Modelos ARCH e GARCH

Análise de Séries Temporais

José Augusto Fiorucci

Section 1

Modelo ARCH

Modelos ARCH

- Autoregressivo Condicionalmente Heterocedástico ARCH(p)
- Proposto em Engle (1982)
- Retornos $\{y_t\}$ são modelados como

$$
\begin{aligned}
y_t &= \sqrt{h_t}\varepsilon_t \\
h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \cdots + \alpha_p y_{t-p}^2
\end{aligned}
$$

em que $\alpha_0, \alpha_1, \ldots, \alpha_p$ são parâmetros e $\{\varepsilon_t\}$ é um ruído branco com $E[\varepsilon_t] = 0$ e $Var[\varepsilon_t] = 1$.

- Espaço Paramétrico
  - $\alpha_0 > 0, \alpha_1 \geq 0, \alpha_2 \geq 0, \ldots, \alpha_p \geq 0$
    - importante para garantir $h_t > 0$ para todo $t = 1, 2, 3, \ldots$.
  - $\alpha_1 + \alpha_2 + \cdots + \alpha_p < 1$
    - importante para garantir que $Var[y_t] > 0$ (será visto a frente).
    - geralmente chamada de restrição de estacionariedade (será visto a frente).
- Note que
  - $E_{t-1}[y_t] = E_{t-1}[\sqrt{h_t}\varepsilon_t] = \sqrt{h_t}\, E_{t-1}[\varepsilon_t] = 0$
  - $Var_{t-1}[y_t] = Var_{t-1}[\sqrt{h_t}\varepsilon_t] = h_t\, Var[\varepsilon_t] = h_t$

Propriedades:

- Média Incondicional

$$E[y_t] = E[E_{t-1}[y_t]] = E[0] = 0$$

  - lembrete: $E[E[X|Y]] = E[X]$

- Autocovariância

$$\begin{aligned} Cov(y_t, y_{t+k}) &= E[y_t * y_{t+k}] \\ &= E[E_{t+k-1}[y_t * y_{t+k}]] \\ &= E[y_t \, E_{t+k-1}[y_{t+k}]] \end{aligned}$$

Como $E_{t+k-1}[y_{t+k}] = 0$, então

$$Cov(y_t, y_{t+k}) = 0, \qquad \forall\, k \neq 0$$

  - ou seja, ao utilizar um modelo ARCH(p) para a série $\{y_t\}$, estamos assumindo que o processo é não autocorrelacionado.
  - um modelo ARMA geralmente é considerado inicialmente em situações em que a série observada $y_1, \ldots, y_n$ apresenta autocorrelações expressivas. Neste caso, o modelo ARCH(p) deve ser aplicado aos resíduos do modelo ARMA, gerando assim um modelo ARMA(p_1, q_1)-ARCH(p).

- Variância Incondicional

$$
\begin{aligned}
Var[y_t] = E[y_t^2] = E[E_{t-1}[y_t^2]] = E[h_t] \\
&= E[\alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \cdots + \alpha_p y_{t-p}^2] \\
&= \alpha_0 + \alpha_1 E[y_{t-1}^2] + \alpha_2 E[y_{t-2}^2] + \cdots + \alpha_p E[y_{t-p}^2] \\
&= \alpha_0 + \alpha_1 Var[y_{t-1}] + \alpha_2 Var[y_{t-2}] + \cdots + \alpha_p Var[y_{t-p}]
\end{aligned}
$$

  Suponto estacionariedade, segue que

$$Var[y_t] = Var[y_{t-1}] = \cdots = Var[y_{t-p}]$$

  e portanto

$$Var[y_t] = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2 + \cdots + \alpha_p)}$$

- Estacionariedade:
  - Média constante?
    - Sempre verdade
  - Autocovariância depende apenas da defasagem?
    - Verdade apenas se a condição $\alpha_1 + \alpha_2 + \cdots + \alpha_p < 1$ esta satisfeita. Caso contrário, a variância não será constante.

- Curtose
  - Afim de simplificar os cálculos, vamos assumir:
    - $\{\varepsilon_t\}$ é i.i.d. N(0,1)
    - $\{y_t\}$ é estacionário de quarta ordem, isto é, $\mu_{(4)} = E[y_t^4], \quad \forall t$.
    - $h_t = \alpha_0 + \alpha_1 y_{t-1}^2 \quad \text{(ARCH(1))}$
  - Assim
    - $E_{t-1}[y_t^4] = E_{t-1}[h_t^2 \varepsilon_t^4] = h_t^2 \, E[\varepsilon_t^4] = 3\, h_t^2 = 3(\alpha_0 + 2\alpha_0\alpha_1 y_{t-1}^2 + \alpha_1^2 y_{t-1}^4)$
    -

$$
\begin{aligned}
\mu_{(4)} &= E[y_t^4] = E[E_{t-1}[y_t^4]] = E[3(\alpha_0 + 2\alpha_0\alpha_1 y_{t-1}^2 + \alpha_1^2 y_{t-1}^4)] \\
&= 3(\alpha_0 + 2\alpha_0\alpha_1 Var[y_{t-1}] + \alpha_1^2 E[y_{t-1}^4]) \\
&= 3(\alpha_0 + 2\alpha_0\alpha_1 \frac{\alpha_0}{1-\alpha_1} + \alpha_1^2 \mu_{(4)})
\end{aligned}
$$

isolando $\mu_{(4)}$ na equação anterior, obtemos $\mu_{(4)} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$

  - Com isso, podemos calcular a Curtose

$$
K = \frac{E[y_t^4]}{E[y_t^2]^2} = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \frac{(1-\alpha_1)^2}{\alpha_0^2} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3
$$

    - Como $K > 3$ temos distribuição leptocúrtica.

Previsão de volatilidade

Seja $y_1, \ldots, y_n$ a série dos retornos observados.

Da definição, temos que a equação da volatilidade no tempo $n+k$ é dada por

$$h_{n+k} = \alpha_0 + \alpha_1 y_{n+k-1}^2 + \alpha_2 y_{n+k-2}^2 + \cdots + +\alpha_p y_{n+k-p}^2$$

Portanto a previsão da volatilidade será consiste no valor esperado condicional $\widehat{h}_{n+k|n} = E_n[h_{n+k}]$, o qual pode ser facilmente calculado de forma iterativa para qualquer $k = 1, 2, 3, \ldots$, isto é,

$$\begin{aligned}
\widehat{h}_{n+1|n} &= \alpha_0 + \alpha_1 y_n^2 + \alpha_2 y_{n-1}^2 + \alpha_3 y_{n-2}^2 + \cdots + \alpha_p y_{n-(p-1)}^2 \\
\widehat{h}_{n+2|n} &= \alpha_0 + \alpha_1 \widehat{h}_{n+1|n} + \alpha_2 y_n^2 + \alpha_3 y_{n-1}^2 + \cdots + \alpha_p y_{n-(p-2)}^2 \\
\widehat{h}_{n+3|n} &= \alpha_0 + \alpha_1 \widehat{h}_{n+2|n} + \alpha_2 \widehat{h}_{n+1|n} + \alpha_3 y_n^2 + \cdots + \alpha_p y_{n-(p-3)}^2 \\
&\vdots
\end{aligned}$$

Section 2

Modelos GARCH

Modelos GARCH(p,q)

- Proposto por Bollerslev (1986)
- GARCH: "Generalized Autoregressive Conditional Heterocedastic"
- De acordo com o modelo GARCH(p,q) os retornos $\{y_t\}$ são modelados como

$$y_t = \sqrt{h_t}\,\varepsilon_t$$
$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2 + \cdots + \alpha_p y_{t-p}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \cdots + \beta_q h_{t-q}$$

  em que $\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ são parâmetros e $\{\varepsilon_t\}$ é um ruído branco com $E[\varepsilon_t] = 0$ e $Var[\varepsilon_t] = 1$.

- As restrições abaixo são necessárias para garantir que $Var_{t-1}[\varepsilon_t] > 0$ e $Var[\varepsilon_t] > 0$:

$$\alpha_0 > 0$$
$$\alpha_i \geq 0, \quad i = 1, \ldots, p$$
$$\beta_j \geq 0, \quad j = 1, \ldots, q$$
$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$$

- Parcimonia:
  - Modelos GARCH tendem a ser mais parcimoniosos que os ARCH
  - Dificilmente será necessário ordens maiores que (1, 1)
    - Na dúvida é recomendado testar as ordens (1,1); (2,1); (1,2) e (2,2) e escolher um deles via algum critério de parcimonia, como o AIC e BIC.

- Os modelos GARCH apresentam as mesmas propriedade vistas para os modelos ARCH,
  - Média zero
  - Sem autocorrelações
  - Distribuição leptocúrtica
  - Pode-se mostrar que a variância incondicional é dada por

$$Var[y_t] = \frac{\alpha_0}{1 - \left(\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j\right)}$$

Função de verossimilhança

- Note que
$$F_{t-1}(y_t) = P[\,\varepsilon_t < y_t \,/\sqrt{h_t}\,] = F_\varepsilon(y_t\,/\sqrt{h_t})$$
em que $f_\varepsilon(.)$ denota a f.d.a de ε_t (lembrando que $F_\varepsilon(.)$ é a mesma para qualquer t, pois $\{\varepsilon_t\}$ é um processo i.i.d).
- Derivando a equação anterior em y_t obtemos a f.d.p condicional de y_t
$$f_{t-1}(y_t) = (h_t)^{-1/2}\, f_\varepsilon(y_t\,/\sqrt{h_t})$$
em que $f_\varepsilon(.)$ denota a f.d.p de ε_t.
- Logo, se θ é conjunto de todos os parâmetros, então a função de verossimilhança dos modelos ARCH/GARCH pode ser escrita como
$$L(\theta) = \prod_{t=1}^{n} f_{t-1}(y_t) = \prod_{t=1}^{n}(h_t)^{-1/2}\, f_\varepsilon(y_t\,/\sqrt{h_t})$$
  - Note que a função anterior se aplica para qualque f.d.p de ε_t, apenas com as restrições $E[\varepsilon_t] = 0$ e $Var[\varepsilon_t] = 1$.

Função de verossimilhança

- Assumindo $\{\varepsilon_t\}$ como um processo i.i.d N(0,1), obtemos

$$L(\theta) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\, h_t}} \exp\left\{ -\frac{1}{2} \left(\frac{y_t}{\sqrt{h_t}} \right)^2 \right\}$$

  - O estimador de máxima verossimilhança pode ser obtido apenas de forma numérica.

- A suposição de normalidade geralmente não é observada em séries de retornos, pois essas demandam de distribuições mais leptocúrticas e algumas vezes assimétricas.
  - A abordagem mais comum é utilizar a distribuição t-Student padronizada (STD), bem como sua respectiva variação assimétrica (SSTD).

Resíduos

- Resíduos são calculados como

$$\varepsilon_t = y_t \,/ \sqrt{h_t}$$

além da verificação de aderência e autocorrelação de $\{\varepsilon_t\}$, deve-se vericar também autocorrelação de $\{\varepsilon_t^2\}$ no qual espera-se não observar a existência de volatilidade na série $\{\varepsilon_t\}$.

Previsão de volatilidade

A previsão de votilidade para k passos a frente de n é facilmente calculada para o modelo GARCH(p,q), uma vez que

$$\begin{aligned}\widehat{h}_{n+k|n} &= E_n[h_{n+k}] \\ &= \alpha_0 + \alpha_1 E_n[y^2_{n+k-1}] + \cdots + \alpha_p E_n[y^2_{n+k-p}] + \beta_1 E_n[h_{n+k-1}] + \cdots + \beta_q E_n[h_{n+k-q}]\end{aligned}$$

em que

$$E_n[y_j^2] = \begin{cases} y_j^2 \ , & j \leq n \\ \widehat{h}_{j|n} \ , & j > n \end{cases}$$

$$E_n[h_j] = \begin{cases} h_j \ , & j \leq n \\ \widehat{h}_{j|n} \ , & j > n \end{cases}$$

Assim as previsões são calculadas recursivamente para $\widehat{h}_{n+1|n}, \widehat{h}_{n+2|n}, \ldots, , \widehat{h}_{n+k|n}$.

Exemplo:

Previsão de volatilidade GARCH(1,1). Equação da volatilidade:
$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}, \quad t = 2, 3, 4, \dots$

- 1 passo a frente

$$\begin{aligned}\widehat{h}_{n+1|n} &= E_n[h_{n+1}] = \alpha_0 + \alpha_1 E_n[y_n^2] + \beta_q E_n[h_n] \\ &= \alpha_0 + \alpha_1 y_n^2 + \beta_1 h_n\end{aligned}$$

- 2 passos a frente

$$\begin{aligned}\widehat{h}_{n+2|n} &= E_n[h_{n+2}] = \alpha_0 + \alpha_1 E_n[y_{n+1}^2] + \beta_1 E_n[h_{n+1}] \\ &= \alpha_0 + \alpha_1 \widehat{h}_{n+1|n} + \beta_1 \widehat{h}_{n+1|n} \\ &= \alpha_0 + (\alpha_1 + \beta_1)\, \widehat{h}_{n+1|n}\end{aligned}$$

- 3 passos a frente

$$\begin{aligned}\widehat{h}_{n+3|n} &= E_n[h_{n+3}] = \alpha_0 + \alpha_1 E_n[y_{n+2}^2] + \beta_1 E_n[h_{n+2}] \\ &= \alpha_0 + (\alpha_1 + \beta_1)\, \widehat{h}_{n+2|n}\end{aligned}$$

Exemplo (continuação):

- Seguindo com o processo interativo, para $k \geq 2$ obtemos

$$\widehat{h}_{n+k|n} = \alpha_0 + (\alpha_1 + \beta_1)\,\widehat{h}_{n+k-1|n} = \alpha_0 \sum_{i=0}^{k-2} (\alpha_1 + \beta_1)^i \; + \; (\alpha_1 + \beta_1)^{k-1}\,\widehat{h}_{n+1|n}$$

- Da restrição $0 \leq \alpha_1 + \beta_1 < 1$ do modelo GARCH(1, 1), temos que no limite

$$\lim_{k \to \infty} \widehat{h}_{n+k|n} = \alpha_0 \sum_{i=0}^{\infty} (\alpha_1 + \beta_1)^i = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} = Var[y_t]$$

ou seja, a previsão de longo prazo da volatilidade (variância condicional) tende a variância incondicional do processo ($Var[y_t]$, cte pra todo t), como era esperado.

- O resultado anterior pode ser expandido para qualquer ordem dos modelos GARCH.

- Códigos e Exemplos direto no R.