Lecture 36

GARCH Models Part 2

Arnab Hazra

Dow Jones Industrial Average (recap)

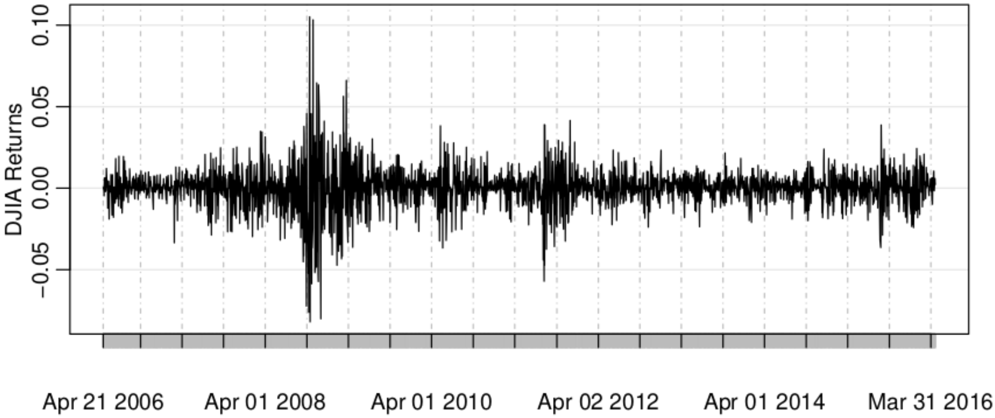

Fig. 1.4. *The daily returns of the Dow Jones Industrial Average (DJIA) from April 20, 2006 to April 20, 2016.*

Dow Jones Industrial Average (recap)

- It is easy to spot the financial crisis of 2008 in the figure.
- The data shown here are typical of return data ($R_t = \frac{X_t - X_{t-1}}{X_{t-1}}$).
- If the return represents a small (in magnitude) percentage change then $\nabla \log(X_t) \approx R_t$. Either value, $\nabla \log(X_t)$ or $\frac{X_t - X_{t-1}}{X_{t-1}}$ are called the return.
- The mean of the series appears to be stable with an average return of nearly zero, however, highly volatile periods tend to be clustered together.
- A problem in the analysis of these type of financial data is to forecast the volatility of future returns.
- Models such as ARCH and GARCH models and stochastic volatility models have been developed to handle these problems.

ARCH(1) model (recap)

- If R_t follows an AR(1) process, $\mathrm{Var}(R_t|R_{t-1}, R_{t-2}, \ldots) = \mathrm{Var}(R_t|R_{t-1}) = \sigma_W^2$.
- Typically, for financial series, R_t does not have a constant conditional variance, and highly volatile periods tend to be clustered together.
- The simplest ARCH model, the ARCH(1), models the return as

$$R_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2$$

  where ε_t's are IID standard Gaussian white noise.
- With ARMA models, we must impose some constraints ($\alpha_0, \alpha_1 \geq 0$) on the model parameters to obtain desirable properties.
- The conditional distribution $R_t|R_{t-1} \sim \mathrm{N}(0, \alpha_0 + \alpha_1 R_{t-1}^2)$.

AR(1)-type representation of ARCH(1) model (recap)

- We can write the ARCH(1) model as a non-Gaussian AR(1) model for R_t^2.
- First, we write as
$$R_t^2 = \sigma_t^2 \varepsilon_t^2, \quad \alpha_0 + \alpha_1 R_{t-1}^2 = \sigma_t^2.$$
- We subtract the two equations to obtain
$$R_t^2 - (\alpha_0 + \alpha_1 R_{t-1}^2) = \sigma_t^2 \varepsilon_t^2 - \sigma_t^2 \overset{\textit{Notation}}{=} V_t.$$
- Here $V_t = \sigma_t^2(\varepsilon_t^2 - 1)$. Because ε_t^2 is the square of a N(0, 1) random variable, $\varepsilon_t^2 - 1$ is a shifted (to have mean-zero), χ_1^2 random variable.
- Overall, $R_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + V_t$ where $V_t | R_{t-1} \sim (\alpha_0 + \alpha_1 R_{t-1}^2) \times (\chi_1^2 - 1)$.

Digression: Martingale (recap)

- A basic definition of a discrete-time martingale is a discrete-time stochastic process $\{X_1, X_2, X_3, \ldots\}$ that satisfies for any time T,

$$E(|X_t|) < \infty,$$

$$E(X_{T+1}|X_1, \ldots, X_T) = X_T.$$

- We can define $X_t^* = X_t - E(X_t|X_{t-1}, X_{t-2}, \ldots, X_1)$.

- Here, clearly,

$$E(X_t^*|X_{t-1}^*, X_{t-2}^*, \ldots, X_1^*) = E(X_t|X_{t-1}, X_{t-2}, \ldots, X_1) - E(X_t|X_{t-1}, X_{t-2}, \ldots, X_1) = 0.$$

- Here X_t^* is called martingale difference.

Properties of GARCH (recap)

- We define $\mathcal{R}_s = \{R_s, R_{s-1}, \ldots\}$.
- Because $E(R_t|\mathcal{R}_{t-1}) = 0$, the process R_t is said to be a martingale difference.
- Because R_t is a martingale difference, it is also an uncorrelated sequence.
- Therefore, $E(R_t^2)$ and $\mathrm{Var}(R_t^2)$ must be constant with respect to time t.

Properties of GARCH (contd., recap)

- $$E(R_t) = EE(R_t|\mathcal{R}_{t-1}) = EE(R_t|R_{t-1}) = 0$$

- $$\mathrm{Cov}(R_{t+h}, R_t) = E(R_t R_{t+h}) = EE(R_t R_{t+h}|R_{t+h-1}) = ER_t E(R_{t+h}|R_{t+h-1}) = 0$$

- $$E(R_t^2) = \mathrm{Var}(R_t) = \frac{\alpha_0}{1-\alpha_1}$$

- $$E(R_t^4) = \frac{3\alpha_0^2}{(1-\alpha_1)^2} \times \frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

Parameter estimation

- 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 $R_2, \ldots, R_T$ given R_1, is given by where the density $f_{\alpha_0,\alpha_1}(R_t|R_{t-1})$.
- Hence, the criterion function to be minimized, $\ell(\alpha_0, \alpha_1) = c - \log[L(\alpha_0, \alpha_1|R_1)]$ is given by

$$\ell(\alpha_0, \alpha_1) = 0.5 \sum_{t=2}^{T} \log(\alpha_0 + \alpha_1 R_{t-1}^2) + 0.5 \sum_{t=2}^{T} \frac{R_t^2}{\alpha_0 + \alpha_1 R_{t-1}^2}.$$

- Estimation is accomplished by numerical methods.

ARCH(p) Models

- The ARCH(1) model can be extended to the general ARCH(p) model by keeping $R_t = \sigma_t \varepsilon_t$ but extending $\sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2$ to $\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j R_{t-j}^2$.
- Estimation for ARCH(p) also follows in an obvious way from the discussion of estimation for ARCH(1) models.
- That is, the conditional likelihood of the data $R_{p+1}, \ldots, R_T$ given $R_1, \ldots, R_p$, is

$$L(\boldsymbol{\alpha}|R_1, \ldots, R_p) = \prod_{t=p+1}^{T} f_{\boldsymbol{\alpha}}(R_t|R_{t-1}, \ldots, R_{t-p}),$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)$ and, under the assumption of normality, the conditional densities $f_{\boldsymbol{\alpha}}(\cdot|\cdot)$ are, for $t > p$, given by

$$R_t|R_{t-1}, \ldots, R_{t-p} \sim \mathrm{N}(0, \alpha_0 + \sum_{j=1}^{p} \alpha_j R_{t-j}^2).$$

Other extensions

- 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, $\boldsymbol{Z}_t = [z_{t1}, \ldots, z_{tp}]'$, and ARCH(1) noise Y_t, say, $X_t = \boldsymbol{Z}_t'\beta + Y_t$,

- Similarly, for example, an AR(1) model for data X_t exhibiting ARCH(1) errors would be $X_t = \phi_0 + \phi_1 X_{t-1} + Y_t$.

GARCH(p, q) model

- Another extension of ARCH is the generalized ARCH or GARCH model.
- For example, a GARCH(1, 1) model retains $R_t = \sigma_t \varepsilon_t$, but

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

- The model admits a non-Gaussian ARMA(1, 1) model for the squared process

$$R_t^2 = \alpha_0 + (\alpha_1 + \beta_1) R_{t-1}^2 + V_t - \beta_1 V_{t-1}.$$

- It follows from the equations

$$R_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1), \quad \beta_1(R_{t-1}^2 - \sigma_{t-1}^2) = \beta_1 \sigma_{t-1}^2(\varepsilon_{t-1}^2 - 1)$$

- A natural extension to GARCH(p, q) model is by retaining $R_t = \sigma_t \varepsilon_t$, but

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_p R_{t-j}^2 + \sum_{k=1}^{q} \beta_k \sigma_{t-k}^2.$$

GARCH fitting to Dow Jones Industrial Average

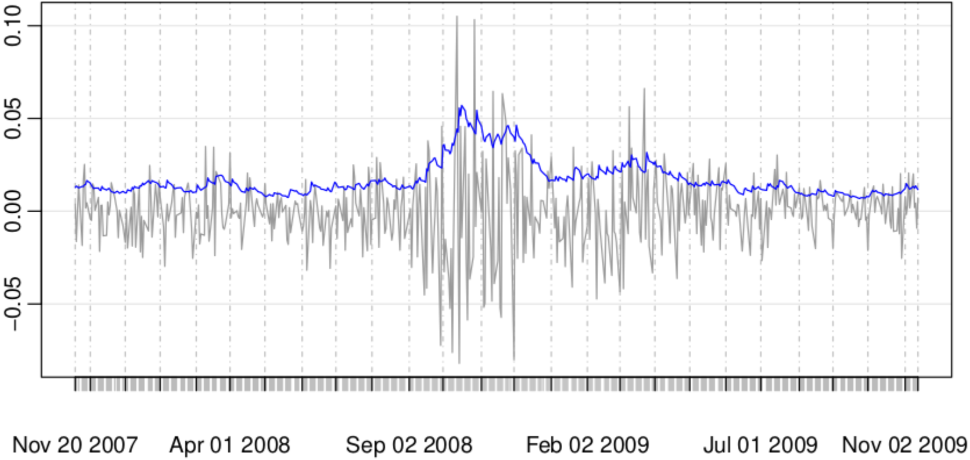

Fig. 5.6. *GARCH one-step-ahead predictions of the DJIA volatility, $\hat{\sigma}_t$, superimposed on part of the data including the financial crisis of 2008.*

Drawbacks of GARCH models

- The volatility component σ_t^2 in GARCH models are conditionally nonstochastic.
- This assumption seems a bit unrealistic and the stochastic volatility model adds a stochastic component to the volatility.
- For GARCH model, $R_t = \sigma_t \varepsilon_t$ and hence, $\log[R_t^2] = \log[\sigma_t^2] + \log[\varepsilon_t^2]$.
- Thus, $\log[R_t^2]$ are generated by two components, the unobserved volatility $\log[\sigma_t^2]$ and the unobserved noise $\log[\varepsilon_t^2]$.

Stochastic volatility models

- The basic stochastic volatility model assumes the logged latent variable is an autoregressive process

$$\log[\sigma^2_{t+1}] = \phi_0 + \phi_1 \log[\sigma^2_t] + W_t,$$

where $W_t \overset{IID}{\sim} \mathrm{N}(0, \sigma^2_W)$.

- The introduction of the noise term W_t makes the latent volatility process stochastic, which is a state-space model.

- Given T observations, the goals are to estimate the parameters ϕ_0, ϕ_1, and σ^2_W, and then predict future volatility using Kalman forecasting.

Thank you!