Advantages of GARCH model

Simplicity

Generates volatility clustering

Heavy tails (high kurtosis)

Weaknesses of GARCH model

Symmetric btw positive & negative prior returns

Restrictive

Provides no explanation

Not sufficiently adaptive in prediction

GARCH-M model:

$$r_t = \mu + c\sigma_t^2 + a_t,$$

$$a_t = \sigma_t \epsilon_t,\ \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where c is referred to as risk premium, which is expected to be positive.

EGARCH model:

Asymmetry in responses to + and - returns:

$$g(\epsilon_t) = \theta\epsilon_t + [|\epsilon_t| - E(|\epsilon_t|)],$$

with $E[g(\epsilon_t)] = 0$. To see asymmetry of $g(\epsilon_t)$, rewrite it as

$$g(\epsilon_t) = \begin{cases} (\theta + 1)\epsilon_t - E(|\epsilon_t|) \; if \; \epsilon_t \geq 0, \\ (\theta - 1)\epsilon_t - E(|\epsilon_t|) \; if \; \epsilon_t < 0. \end{cases}$$

An EGARCH(m, s) model:

$$a_t = \sigma_t\epsilon_t, \;\; \ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{m} \alpha_i g(\epsilon_{t-i}) + \sum_{i=1}^{s} \beta_i \ln(\sigma_{t-i}^2).$$

Some features of EGARCH models:

uses log trans. to relax the positiveness constraint

asymmetric responses.

Consider an EGARCH(1,1) model

$$a_t = \sigma_t \epsilon_t, \quad (1 - \beta B)\ln(\sigma_t^2) = \alpha_0 + \alpha g(\epsilon_{t-1}),$$

Under normality, $E(|\epsilon_t|) = \sqrt{2/\pi}$ and the model becomes

$$(1 - \beta B)\ln(\sigma_t^2) = \begin{cases} \alpha_* + \alpha(\theta + 1)\epsilon_{t-1} \ if \ \epsilon_{t-1} \geq 0, \\ \alpha_* + \alpha(\theta - 1)\epsilon_{t-1} \ if \ \epsilon_{t-1} < 0, \end{cases}$$

where $\alpha_* = \alpha_0 - \alpha\sqrt{2/\pi}$.

This is a nonlinear fun. similar to that of the threshold AR model of Tong (1978, 1990). Specifically, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\beta} \exp(\alpha_*) \begin{cases} \exp[(\theta_1 + \alpha)\epsilon_{t-1}] \ if \ a_{t-1} \geq 0, \\ \exp[(\theta_1 - \alpha)\epsilon_{t-1}] \ if \ a_{t-1} < 0, \end{cases}$$

where $\theta_1 = \theta\alpha$.

The coefficients $(\theta_1 + \alpha)$ and $(\theta_1 - \alpha)$ show the asymmetry in response to positive and negative a_{t-1}. The model is, therefore, nonlinear if $\theta_1 \neq 0$. Thus, θ_1 (or θ) is referred to as the leverage parameter. It shows the effect of the sign of a_{t-1} whereas α denotes the magnitude effect. See Nelson (1991) for an example of EGARCH model.

ARMA-GARCH model:

$$\begin{aligned} r_t &= \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j}, \\ a_t &= \eta_t \sigma_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2. \end{aligned}$$

r_t is called ARMA$(p,q)-$GARCH(r,s) model.

ARIMA-GARCH model:

$$\begin{aligned} \phi_p(B)(1-B)\log P_t &= \theta_q(B) a_t, \\ a_t &= \eta_t \sigma_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2. \end{aligned}$$

$\log P_t$ is called ARIMA$(p,1,q)-$GARCH(m,s) model.

Estimation: Maximum Likelihood Estimation

We consider the case with $p = 1$ and $q = 0$ and $s = m = 1$. Assume that random sample $\{r_1, \cdots, r_n\}$ is from the AR(1)-GARCH(1,1) model:

$$\begin{aligned} &r_t = \phi_{10} r_{t-1} + a_t, \\ &a_t = \sigma_t \epsilon_t, \\ &\sigma_t^2 = \alpha_{00} + \alpha_{10} a_{t-1}^2 + \beta_{10} \sigma_{t-1}^2, \end{aligned}$$

where $\lambda_0 = (\phi_{10}, \alpha_{00}, \alpha_{10}, \beta_{10})'$ is called the true parameters.

Denote $\tilde{Z}_t = (r_t, r_{t-1}, \cdots)$. Given $\tilde{Z}_{t-1}$, the conditional density function of r_t is

$$f(r_t|\tilde{Z}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(r_t - \phi_{10} r_{t-1})^2}{2\sigma_t^2}\right).$$

Given $\tilde{Z}_0$, the conditional joint density function of $(r_n, r_{n-1}, \cdots, r_1)$:

$$f(r_t, \cdots, r_1|\tilde{Z}_0) = \prod_{t=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(r_t - \phi_{10} r_{t-1})^2}{2\sigma_t^2}\right) \right\},$$

where

$$\sigma_t^2 = \alpha_{00} + \alpha_{10}(r_{t-1} - \phi_{10} r_{t-2})^2 + \beta_{10}\sigma_{t-1}^2.$$

Replaced λ_0 by its unknown parameter $\lambda = (\phi, \alpha_0, \alpha_1, \beta_1)'$, we get

$$\begin{aligned} a_t(\phi) &= r_t - \phi r_{t-1}, \\ \sigma_t^2(\lambda) &= \alpha_0 + \alpha_1(r_{t-1} - \phi r_{t-2})^2 + \beta_1 \sigma_{t-1}^2(\lambda). \end{aligned}$$

The conditional likelihood function of $(r_n, r_{n-1}, \cdots, r_1)$:

$$f(r_t, \cdots, r_1|\tilde{Z}_0) = \prod_{t=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma_t^2(\lambda)}} \exp\left(-\frac{a_t^2(\phi)}{2\sigma_t^2(\lambda)}\right) \right\}.$$

Log -conditional likelihood function of $(r_n, r_{n-1}, \cdots, r_1)$:

$$\begin{aligned} L(\lambda) &\equiv \ln f(r_t, \cdots, r_1 | \tilde{Z}_0) \\ &= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left\{\ln \sigma_t^2(\lambda) + \frac{a_t^2(\phi)}{\sigma_t^2(\lambda)}\right\}. \end{aligned}$$

The MLE of λ is the maximizer of $L(\lambda)$, denote by $\widehat{\lambda}$. If $E\varepsilon_t^4 < \infty$, then

$$\widehat{\lambda} \longrightarrow \lambda_0 \text{ as } n \to \infty,$$
$$\sqrt{n}(\hat{\lambda} - \lambda_0) \sim N(0, \widehat{\Omega}),$$

where

$$\widehat{\Omega} = E\left[\frac{\partial^2 L(\hat{\lambda})}{\partial\lambda\partial\lambda'}\right]^{-1} E\left[\frac{\partial L(\hat{\lambda})}{\partial\lambda}\frac{\partial L(\hat{\lambda})}{\partial\lambda'}\right] E\left[\frac{\partial^2 L(\hat{\lambda})}{\partial\lambda\partial\lambda'}\right]^{-1}.$$

Diagnostic Checking and Model selection

The formal method is not provided in SAS. AIC is a main tool for model selection.

We can use Ljung-Box test for squared standardized residuals:

$$\widehat{\epsilon}_t = \frac{a_t(\widehat{\phi})}{\sigma_t(\widehat{\lambda})} \; and \; \widehat{\epsilon}_t^2 = \frac{a_t^2(\widehat{\phi})}{\sigma_t^2(\widehat{\lambda})}.$$

Forecasting

The forecast value $\widehat{r}_t(l)$ of r_{t+l} is calculated by

$$\widehat{r}_t(l) = E(r_{t+l}|r_t, r_{t-1}, \cdots).$$

The formulas is the same as the one in the ARMA model with a constant variance.

The one-step forecast interval for the ARIMA-GARCH model:

$$\left[\widehat{r}_t(1) - \mathcal{N}_{\frac{\alpha}{2}}\sigma_t(\widehat{\lambda}),\; \widehat{r}_t(1) + \mathcal{N}_{\frac{\alpha}{2}}\sigma_t(\widehat{\lambda})\right]$$

where $\mathcal{N}_{\frac{\alpha}{2}}$ is the $\alpha/2-$ quantile of $\mathcal{N}(0,1)$.
SAS only provides the forecasting value and forecast interval for AR-GARCH model.
You have to write your own programs in practice.

Other GARCH-type Models

The Threshold GARCH (TGARCH) or GJR Model
A TGARCH(s,m) or GJR(s,m) model is defined as

$$r_t = \mu_t + a_t, \;\; a_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2,$$

where N_{t-i} is an indicator variable such that

$$N_{t-i} = 1 \; if \; a_{t-i} < 0, \;\; and \;\; = 0 \;\; otherwise$$

One expects γ_i to be positive so that prior negative returns have higher impact on the volatility.

TGARCH(1,1) model:

$$\begin{aligned} \sigma_t^2 &= (\alpha_{10} + \alpha_{11} a_{t-1}^2 + \beta_{11} \sigma_{t-1}^2) I\{a_{t-1} \le 0\} \\ &\quad + (\alpha_{20} + \alpha_{21} a_{t-1}^2 + \beta_{21} \sigma_{t-1}^2) I\{a_{t-1} > 0\}. \end{aligned}$$

The CHARMA Model

Make use of interaction btw past shocks.

A CHARMA model is defined as

$r_t = \mu_t + a_t,\ a_t = \delta_{1t}a_{t-1} + \delta_{2t}a_{t-2} + \cdots + \delta_{mt}a_{t-m} + \eta_t,$

where $\{\eta_t\}$ is iid $N(0, \sigma_\eta^2)$, $\{\delta_t\} = \{(\delta_{1t}, \cdots, \delta_{mt})'\}$ is a sequence of iid random vectors $D(0, \Omega)$, $\{\delta_t\}$ and $\{\eta_t\}$ are independent. The model can be written as

$$a_t = (a_{t-1}, \cdots, a_{t-m})\delta_t + \eta_t,$$

with conditional variance

$$\sigma_t^2 = \sigma_\eta^2 + (a_{t-1}, \cdots, a_{t-m})\Omega(a_{t-1}, \cdots, a_{t-m})',$$

where $\Omega = E(\delta_t \delta_t')$.

RCA Model

A time series r_t is a RCA(p) model if

$$r_t = \phi_0 + \sum_{i=1}^{p} (\phi_i + \delta_{it}) r_{t-i} + a_t,$$

where $\{a_t\}$ is i.i.d. with mean 0 and variance σ_a^2, $\{\delta_t\} = \{(\delta_{1t}, \cdots, \delta_{pt})'\}$ is a sequence of iid random vectors $D(0, \Omega_\delta)$, $\{\delta_t\}$ and $\{a_t\}$ are independent. For this model, we have

$$\mu_t = E(r_t | F_{t-1}) = \sum_{i=1}^{p} \phi_i r_{t-i},$$

$$\sigma_t^2 = \sigma_a^2 + (r_{t-1}, \cdots, r_{t-p}) \Omega_\delta (r_{t-1}, \cdots, r_{t-p})',$$

where $\Omega_\delta = E(\delta_t \delta_t')$.

Stochastic volatility model

A (simple) SV model is

$$a_t = \sigma_t \epsilon_t, \;\; (1 - \alpha_1 B - \cdots - \alpha_m B^m) \ln(\sigma_t^2) = \alpha_0 + v_t,$$

where ϵ_t's are iid $N(0,1)$, v_t's are iid $N(0, \sigma_v^2)$, $\{\epsilon_t\}$ and $\{v_t\}$ are independent.

Alternative Approaches to Volatility

Two alternative methods:

Use of high-frequency financial data

Use of daily open, high, low and closing prices

Use of High-Frequency Data

Purpose: monthly volatility

Data: Daily returns

Let r_t^m be the $t-$th month log return. Let $\{r_{t,i}\}_{i=1}^n$ be the daily log returns within the $t-$th month.

Using properties of log returns, we have

$$r_t^m = \sum_{i=1}^n r_{t,i}.$$

Assuming that the conditional variance and covariance exist, we have

$$\begin{aligned} Var(r_t^m|F_{t-1}) &= \sum_{i=1}^n Var(r_{t,i}|F_{t-1}) \\ &\quad +2\sum_{i<j} Cov[(r_{t,i}, r_{t,j})|F_{t-1}], \end{aligned}$$

where F_{t-1} = the information available at month $t-1$ (inclusive). Further simplification possible under additional assumptions.

If $\{r_{t,i}\}$ is a white noise series, then

$$Var(r_t^m|F_{t-1}) = nVar(r_{t,1}),$$

where $Var(r_{t,1})$ can be estimated from the daily returns $\{r_{t,i}\}_{i=1}^n$ by

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2}{n-1},$$

where $\bar{r}_t$ is the sample mean of the daily log returns in month t (i.e., $\bar{r}_t = \sum_{i=1}^n r_{t,i}/n$).

The estimated monthly volatility is then

$$\hat{\sigma}_m^2 = \frac{n}{n-1} \sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2.$$

If $\{r_{t,i}\}$ follows an MA(1) model, then

$$Var(r_t^m|F_{t-1}) = nVar(r_{t,1}) + 2(n-1)Cov(r_{t,1}, r_{t,2}),$$

which can be estimated by

$$\hat{\sigma}_m^2 = \frac{n}{n-1} \sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2 + 2 \sum_{i=1}^{n-1} (r_{t,i} - \bar{r}_t)(r_{t,i+1} - \bar{r}_t).$$

Advantage: Simple

Weaknesses:

Model for daily returns $\{r_{t,i}\}$ is unknown.

Typically, 21 trading days in a month, resulting in a small sample size.

Realized integrated volatility

If the sample mean $\bar{r}_t$ is zero, then $\sigma_m^2 \approx \sum_{i=1} r_{t,i}^2$.

Use cumulative sum of squares of daily log returns within a month as an estimate of monthly volatility.

Apply the idea to *intrdaily* log returns and obtain realized integrated volatility.

Assume daily log return $r_t = \sum_{i=1}^n r_{t,i}$. The quantity

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2,$$

is called the realized volatility of r_t.

Advantages: simplicity and using intraday information

Weaknesses:

Effects of market microstructure (noises)

Overlook overnight return

Use of Daily Open, High, Low and Close Prices: See text book.