Modeling the conditional variance

- Time series with stochastic volatility

unpredictable/depends

One period return

$$y_t = p_t - p_{t-1}$$

where $p_t = \log P_t$, P_t price at time t

decomposed in

conditional mean — conditional volatility

$$y_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} \cdot z_t$$

$\mu_{t|t-1}$: assume $= 0$

where $z_t \sim NID(0,1)$ and given the filtration $\mathcal{F}$

$$\mu_{t|t-1} = E(y_t \mid \mathcal{F}_{t-1})$$ → best estimate of the mean based on past information

$$\sigma^2_{t|t-1} = Var(y_t \mid \mathcal{F}_{t-1}) = Var(\varepsilon_t \mid \mathcal{F}_{t-1})$$ → volatility associated with the observation at time t

Remark

When $\mu_{t|t-1} = 0$ then

$$y_t = \varepsilon_t = \overbrace{\sigma_{t|t-1}}^{\text{scaling factor}} \cdot z_t$$

z_t → iid Gaussian r.v.

Now

$$E(y_t \mid \mathcal{F}_{t-1}) = E(\sigma_{t|t-1} \cdot z_t \mid \mathcal{F}_{t-1}) = \sigma_{t|t-1} \cdot E(z_t) = 0$$

is a martingale difference sequence so its uncond. exp. is also zero

sequence of r.v. where each term represents the diff between consecutive obs. in a mart. process

$$E(y_t) = 0$$

And the conditional variance of y_t given $\mathcal{F}_{t-1}$

$$E(y^2 \mid \mathcal{F}_{t-1}) = E(\sigma^2_{t|t-1} \cdot z_t^2 \mid \mathcal{F}_{t-1}) = \underbrace{\sigma^2_{t|t-1}}_{\text{since it's indep from the filtration}} \cdot \underbrace{E(z_t^2 \mid \mathcal{F}_{t-1})}_{\text{variance of a } N(0,1)} = \sigma^2_{t|t-1} \cdot E(z_t^2)$$

$$= \sigma^2_{\underbrace{t}_{\text{time varying}} | \underbrace{t-1}_{\mathcal{F}_{t-1}\text{-measurable}}}$$

Model the cond. variance

↓

we search for a function of $y_t, y_{t-1}, \dots$ and the initial condition $\sigma_{1|0}$ that describes the dynamics of $\sigma^2_{t|t-1}$ and allows for the stylised facts that we observe in real data.

$$\underline{Var(y_t \mid \mathcal{Y}_{t-1})} = \underline{\sigma^2_{t|t-1}}$$

The ARCH model relates $\underline{\sigma^2_{t|t-1}}$ to the past squared observations $\underline{y_t^2}$

- Stylised facts characterising returns
  - returns are serially uncorrelated or weakly dependent over time
  - returns are not independent
  - Volatility changes slowly over time
  - Financial t.s. exhibit volatility clustering
  - Heavy tail modeling distribution

- $\underbrace{A}_{\text{autoregressive}}\underbrace{RC}_{\text{conditional}}\underbrace{H}_{\text{Heteroscedastic}}$ models

$$y_t = \sigma_{t|t-1} \cdot z_t$$

and

$$\sigma^2_{t|t-1} = \overset{\text{constant}}{\underbrace{\omega}_{\text{the unconditional variance}}} + \overbrace{\alpha_1 \cdot y^2_{t-1} + \dots + \alpha_p \cdot y^2_{t-p}}^{\text{linear combination of past obs.}}$$

with

$$\omega > 0, \quad \alpha_i \geq 0, \quad \text{for } i = 1, \dots, p$$

and

$$z_t \sim N(0, 1)$$

• ARCH (1) process

Let $z_t \sim N(0,1)$

$$y_t = \sigma_{t|t-1} \cdot z_t$$

and with

$$\sigma^2_{t|t-1} = \omega + \alpha y^2_{t-1} \qquad \omega > 0, \ \alpha \geq 0$$

• Remarks

- $y_t \mid \mathcal{Y}_{t-1} \sim N(0, \sigma_{t|t-1})$

- The cond. var. $\sigma^2_{t|t-1}$ change over time

  ↳ large past squared obs imply a large cond. var. for the obs. y_t

  $$\sigma^2_{t|t-1} = \omega + \alpha y^2_{t-1}$$

- ε_t is a martingale difference sequence

  $$E[\varepsilon_t \mid \mathcal{Y}_{t-1}] = 0 \implies E[\varepsilon_t] = 0 \text{ and } E[\varepsilon_t \cdot g(\mathcal{Y}_{t-1})] = 0$$

Unconditional Variance

by the law of iterated expectations

$$E_{t-1}(y_t^2) = \omega + \alpha y^2_{t-1}$$

$$E_{t-2}[E_{t-1}(y_t^2)] = \omega + \omega\alpha + \alpha^2 \cdot y^2_{t-2}$$

$$\vdots$$

$$E_{t-j} \dots E_{t-2} E_{t-1}(y_t^2) = \omega + \omega \cdot \alpha + \omega \cdot \alpha^2 + \dots + \omega \alpha^{j-1} + \alpha^j \cdot y^2_{t-j}$$

So, letting $j \to +\infty$, $(0 < \alpha < 1)$

- $\lim_{j \to +\infty} E_{t-j}[\sigma^2_{t|t-1}] \Rightarrow \alpha^j \to 0$
- $\omega \sum_{h=0}^{j-1} \alpha^h \longrightarrow \dfrac{\omega}{1-\alpha}$ since $0<\alpha<1$ this is a geometric series so it converges to $\frac{1}{1-\alpha}$
- $E_{t-j} E_{t-j+1} \dots E_{t-2} E_{t-1}(y_t^2) \longrightarrow E(y_t^2) = Var(y_t) = E(y_t^2) = \dfrac{\omega}{1-\alpha}$

• Remarks

- Conclusion:

$$\left.\begin{aligned} E(y_t) &= 0 \quad \text{(martingale difference sequence)} \\ Var(y_t) &= \frac{w}{1-\alpha} \quad \text{(by law of iterated expectations)} \\ E(y_t y_{t-k}) &= 0 \end{aligned}\right\} \Rightarrow \text{unconditionally } y_t \sim WN\left(0, \frac{w}{1-\alpha}\right)$$

- y_t are serially uncorrelated but not independent
- $z_t \sim N(0,1)$ where the odd moments of y_t are all equal to zero

• Fourth moment of an ARCH (1) process

$$E(y_t^4) = E(\sigma_{t|t-1}^4 \cdot z_t^4)$$

$$= E\left[(w + \alpha y_{t-1}^2)^2 \cdot z_t^4\right]$$

$$= E\left[(w + \alpha y_{t-1}^2)^2\right] E(z_t^4)$$

the fourth moment of a normal distribution is $3 \cdot \sigma^4$

$$= E\left[w^2 + \alpha^2 y_{t-1}^4 + 2\alpha w y_{t-1}^2\right] \cdot 3$$

$$E(y_t^4) = 3 \cdot \left(w^2 + \alpha^2 E(y_{t-1}^4) + 2\alpha w\, E(y_{t-1}^2)\right)$$

$$E(y_t^4) = 3 \cdot \left(w^2 + \alpha^2 E(y_{t-1}^4) + 2\alpha \frac{w^2}{1-\alpha}\right)$$

$$= 3w^2 + 3\alpha^2 \cdot E\left[y_{t-1}^4\right] + 3\left(2\alpha \cdot \frac{w^2}{1-\alpha}\right)$$

$$(1-3\alpha^2) \cdot E(y_t^4) = 3w^2\left[1 + \frac{2\alpha}{1-\alpha}\right]$$

$$(1-3\alpha^2)\, E(y_t^4) = 3w^2\left(1 + \frac{2\alpha}{1-\alpha}\right)$$

$$E(y_t^4) = \frac{3w^2}{1-3\alpha^2}\left[\frac{1+\alpha}{1-\alpha}\right]$$

Kurtosis of an ARCH (1) process

$$\frac{E[y_t^4]}{[E(y_t^2)]^2} = \frac{3w^2}{1-3\alpha^2}\left[\frac{1+\alpha}{1-\alpha}\right] \cdot \frac{(1-\alpha)^2}{w^2}$$

$$= 3 \cdot \left[\frac{1-\alpha^2}{1-3\alpha^2}\right] \geq 3$$

- $E(y_t^4) < +\infty \iff 3\alpha^2 < 1 \rightarrow \alpha < \frac{1}{\sqrt{3}}$
- The distribution of y_t is leptokurtic

$AR(1)$ representation of a squared $ARCH(1)$ model

The squared process y_t^2 is an $AR(1)$ process:

$$y_t^2 = \sigma^2_{t|t-1} + \left(y_t^2 - \sigma^2_{t|t-1}\right)$$

$$= \sigma^2_{t|t-1} + v_t$$

where $v_t = y_t^2 - \sigma^2_{t|t-1}$ is a martingale diff. sequence

$$= \omega + \alpha y^2_{t-1} + v_t$$

The autocorrelation function of y_t^2 is $\rho_k = \alpha^{|k|}$

Weaknesses of ARCH models

- Build upon the square returns, the model assume that positive and negative shocks have the same effects on volatility
- ARCH models are restrictive in capturing heavy tails in the Gaussian errors assumption given that $0 < \alpha^2 < \frac{1}{3}$
- ARCH models are likely to overpredict the volatility

Generalised ARCH models

Let's recall that an ARCH (p) model is:

$$y_t = \sigma_{t|t-1} \cdot z_t \quad , \quad z_t \sim \text{IID } N(0,1)$$

$$\sigma^2_{t|t-1} = \omega + \alpha_1 y^2_{t-1} + \dots + \alpha_p y^2_{t-p} \quad , \text{ with } \quad 0 < \alpha < \frac{1}{\sqrt{3}}$$

but when $p > 1$ they are difficult to estimate due to the non negativity constraints.

So an alternative is GARCH (1, 1) model

let's define the model as:

$$y_t = \sigma_{t|t-1} \cdot z_t$$] → Gaussian ARCH(1)

and

$$\sigma^2_{t|t-1} = \omega + \alpha y^2_{t-1} + \overbrace{\beta \sigma^2_{t-1|t-2}}^{\text{GARCH component}}$$

with $\omega > 0$, $0 \leq \alpha \, \beta \leq 1$, $\alpha + \beta < 1$

The constraint $\alpha + \beta < 1$ guarantees that the unconditional variance of y_t is finite.

In general, a GARCH(p, q)

$$\sigma^2_{t|t-1} = \omega + \alpha_1 y^2_{t-1} + \dots + \alpha_p \cdot y^2_{t-p} + \beta_1 \alpha^2_{t-1} + \dots + \beta_q \cdot \alpha^2_{t-q}$$

ARMA (1,1) representation of a GARCH (1,1) model

Let's define

$$v_t = y_t^2 - \sigma^2_{t|t-1} \quad , \quad \text{so that} \quad \sigma^2_{t-1} = y^2_{t-1} - v_{t-1}$$

so we can write the GARCH model as an ARMA (1,1) model for the squared observations

$$y_t^2 = \sigma^2_{t|t-1} + y_t^2 - \sigma^2_{t|t-1}$$

$$= \sigma^2_{t|t-1} + v_t$$

$$= \omega + \alpha y^2_{t-1} + \beta \sigma^2_{t-1} + v_t$$

$$= \omega + \alpha y^2_{t-1} + \beta \left(y^2_{t-1} - v_{t-1} \right) + v_t$$

$$= \omega + (\alpha + \beta) y^2_{t-1} + v_t - \beta \cdot v_{t-1}$$

Recall that v_t is a martingale difference sequence, and thus it is white noise in a weak sense

Unconditional moments of a GARCH (1,1)

As y_t is a m.d.s., $E(y_t) = 0$

If $\alpha + \beta < 1$, the ARMA (1,1) process is stationary so

$$E(y_t^2) = \frac{\omega}{1 - (\alpha + \beta)}$$

And it can be shown that

$$\frac{E(y_t^4)}{E[(y_t^2)]^2} = 3 \cdot \frac{1 - (\alpha + \beta)^2}{1 - (\alpha + \beta)^2 - 2\alpha^2} > 3 \qquad \text{Kurtosis} > \text{Normal}$$

GARCH process as a filter

$$E(y_t^2 \mid \mathcal{F}_{t-1}) = \sigma^2_{t|t-1}$$

$$= \omega + \alpha y^2_{t-1} + \beta \sigma^2_{t-1|t-2}$$

$$= \omega + \alpha y^2_{t-1} + \beta \left(\omega + \alpha y^2_{t-2} + \beta \sigma^2_{t-2|t-3} \right)$$

$$= \omega \left(1 + \beta + \dots + \beta^{j-1} \right) + \alpha \left(y^2_{t-1} + \beta y^2_{t-2} + \dots + \beta^{j-1} y^2_{t-j} \right)$$

$$+ \beta^j \sigma^2_{t-j|t-j+1}$$

and when $j \to +\infty$

$$\lim_{j \to +\infty} E(y_t^2 \mid \mathcal{F}_{t-1}) = \frac{\omega}{1-\beta} + \alpha \sum_{j=1}^{+\infty} \beta^{j-1} \cdot y^2_{t-1}$$

Prediction

The one step ahead predictive variance is given by the model

$$\mathrm{Var}(y_{t+1} \mid \mathcal{F}_t) = E_t(y^2_{t+1})$$

$$= \sigma^2_{t+1}$$

$$= \omega + \alpha y_t^2 + \beta \sigma^2_{t|t-1}$$

distribution:

$$y_{t+1} \mid \mathcal{F}_t \sim N(0, \sigma^2_{t+1|t})$$

In general:

$$\mathrm{Var}(y_{t+\ell} \mid \mathcal{F}_t) = E_t(y^2_{t+\ell})$$

$$= E_t \dots E_{t+\ell-1} (y^2_{t+\ell})$$

$$= E_t(\sigma^2_{t+\ell})$$

$$= E_t \left(\omega + \alpha y^2_{t+\ell-1} + \beta \sigma^2_{t+\ell-1} \right)$$

$$= E_t \dots E_{t+\ell-2} \left(\omega + \alpha y^2_{t+\ell-1} + \beta \sigma^2_{t+\ell-1} \right)$$

$$= \omega + (\alpha + \beta) E_t (\sigma^2_{t+\ell-1})$$

Thus

ℓ step ahead forecast of the conditional variance is given by this difference equation

$$E_t\left(\sigma^2_{t+\ell}\right) = \omega + (\alpha+\beta)\, E_t\left(\sigma^2_{t+\ell-1}\right)$$

$$E\left(y^2_t\right) = \frac{\omega}{1-(\alpha+\beta)} = \sigma^2, \text{ as } \ell \to +\infty$$

We have derived

$$E_t\left(\sigma^2_{t+\ell}\right) = \omega + (\alpha+\beta)\, E_t\left(\sigma^2_{t+\ell-1}\right)$$

and we can write

$$\begin{aligned}
\sigma^2_{t+\ell|t} &= \omega + (\alpha+\beta)\,\sigma^2_{t+\ell-1|t} \\
&= \omega + (\alpha+\beta)\cdot\left[\omega + (\alpha+\beta)\cdot\sigma^2_{t+\ell-2|t}\right] \\
&= \omega \sum_{j=0}^{h-2} (\alpha+\beta)^j + (\alpha+\beta)^{h-1}\cdot\sigma^2_{t+1|t} \\
&= \omega\,\frac{1-(\alpha+\beta)^{h-1}}{1-(\alpha+\beta)} + (\alpha+\beta)^{h-1}\cdot\sigma^2_{t+1|t} \\
&= \sigma^2\left[1-(\alpha+\beta)^{h-1}\right] + (\alpha+\beta)^{h-1}\cdot\sigma^2_{t+1|t} \\
&= \sigma^2 + (\alpha+\beta)^{h-1}\cdot\left(\sigma^2_{t+1|t} - \sigma^2\right)
\end{aligned}$$

• Models with different error distributions

Let's assume that z_t has a distribution with heavier tails than $N(0,1)$

For instance let's take a std. Student t distribution with $\nu > 2$ degree of freedoms

Gamma function

$$f(y_t \mid y_{t-1}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)\sigma_t^2}} \left[1 + \frac{(y_t - \mu_{t|t-1})^2}{(\nu-2)\sigma^2_{t|t-1}}\right]^{-\left(\frac{\nu+1}{2}\right)}$$

GARCH (1, 1) with dynamics driven by different MDS

Let $\quad y_t = \sigma_{t|t-1} \cdot z_t \;, \quad z_t \sim NID(0,1)$

with conditional variance

$$\sigma^2_{t|t-1} = \omega + \alpha \cdot y^2_{t-1} + \beta\, \sigma^2_{t-1|t-2}$$

or equivalently $\quad \sigma^2_{t|t-1} = \omega + (\alpha+\beta)\sigma^2_{t-1} + \alpha \cdot v_{t-1}$

one can write

$$v_t = \sigma^2_{t|t-1}\, w_t$$

where

$$w_t = \frac{y_t^2}{\sigma^2_{t|t-1}} - 1 \quad \Longrightarrow \text{ is a martingale diff. sequence}$$

so

$$\sigma^2_{t|t-1} = \omega + (\alpha+\beta)\sigma^2_{t-1|t} + \alpha\, \sigma^2_{t-1} \cdot w_{t-1}$$

Summarising

for a GARCH (1,1), $\quad y_t = \sigma_{t|t-1} \cdot z_t \quad , \quad z_t \sim IID(0,1)$

the conditional variance is $\quad \sigma^2_{t+1|t} = \omega + \underbrace{\phi}_{(\alpha+\beta)}\, \sigma^2_{t|t-1} + \alpha\, \sigma^2_{t|t-1} \cdot w_t$

and $\quad v_t = y_t^2 - \sigma_t^2$

First order dynamic conditional score model (DCS)

$$\sigma^2_{t+1|t} = \omega + \phi\,\sigma^2_{t|t-1} + \kappa\,\underbrace{\sigma^2_{t|t-1}\cdot u_t}_{}$$

martingale difference sequence different from v_t when z_t is non Gaussian

- dynamic: the model accounts for changes and dependencies over time
- conditional: conditional score is the $\frac{d}{dx}$ of the log-likelihood function
- score: measures the sensitivity of the likelihood to changes in the model parameters
- first order: consider only the immediate past information

- ## Beta-t-GARCH(1,1) processes

$$y_t = \sigma_{t|t-1}\cdot z_t \quad , \qquad \nu > 2$$

$z_t = \left(\frac{(\nu-2)}{\nu}\right)^{\frac{1}{2}}\cdot z^*_t$ has a $t_\nu(0,1)$ distribution

Assume $\quad y_t \mid y_{t-1} \sim \phi(\underset{\substack{\downarrow \\ \text{time varying parameter}}}{\mu_t}, \sigma)$

Classical models for time varying mean

UC ARMA models

$\Downarrow$

State space form

Kalman filter

$$y_t \mid y_{t-1} \sim \phi(\underset{\substack{\downarrow \\ \mu=0}}{0}, \sigma_t) \Longrightarrow \text{ARCH/GARCH models}$$

Classical models for time varying variance

- Let's consider AR(1) process

$$y_t =$$

Conditional mean and unconditional mean

$\underbrace{\text{Conditional variance}}_{\text{does not depend on time}}$ and unconditional variance

Score driven models

parametric specific.

idea: cond distribution of the dgp is given, can be heavy-tailed
and the location or scale parameters change over time

transformation of the prediction error

$$y_t \mid \mathcal{Y}_{t-1} \sim \phi(y_t ; \lambda_t , \theta) \quad \text{where} \quad \lambda_t = \omega + \phi \cdot \lambda_{t-1} + \kappa \cdot u_{t-1}$$

$$u_t \propto \underbrace{\frac{\partial}{\partial \lambda_t} \log p(y_t ; \lambda_t , \theta)}_{\text{score function}}$$

→ the dynamics of the time varying parameters are updated by a filter and driven by the score of the conditional distribution

- Key property of the score

Why the score? → Robustness

generic specific

$$\lambda_t = \mu_t \quad , \quad \phi(y_t ; \lambda_t , \theta) \sim t_\nu(\mu_t , \sigma)$$

$$\mu_t = \omega + \phi \, \mu_{t-1} + \kappa \cdot u_{t-1} \quad , \quad \theta = (\nu, \sigma, \phi, \omega, \kappa)$$

- Innovation form of KF

$$v_t = y_t - \mu_{t|t-1}$$

- First order DCS model for the location

DCS: dynamic conditional score

Let

$$p(y_t \mid \mathcal{Y}_{t-1}) \sim t_\nu(\mu_{t|t-1}, e^\lambda)$$

where

$$\mu_{t+1|t} = \delta + \phi\, \mu_{t|t-1} + \kappa\, u_t$$

with

$$u_t \propto \frac{\partial}{\partial \mu_{t|t-1}} \ell_t(\mu_{t|t-1})$$

we can write equivalently as above

$$y_t = \mu_{t|t-1} + v_t$$

$$\mu_{t+1|t} = \delta + \phi \, \mu_{t|t-1} + \kappa \cdot u_t$$

with

$$u_t = \left(1 + \frac{v_t^2}{\nu \cdot e^{2\lambda}} \right)^{-1} \cdot v_t \qquad \text{as } \nu \xrightarrow{+\infty} t_\infty \Rightarrow \text{Normal}$$

the idea of the score: