Chapter 7 Conditional Heteroscedastic Models (Part B)

GARCH Model

(a) A general ARMA(p,q) + GARCH(m, s) model is

$$r_t = \mu_t + u_t, \quad u_t = \sigma_t \epsilon_t;$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

where μ_t is ARMA(p,q) model

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i u_{t-i};$$

where ϵ_t are IID with,

$$E\epsilon_t = 0, \quad E\epsilon_t^2 = 1,$$

$\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and

$$\sum_{i=1}^{max(m;s)} (\alpha_i + \beta_i) < 1.$$

Here $\alpha_i = 0$ if $i > m$ and $\beta_i = 0$ if $i > s$.

(b) Let $\eta_t = u_t^2 - \sigma_t^2$, then $\{\eta_t\}$ is uncorrelated series. The GARCH model becomes

$$u_t^2 = \omega + \sum_{i=1}^{\max(m;s)} (\alpha_i + \beta_i) u_{t-i}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}$$

(c) This is an ARMA form for the squared series u_t^2. It can be used to understand properties of GARCH models, e.g. moment equations, forecasting, etc.

(d) If we further assume μ_t is ARMA(p,q) model, then together the model is called ARMA(p,q)+GARCH(m,s) model

Probability properties

Focus on a GARCH(1,1) model

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2;$$

(a) Weakly stationary: $0 \leq \alpha_1, \quad \beta_1 \leq 1, \quad (\alpha_1 + \beta_1) < 1.$

(b) Volatility clustering

(c) Unconditional variance

$$E(u_t^2) = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

(d) Heavy tails: if $\epsilon_t \sim N(0,1)$, and $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then

$$\frac{E(u_t^4)}{[E(u_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3$$

Forecast of ARMA(p,q) + GARCH(m,s)

(a) Usually, for $r_1, ..., r_T$, the residuals $u_1, ..., u_T$ and fitted σ_t^2 are calculated.

(b) The forecast include forecast of expected return (conditional mean) and of volatility, i.e.

$$E(r_{t+h}|r_t, r_{t-1}, ...)$$

and

$$\sigma_{t+h|t}$$

Note that for conditional mean, we can do it exactly as before by ignoring the volatility model, i.e. ARCH or GARCH.

Next we only focus on the prediction of volatility

(c) For 1-step ahead forecast,

$$\hat{\sigma}^2_{T+1} = \omega + \alpha_1 u^2_T + ... + \alpha_m u^2_{T+1-m} + \beta_1 \sigma^2_T + ... + \beta_s \sigma^2_{T+1-s}$$

(d) For 2-step ahead forecasts

$$\begin{aligned}
\hat{\sigma}^2_{T+2} &= \omega + \alpha_1 u^2_{T+1} + \alpha_2 u^2_T + ... + \alpha_m u^2_{T+2-m} + \beta_1 \sigma^2_{T+1} + \beta_2 \sigma^2_T + ... + \beta_s \sigma^2_{T+2-s} \\
&\approx \omega + \alpha_1 \hat{\sigma}^2_{T+1} + \alpha_2 u^2_T + ... + \alpha_m u^2_{T+2-m} + \beta_1 \hat{\sigma}^2_{T+1} + \beta_2 \sigma^2_T + ... + \beta_s \sigma^2_{T+2-s}
\end{aligned}$$

(e) For 3-step ahead,

$$\begin{aligned}
\hat{\sigma}^2_{T+3} &= \omega + \alpha_1 u^2_{T+2} + \alpha_2 u^2_{T+1} + ... + \alpha_m u^2_{T+3-m} + \beta_1 \sigma^2_{T+2} + \beta_2 \sigma^2_{T+1} + ... + \beta_s \sigma^2_{T+3-s} \\
&\approx \omega + \alpha_1 \hat{\sigma}^2_{T+2} + \alpha_2 \hat{\sigma}^2_{T+1} + ... + \alpha_m u^2_{T+3-m} + \beta_1 \hat{\sigma}^2_{T+2} + \beta_2 \hat{\sigma}^2_{T+1} + ... + \beta_s \sigma^2_{T+3-s}
\end{aligned}$$

(f) R package

```
[object] = garchFit(~arma(p,q) + garch(m,s), data = ??, cond.dist = "std", ...)

coef = [object]@fit$coef;

# get residuals

ut = [object]@residuals        # this u_t is useful for prediction of returns, and also the volatility

ut = [object]@sigma.t     # fitted volatility

predict([object], n.ahead=??, conf=0.95, plot=TRUE)

                                                    # conf= confidence

plot([object])
```

Example 0.1 *For the same time series above, if we fit it by a ARMA(1,1) + GARCH(2,1) with normal and t(7) for the innovation, the AIC are respectively 3.105409 and 3.085261. See code06A3.R.*

As comparing AIC and BIC, GARCH(2,1) is preferred. The fitted model is

$$
\begin{aligned}
r_t &= 3.659e-03 + 9.457e-01 * (r_{t-1} - 3.659e-03) - 9.612e-01 * u_{t-1} + u_t \\
u_t &= \sigma_t \epsilon_t \\
\sigma_t^2 &= 2.573e-02 + 1.000e-08 * u_{t-1}^2 + 1.572e-01 u_{t-2}^2 + 8.367e-01 \sigma_{t-1}^2
\end{aligned}
$$

ϵ_t *follows t(7)*

> predict(mGARCHtA, 5)

meanForecast meanError standardDeviation

1 0.09513836 0.8225427 0.8225427

2 0.09364919 0.8968292 0.8967397

3 0.09224061 0.8976294 0.8974431

4 0.09090824 0.9094484 0.9091786

5 0.08964797 0.9193364 0.9189909

The GARCH-M model (GARCH-in-Mean)

$$r_t = \mu_t + c\sigma_t + u_t, \quad u_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

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

APARCH Models (Asymmetric Power ARCH)

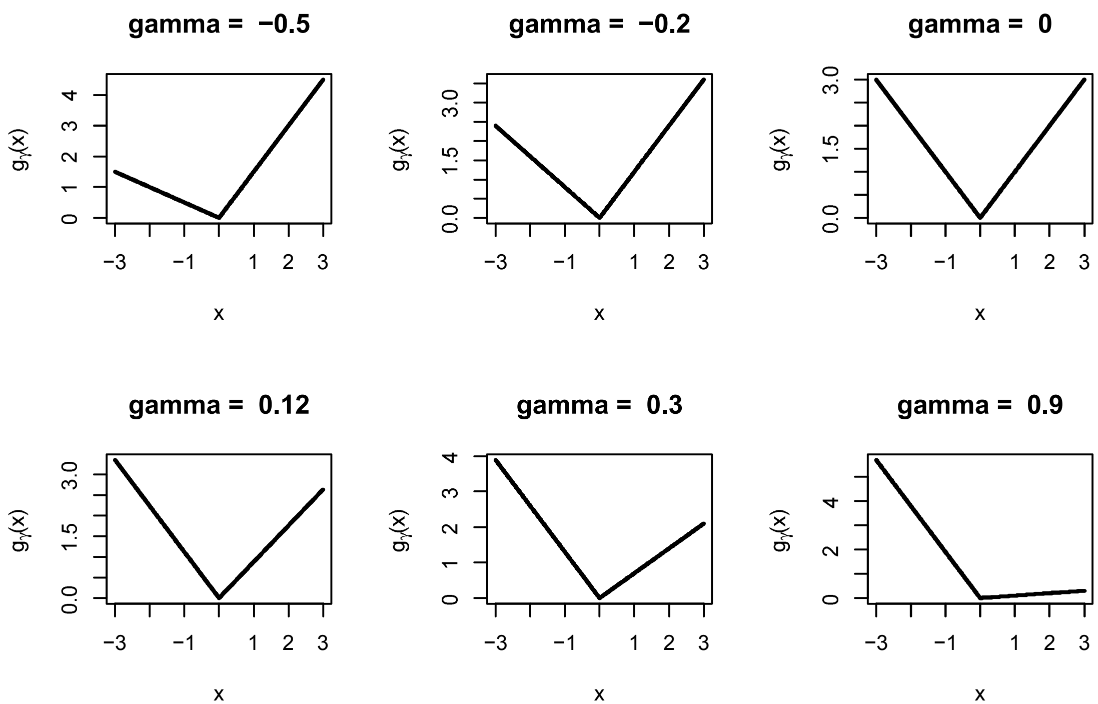
(a) In some financial time series, large negative returns appear to increase volatility more than do positive returns of the same magnitude. This is called the leverage effect.

(b) The APARCH(p; q) model for the conditional standard deviation is

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i (|u_{t-1}| - \gamma_i u_{t-1})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta;$$

where $\delta > 0$ and $-1 < \gamma_j < 1, j = 1, ...p$. Note that $\delta = 2$ and $\gamma_1 = ... = \gamma_p = 0$ give a standard GARCH model.

(c) $g_\gamma(x) = |x| - \gamma x$ looks as follows: positive past values of at increase volatility more than negative past values of the same magnitude.



Example 0.2 *For the above data, if we fit an APARCH(1,1), the estimated model for $y_t = 100 * r_t$ is*

$$\begin{aligned} y_t &= 0.003659 + u_t, \quad u_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 0.018004 + 0.050097(|u_{t-1}| - 0.999120 * u_{t-1})^2 + 0.895382\sigma_{t-1}^2 \end{aligned}$$

where ϵ_t is t(7). By comparing AIC, it seems APARCH(1,1) is even better.

See code06A4.R.

Note that γ is usually significantly positive, meaning there is strong leverage effect. As a consequence, **it can explain another stylized fact: Bull market is less volatile than the bear market**

Value at Risk (VaR) Conditional on the past

Given the past information $r_1, ..., r_{t-1}$, for a small $\alpha > 0$ we define the value at risk as

$$VaR_t(\alpha) = -\max\{v : P(r_t \leq v | r_{t-1}, r_{t-2}, ...) \leq \alpha\}$$

If we assume

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

and ϵ_t is independent of σ_t, then

$$VaR_t(\alpha) = -\mu_t - \sigma_t z_\alpha = -\mu_t + \sigma_t VaR_\epsilon(\alpha)$$

where z_α is the αth quantile of ϵ_t

- if $\epsilon_t \sim N(0,1)$, then

$$z_\alpha = \Phi^{-1}(\alpha)$$

  the α quantile of standard normal.

- If t-distribution is considered for the residuals with degree of freedom ν, because ϵ_t is standardized, i.e.

$$\epsilon_t = \epsilon_t'/\sigma$$

where ϵ_t' follows $t(\nu)$, and

$$\sigma = \sqrt{\frac{\nu}{\nu - 2}}$$

Thus

$$z_\alpha = t_\nu^{-1}(\alpha)/\sqrt{\frac{\nu}{\nu - 2}}$$

where $t_\nu^{-1}(\alpha)$ is the α quantile of $t(\nu)$.

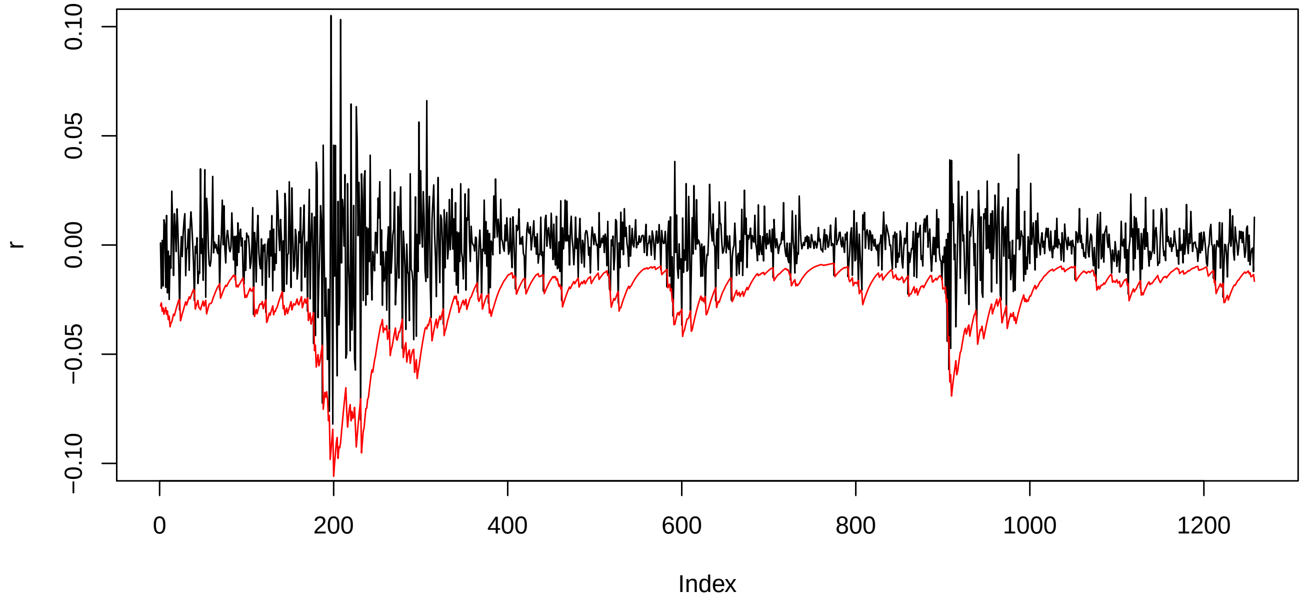
- As comparison, we studied the "unconditional" VaR of r_t

$$VaR(\alpha) = -\max\{v : P(r_t \leq v) \leq \alpha\}$$

If $(r_t - m)/s \sim t(\nu)$, then

$$VaR(\alpha) = -m - s * t_\nu^{-1}(\alpha)$$

Example 0.3 *For the above data, if we fit the model by GARCH(1,1) with t(7) as innovation, the* $VaR_t(0.05)$ *is shown in red below.*



see code06A5.R.

Validation of VaR model using garchFit

The VaR$_t$ with level α means

$$Pr(u_t > -VaR_t(\alpha)) = 100(1-\alpha)\% \tag{1}$$

To validate the calculation of $VaR_t(\alpha)$, the estimated probability

$$\hat{p} = \frac{\#\{u_t > -VaR_t(\alpha), t = 1, ..., T\}}{T}$$

should be approximately $100(1-\alpha)\%$.

In practice, T is usually 1 to 5 years, and $\alpha = 0.01$ or 0.001.

(a) if $\hat{p} > 100(1-\alpha)\%$, then the model is too conservative

(b) if $\hat{p} < 100(1-\alpha)\%$, then the model is too aggressive

It may be preferable for models to be a bit conservative rather than too aggressive

Example 0.4 *For the same data above, if we using aparch(1,1) with t(7) innovation and* $\alpha = 0.01$, *the proportion for the residuals above the VaR is*

$$\hat{p} = 0.9881$$

If we use normal innovation

$$\hat{p} = 0.975$$

therefore, aparch(1,1) with t(7) is more preferred.

see code06A6.R.

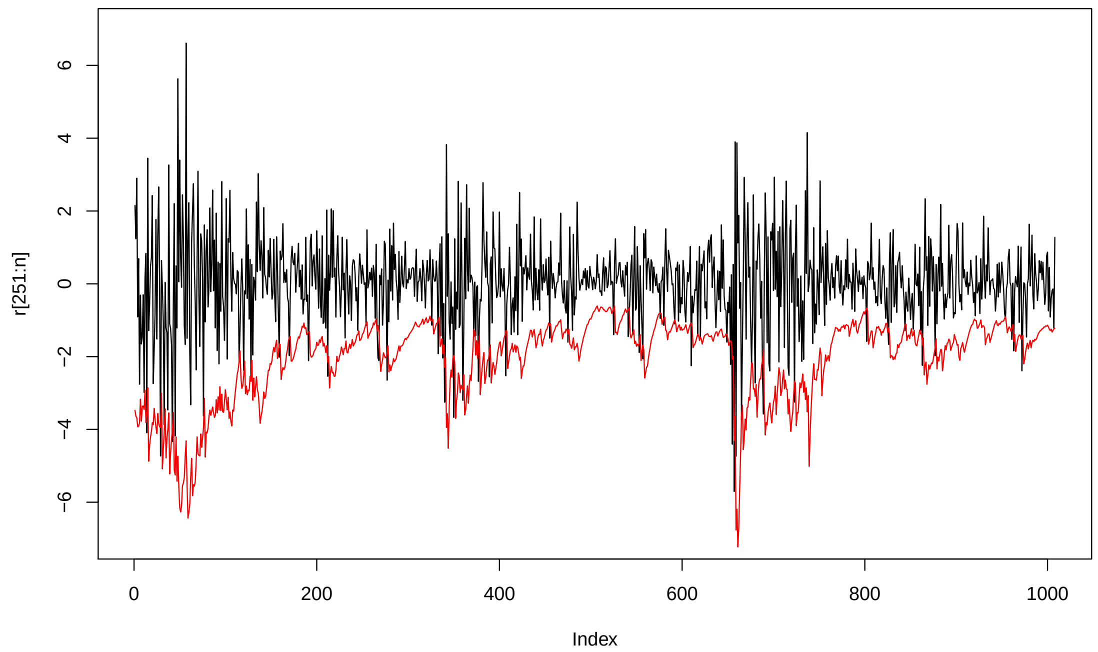
Prediction of VaR

We first predict σ_{T+h} then, together with the distribution, calculate the VaR at day $T+h$ at level α.

$$-\mu_{T+h|T} - \sigma_{T+h|T} Q(\alpha)$$

where $Q(\alpha)$ is the quantile for the standardized residuals, and $\mu_{T+h|T}, \sigma_{T+h|T}$ are the prediction of conditional mean and conditional standard deviation of the time series.

Example 0.5 (Continued) Based one each window of $T = 250$ days, we can estimate the model and make prediction of the conditional variance for the next day. Below is a plot of the predict VaR at level 0.05 and the observed returns based on GARCH(2,1) with t(7) for the standardized residuals. The empirical value of the level is 0.05357143 (a little too aggressive)



see code06Apred.R.