 3.29, can be obtained by straightforward calculations. For example, the 2×1 gradient vector, $l^{(1)}(\alpha_0, \alpha_1)$, is given by

$$\begin{pmatrix} \partial l/\partial\alpha_0 \\ \partial l/\partial\alpha_1 \end{pmatrix} = \sum_{t=2}^{n} \begin{pmatrix} 1 \\ y_{t-1}^2 \end{pmatrix} \times \frac{\alpha_0 + \alpha_1 y_{t-1}^2 - y_t^2}{2\left(\alpha_0 + \alpha_1 y_{t-1}^2\right)^2}.$$

The calculation of the Hessian matrix is left as an exercise (Problem 5.9). The likelihood of the ARCH model tends to be flat unless n is very large. A discussion of this problem can be found in Shephard (1996).

It is also possible to combine a regression or an ARMA model for the mean with an ARCH model for the errors. For example, a regression with ARCH(1) errors model would have the observations x_t as linear function of p regressors, $\boldsymbol{z}_t = (z_{t1}, ..., z_{tp})'$, and ARCH(1) noise y_t, say,

$$x_t = \boldsymbol{\beta}'\boldsymbol{z}_t + y_t,$$

where y_t satisfies (5.36)-(5.37), but, in this case, is unobserved. Similarly, for example, an AR(1) model for data x_t exhibiting ARCH(1) errors would be

$$x_t = \phi_0 + \phi_1 x_{t-1} + y_t.$$

These types of models were explored by Weiss (1984).

Example 5.4 Analysis of U.S. GNP

In Example 3.38, we fit an MA(2) model and an AR(1) model to the U.S. GNP series and we concluded that the residuals from both fits appeared to behave like a white noise process. In Example 3.42 we concluded that the AR(1) is probably the better model in this case. It has been suggested that the U.S. GNP series has ARCH errors, and in this example, we will investigate this claim. If the GNP noise term is ARCH, the squares of

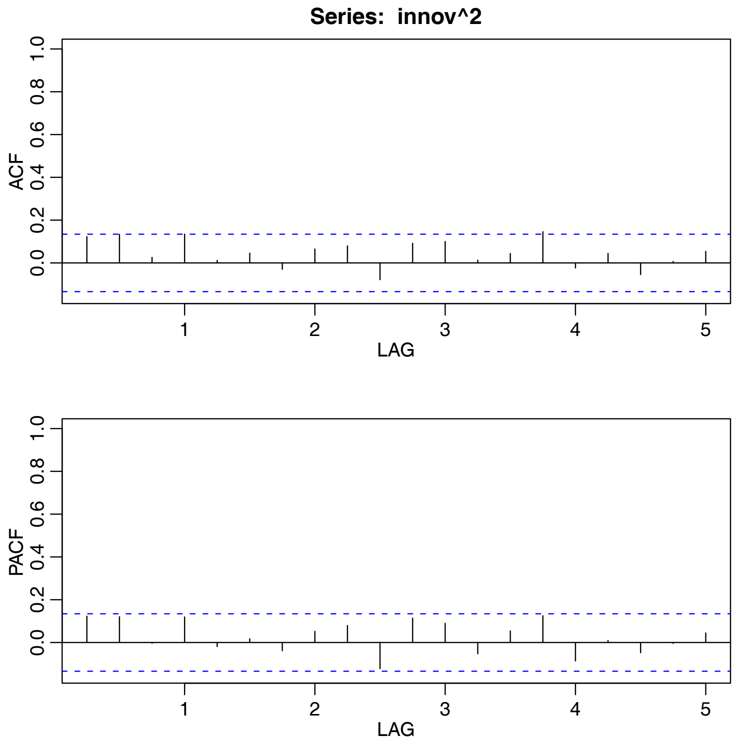


Fig. 5.5. ACF and PACF of the squares of the residuals from the AR(1) fit on U.S. GNP.

the residuals from the fit should behave like a non-Gaussian AR(1) process, as pointed out in (5.39). Figure 5.5 shows the ACF and PACF of the squared residuals it appears that there may be some dependence, albeit small, left in the residuals. The figure was generated in R as follows.

```
gnpgr = diff(log(gnp))  # get the growth rate
sarima(gnpgr, 1, 0, 0)  # fit an AR(1)
acf2(innov^2, 24)  # get (p)acf of the squared residuals
```

We used the R package `fGarch` to fit an AR(1)-ARCH(1) model to the U.S. GNP returns with the following results. A partial output is shown; we note that `garch(1,0)` specifies an `arch(1)` in the code below (details later).

```
library(fGarch)
summary(garchFit(~arma(1,0)+garch(1,0), gnpgr))
```

```
         Estimate  Std. Error  t value Pr(>|t|)
mu      5.278e-03   8.996e-04    5.867 4.44e-09
ar1     3.666e-01   7.514e-02    4.878 1.07e-06
omega   7.331e-05   9.011e-06    8.135 4.44e-16
alpha1  1.945e-01   9.554e-02    2.035   0.0418
```

```
Standardised Residuals Tests:  Statistic p-Value
 Jarque-Bera Test   R    Chi^2  9.118036  0.01047234
 Shapiro-Wilk Test  R    W      0.9842407 0.01433690
 Ljung-Box Test     R    Q(10)  9.874326  0.4515875
 Ljung-Box Test     R    Q(15)  17.55855  0.2865844
 Ljung-Box Test     R    Q(20)  23.41363  0.2689437
 Ljung-Box Test     R^2  Q(10)  19.2821   0.03682246
 Ljung-Box Test     R^2  Q(15)  33.23648  0.004352736
 Ljung-Box Test     R^2  Q(20)  37.74259  0.009518992
 LM Arch Test       R    TR^2   25.41625  0.01296901
```

In this example, we obtain $\widehat{\phi}_0 = .005$ (called `mu` in the output) and $\widehat{\phi}_1 = .367$ (called `ar1`) for the AR(1) parameter estimates; in Example 3.38 the values were .005 and .347, respectively. The ARCH(1) parameter estimates are $\widehat{\alpha}_0 = 0$ (called `omega`) for the constant and $\widehat{\alpha}_1 = .195$, which is significant with a p-value of about .04. There are a number of tests that are performed on the residuals [`R`] or the squared residuals [`R^2`]. For example, the Jarque–Bera statistic tests the residuals of the fit for normality based on the observed skewness and kurtosis, and it appears that the residuals have some non-normal skewness and kurtosis. The Shapiro–Wilk statistic tests the residuals of the fit for normality based on the empirical order statistics. The other tests, primarily based on the Q-statistic, are used on the residuals and their squares.

The ARCH(1) model can be extended to the general ARCH(m) model in an obvious way. That is, (5.36), $y_t = \sigma_t \epsilon_t$, is retained, but (5.37) is extended to

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \cdots + \alpha_m y_{t-m}^2. \tag{5.47}$$

Estimation for ARCH(m) also follows in an obvious way from the discussion of estimation for ARCH(1) models. That is, the conditional likelihood of the data $y_{m+1}, ..., y_n$ given $y_1, \ldots, y_m$, is given by

$$L(\boldsymbol{\alpha} \mid y_1, \ldots, y_m) = \prod_{t=m+1}^{n} f_{\boldsymbol{\alpha}}(y_t \mid y_{t-1}, \ldots, y_{t-m}), \tag{5.48}$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_m)$ and the conditional densities $f_{\boldsymbol{\alpha}}(\cdot|\cdot)$ in (5.48) are normal densities; that is, for $t > m$,

$$y_t \mid y_{t-1}, \ldots, y_{t-m} \sim \mathrm{N}(0, \alpha_0 + \alpha_1 y_{t-1}^2 + \cdots + \alpha_m y_{t-m}^2).$$

Another extension of ARCH is the generalized ARCH or GARCH model developed by Bollerslev (1986). For example, a GARCH(1,1) model retains (5.36), $y_t = \sigma_t \epsilon_t$, but extends (5.37) as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{5.49}$$

Under the condition that $\alpha_1 + \beta_1 < 1$, using similar manipulations as in (5.39), the GARCH(1,1) model, (5.36) and (5.49), admits a non-Gaussian ARMA(1,1) model for the squared process

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta_1)y_{t-1}^2 + v_t - \beta_1 v_{t-1}, \tag{5.50}$$

where v_t is as defined in (5.39). Representation (5.50) follows by writing (5.36) as

$$y_t^2 - \sigma_t^2 = \sigma_t^2(\epsilon_t^2 - 1)$$
$$\beta_1(y_{t-1}^2 - \sigma_{t-1}^2) = \beta_1\sigma_{t-1}^2(\epsilon_{t-1}^2 - 1),$$

subtracting the second equation from the first, and using the fact that, from (5.49), $\sigma_t^2 - \beta_1\sigma_{t-1}^2 = \alpha_0 + \alpha_1 y_{t-1}^2$, on the left-hand side of the result. The GARCH(m, r) model retains (5.36) and extends (5.49) to

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{m} \alpha_j y_{t-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2. \tag{5.51}$$

Conditional maximum likelihood estimation of the GARCH(m, r) model parameters is similar to the ARCH(m) case, wherein the conditional likelihood, (5.48), is the product of N$(0, \sigma_t^2)$ densities with σ_t^2 given by (5.51) and where the conditioning is on the first $\max(m, r)$ observations, with $\sigma_1^2 = \cdots = \sigma_r^2 = 0$. Once the parameter estimates are obtained, the model can be used to obtain one-step-ahead forecasts of the volatility, say $\widehat{\sigma}_{t+1}^2$, given by

$$\widehat{\sigma}_{t+1}^2 = \widehat{\alpha}_0 + \sum_{j=1}^{m} \widehat{\alpha}_j y_{t+1-j}^2 + \sum_{j=1}^{r} \widehat{\beta}_j \widehat{\sigma}_{t+1-j}^2. \tag{5.52}$$

We explore these concepts in the following example.

Example 5.5 GARCH Analysis of the NYSE Returns

As previously mentioned, the daily returns of the NYSE shown in Figure 1.4 exhibit classic GARCH features. We used the R `fGarch` package to fit a GARCH$(1, 1)$ model to the series with the following results:

```
library(fGarch)
summary(nyse.g <- garchFit(~garch(1,1), nyse))
```

```
         Estimate  Std. Error  t value Pr(>|t|)
mu      7.369e-04   1.786e-04    4.126 3.69e-05
omega   6.542e-06   1.455e-06    4.495 6.94e-06
alpha1  1.141e-01   1.604e-02    7.114 1.13e-12
beta1   8.061e-01   2.973e-02   27.112  < 2e-16

Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  3628.415  0
 Shapiro-Wilk Test  R    W      0.9515562 0
 Ljung-Box Test     R    Q(10)  29.69242  0.0009616813
 Ljung-Box Test     R    Q(15)  30.50938  0.01021164
 Ljung-Box Test     R    Q(20)  32.81143  0.03538324
```

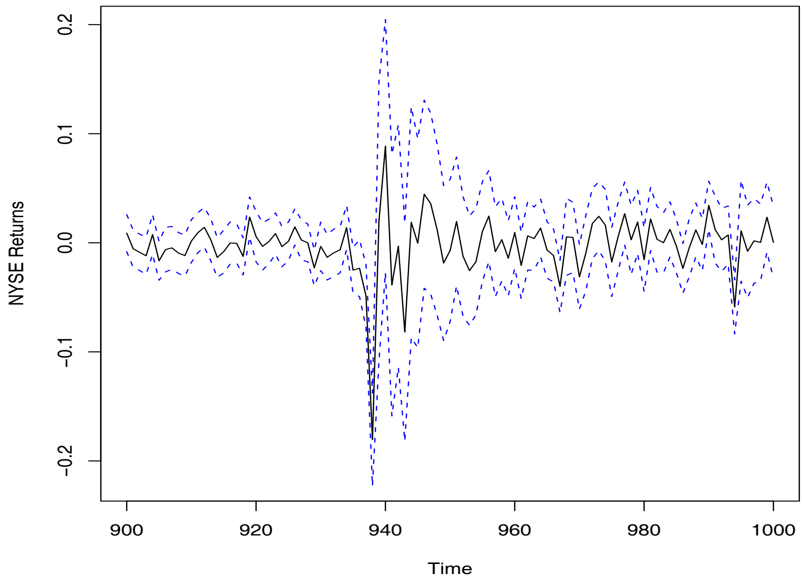


Fig. 5.6. GARCH predictions of the NYSE volatility, $\pm 2\widehat{\sigma}_t$, displayed as dashed lines.

```
Ljung-Box Test     R^2  Q(10)  3.510505  0.9667405
Ljung-Box Test     R^2  Q(15)  4.408852  0.9960585
Ljung-Box Test     R^2  Q(20)  6.68935   0.9975864
LM Arch Test       R    TR^2   3.967784  0.9840107
```

To explore the GARCH predictions of volatility, we calculated and plotted the 100 observations from the middle of the data (which includes the October 19, 1987 crash) along with the one-step-ahead predictions of the corresponding volatility, σ_t^2. The results are displayed as the data $\pm 2\widehat{\sigma}_t$ as a dashed line surrounding the data in Figure 5.6.

```
u = nyse.g@sigma.t
plot(window(nyse, start=900, end=1000), ylim=c(-.22,.2), ylab="NYSE
    Returns")
lines(window(nyse-2*u, start=900, end=1000), lty=2, col=4)
lines(window(nyse+2*u, start=900, end=1000), lty=2, col=4)
```

Some key points can be gleaned from the examples of this section. First, it is apparent that the conditional distribution of the returns is rarely normal. `fGarch` allows for various distributions to be fit to the data; see the help file for information. Some drawbacks of the GARCH model are: (i) the model assumes positive and negative returns have the same effect because volatility depends on squared returns; (ii) the model is restrictive because of the tight constraints on the model parameters (e.g., for an ARCH(1), $0 \le \alpha_1^2 < \frac{1}{3}$); (iii) the likelihood is flat unless n is very large; (iv) the model tends to overpredict volatility because it responds slowly to large isolated returns.

Various extensions to the original model have been proposed to overcome some of the shortcomings we have just mentioned. For example, we have already discussed the fact that the S-PLUS Garch module will fit some non-normal, albeit symmetric, distributions. For asymmetric return dynamics, one can use the EGARCH (exponential GARCH) model, which is a complex model that has different components for positive returns and for negative returns. In the case of persistence in volatility, the integrated GARCH (IGARCH) model may be used. Recall (5.50) where we showed the GARCH(1, 1) model can be written as

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta_1) y_{t-1}^2 + v_t - \beta_1 v_{t-1}$$

and y_t^2 is stationary if $\alpha_1 + \beta_1 < 1$. The IGARCH model sets $\alpha_1 + \beta_1 = 1$, in which case the IGARCH(1, 1) model is

$$y_t = \sigma_t \epsilon_t \quad \text{and} \quad \sigma_t^2 = \alpha_0 + (1 - \beta_1) y_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

There are many different extensions to the basic ARCH model that were developed to handle the various situations noticed in practice. Interested readers might find the general discussions in Engle et al. (1994) and Shephard (1996) worthwhile reading. Also, Gouriéroux (1997) gives a detailed presentation of ARCH and related models with financial applications and contains an extensive bibliography. Two excellent texts on financial time series analysis are Chan (2002) and Tsay (2002).

Finally, we briefly discuss stochastic volatility models; a detailed treatment of these models is given in Chapter 6. The volatility component, σ_t^2, in the GARCH model is conditionally nonstochastic. In the ARCH(1) model for example, any time the previous return is zero, i.e., $y_{t-1} = 0$, it must be the case that $\sigma_t^2 = \alpha_0$, and so on. This assumption seems a bit unrealistic. The stochastic volatility model adds a stochastic component to the volatility in the following way. In the GARCH model, a return, say y_t, is

$$y_t = \sigma_t \epsilon_t \quad \Rightarrow \quad \log y_t^2 = \log \sigma_t^2 + \log \epsilon_t^2. \tag{5.53}$$

Thus, the observations $\log y_t^2$ are generated by two components, the unobserved volatility $\log \sigma_t^2$ and the unobserved noise $\log \epsilon_t^2$. While, for example, the GARCH(1, 1) models volatility without error, $\sigma_{t+1}^2 = \alpha_0 + \alpha_1 r_t^2 + \beta_1 \sigma_t^2$, the basic stochastic volatility model assumes the latent variable is an autoregressive process,

$$\log \sigma_{t+1}^2 = \phi_0 + \phi_1 \log \sigma_t^2 + w_t \tag{5.54}$$

where $w_t \sim$ iid $\mathrm{N}(0, \sigma_w^2)$. The introduction of the noise term w_t makes the latent volatility process stochastic. Together (5.53) and (5.54) comprise the stochastic volatility model. Given n observations, the goals are to estimate the parameters ϕ_0, ϕ_1 and σ_w^2, and then predict future observations $\log y_{n+m}^2$. Details are provided in §6.10.