2
Modelos de Volatilidade Condicional

2.1
Introdução

A teoria de análise de séries temporais, em particular os modelos ARIMA Box-Jenkins, permite a modelagem de variáveis de interesse, a partir da formulação explícita da dependência linear apresentada pelas mesmas.

Dentre os modelos empregados, podemos destacar os autoregressivos. Um processo autoregressivo de ordem p (AR(p)), para uma determinada variável y_t, tem a seguinte forma:

$$y_t = c + \phi_1 \,.\, y_{t-1} + \phi_2 \,.\, y_{t-2} + \ldots + \phi_p \,.\, y_{t-p} + u_t \qquad (2.1.1)$$

Onde, u_t assume a forma de um ruído branco.

$$E(w_t) = 0$$
$$E(w_t w_\tau) = \begin{cases} \lambda^2 & \rightarrow t = \tau \\ 0 & \rightarrow t \neq \tau \end{cases}$$

Dessa forma, pode-se afirmar que:

- O processo será estacionário de segunda ordem, caso as raízes do polinômio em “z” estejam localizadas fora do círculo unitário;

$$1 - \phi_1 \,.\, z - \phi_2 \,.\, z^2 - \ldots - \phi_p \,.\, z^p = 0 \qquad (2.1.2)$$

- A previsão um passo à frente, condicional a toda informação disponível no instante de tempo “t-1”, corresponde à projeção linear da variável de interesse sobre uma constante e seus próprios defasamentos (média condicional);
- Enquanto o valor da média condicional do processo é expresso como função daqueles observados para os respectivos defasamentos, a média incondicional, considerando o processo como sendo estacionário de segunda ordem, se mostra constante.

Para determinadas séries temporais, em particular aquelas observadas no mercado financeiro, pode ser interessante a realização de previsões, não somente relacionadas ao nível da variável, como também à variância da mesma.

A variância de séries financeiras representa um dos assuntos de maior importância na atualidade, uma vez que investidores racionais exigirão um prêmio maior, à medida que a volatilidade (incerteza) dos retornos de seus investimentos aumenta.

Aos pontos enumerados acima, soma-se o fato de que mudanças no regime da volatilidade afetam a própria eficiência das inferências estatísticas com relação aos parâmetros do modelo (c, ϕ_1, ϕ_2, ..., ϕ_p).

2.2
Autoregressive Conditional Heteroskedasticity (ARCH)

Introduzidos por Engle (1982), o modelo ARCH propõe uma formulação autoregressiva para capturar a estrutura de dependência linear existente na variância do ruído de uma determinada série de interesse.

Dessa forma, sendo o processo estocástico definido pela equação 2.1.1, diz-se que u_t^2 segue um processo ARCH (m), desde que este assuma a seguinte forma funcional:

$$u_t^2 = \varsigma + \alpha_1 \, . \, u_{t-1}^2 + \alpha_2 \, . \, u_{t-2}^2 + \ldots + \alpha_m \, . \, u_{t-m}^2 + w_t \qquad (2.2.1)$$

Onde, w_t corresponde a um ruído branco.

- Esperança condicional:

$$E\left[u_t^2 \middle| I_{t-1}\right] = \varsigma + \alpha_1 \, . \, u_{t-1}^2 + \alpha_2 \, . \, u_{t-2}^2 + \ldots + \alpha_m \, . \, u_{t-m}^2 \qquad (2.2.2)$$

- Como $u_t^2 \geq 0$, então, tem-se:
  - Equação 2.2.1 – deve ser não-negativa;
  - Equação 2.2.2 – deve ser positiva.
- Pode-se garantir que as condições acima sejam atendidas, se:
  - $-\varsigma < w_t$;
  - $0 < \varsigma$;
  - $\alpha_j \geq 0 \quad \rightarrow \quad j = 1, 2, \ldots, m$.
- Estacionaridade de segunda ordem (covariância):
  - Raízes do polinômio em Z devem corresponder a valores fora do círculo unitário.

$$1 - \alpha_1 \, . \, z - \alpha_2 \, . \, z^2 - \ldots - \alpha_m \, . \, z^m = 0 \qquad (2.2.3)$$

Como $\alpha_j \geq 0 \quad \rightarrow \quad j = 1, 2, \ldots, m$, então: $\alpha_1 + \alpha_2 + \ldots + \alpha_m < 1$

Uma vez satisfeitas as condições de estacionaridade de segunda ordem e aquelas referentes à própria característica dos dados de interesse, tem-se:

- Variância incondicional de u_t:

$$\sigma^2 = E\left[u_t^2\right] = \varsigma \big/ \left(1 - \alpha_1 - \alpha_2 - \ldots - \alpha_m\right) \qquad (2.2.4)$$

- Previsão "s" passos-à-frente:

$$\hat{u}_{t+s|t}^2 = E\left(u_{t+s}^2 \middle| u_t^2, u_{t-1}^2, \ldots\right) \qquad (2.2.5)$$

$$\left(\hat{u}_{t+j|t}^2 - \sigma^2\right) = \alpha_1\left(\hat{u}_{t+j-1|t}^2 - \sigma^2\right) + \alpha_2\left(\hat{u}_{t+j-2|t}^2 - \sigma^2\right) + \ldots + \alpha_m\left(\hat{u}_{t+j-m|t}^2 - \sigma^2\right) \qquad (2.2.6)$$

$j = 1, 2, \ldots, s$

Dessa forma, $\hat{u}^2_{t+s|t} \xrightarrow{p} \sigma^2$ (considerando os pressupostos assumidos).

2.2.1
Representação alternativa

Corresponde a um pressuposto mais forte sobre a dependência de u_t.
Seja:

$$u_t = \sqrt{h_t} \,.\, \nu_t \quad \rightarrow \quad \nu_t \sim iid\,(0,1) \tag{2.2.7}$$

$$E(\nu_t) = 0$$

$$E(\nu_t^2) = 1$$

Se $h_t = \varsigma + \alpha_1 \,.\, u^2_{t-1} + \alpha_2 \,.\, u^2_{t-2} + \ldots + \alpha_m \,.\, u^2_{t-m}$ (2.2.8)

Então:

ARCH(m): $E\left(u^2_{t+s} \middle| u^2_t, u^2_{t-1}, \ldots\right) = \varsigma + \alpha_1 \,.\, u^2_{t-1} + \alpha_2 \,.\, u^2_{t-2} + \ldots + \alpha_m \,.\, u^2_{t-m}$ (2.2.9)

Substituindo (2.2.7) e (2.2.8) em (2.2.1), segue:

$$h_t \,.\, \nu_t = h_t + w_t \quad \therefore \quad w_t = h_t\left(\nu_t^2 - 1\right) \tag{2.2.10}$$

Assim, a variância (condicional) de w_t será: $E\left(w_t^2\right) = \lambda^2$.

Cabe ressaltar que a variância incondicional de w_t (quarto momento – curtose) não existe para todos os modelos ARCH estacionários.

2.2.2
Estimação dos parâmetros (máxima verossimilhança)

- *ν_t é Gaussiano:*

Seja a equação de regressão

$$y_t = X_t^T \,.\, \beta + u_t \tag{2.2.11}$$

Onde,

X_t – variáveis explicativas (incluindo defasamentos de y_t)

$$u_t = \sqrt{h_t} \,.\, \nu_t \qquad\qquad \nu_t \sim N(0,1) \tag{2.2.12}$$

$$h_t = \xi + \alpha_1 \,.\, u_{t-1}^2 + \alpha_2 \,.\, u_{t-2}^2 + \ldots + \alpha_m \,.\, u_{t-m}^2$$

Condicionando a estimação nas m primeiras observações (t = -m+1, -m+2,..., 0). t = 1, 2, ..., T.

Seja Y_t = (y_t, y_{t-1},..., y_1, y_0,..., y_{-m+1}, X'_t, X'_{t-1},..., X'_1, X'_0,..., X'_{-m+1}). Então, sendo ν_t ~ i.i.d. N(0, 1) e ν_t independente de X_t e Y_t, tem-se que a distribuição condicional de y_t é gaussiana com média X'β e variância h_t.

$$f\left(y_t \middle| X_t ; Y_{t-1}\right) = \frac{1}{\sqrt{2 \,.\, \pi \,.\, h_t}} \,.\, \exp\left(-\frac{1}{2} \,.\, \frac{\left(y_t - X_t^T \,.\, \beta\right)^2}{h_t} \right) \tag{2.2.13}$$

Onde,

$$h_t = \xi + \alpha_1 \,.\, \left(y_{t-1} - X_{t-1}^T \,.\, \beta\right)^2 + \alpha_2 \,.\, \left(y_{t-2} - X_{t-2}^T \,.\, \beta\right)^2 + \ldots + \alpha_m \,.\, \left(y_{t-m} - X_{t-m}^T \,.\, \beta\right)^2$$

$$h_t = \left[Z(\beta)\right]^T \,.\, \delta \tag{2.2.14}$$

$$\delta = \left(\xi, \alpha_1, \alpha_2, \ldots, \alpha_m\right) \tag{2.2.15}$$

$$[Z(\beta)]^T = \left[1, \left(y_{t-1} - X_{t-1}^T \,.\, \beta\right)^2, \left(y_{t-2} - X_{t-2}^T \,.\, \beta\right)^2, \ldots, \left(y_{t-m} - X_{t-m}^T \,.\, \beta\right)^2\right] \qquad (2.2.16)$$

Sendo θ o vetor de hiperparâmetros.

$$\theta_{(ax1)} \equiv (\beta', \delta')$$

Log-verossimilhança:

$$L(\theta) = \sum_{t=1}^{T} \log\left[f\left(y_t \middle| X_t; Y_{t-1}; \theta\right)\right]$$

$$L(\theta) = -\left(\frac{T}{2}\right) . \log(2 \,.\, \pi) - \left(\frac{1}{2}\right) . \sum_{t=1}^{T} \log(h_t) - \left(\frac{1}{2}\right) . \sum_{t=1}^{T} \left[\frac{\left(y_t - X_t^T \,.\, \beta\right)^2}{h_t}\right] \qquad (2.2.17)$$

Trata-se, portanto, de um problema de otimização não-linear com restrições (estacionaridade).

$$Max\ L(\theta) = -\left(\frac{T}{2}\right) . \log(2 \,.\, \pi) - \left(\frac{1}{2}\right) . \sum_{t=1}^{T} \log(h_t) - \left(\frac{1}{2}\right) . \sum_{t=1}^{T} \left[\frac{\left(y_t - X_t^T \,.\, \beta\right)^2}{h_t}\right]$$

$$s.a.$$

$$\alpha_j \geq 0 \quad \rightarrow \quad para\ todo\ j$$

$$\sum_{j=1}^{m} \alpha_j < 1$$

2.3
Outras formulações propostas

2.3.1
GARCH

Seja:

$$u_t = \sqrt{h_t} \,.\, \nu_t \quad \rightarrow \quad \nu_t \sim iid\ (0, 1)$$

$$h_t = \varsigma + \alpha_1 \,.\, u_{t-1}^2 + \alpha_2 \,.\, u_{t-2}^2 + \ldots + \alpha_m \,.\, u_{t-m}^2$$

ARCH(m)

Pode-se imaginar um processo no qual a variância condicional dependa de um número infinito de defasamentos da variável de interesse.

$$h_t = \varsigma + \pi(L)\,.\,u_t^2 \quad (2.3.1)$$

$$\pi(L) = \sum_{j=1}^{\infty} \pi_j \,.\, L^j \quad (2.3.2)$$

Dessa forma, poder-se-ia definir π(L) como sendo a razão de dois polinômios de ordem finita, onde as raízes daquele contido no denominador fossem maiores que 1 (um) em módulo.

$$\pi(L) = \frac{\alpha(L)}{1-\delta(L)} = \frac{\alpha_1 \,.\, L^1 + \alpha_2 \,.\, L^2 + \ldots + \alpha_m \,.\, L^m}{1-\delta_1 \,.\, L^1 - \delta_2 \,.\, L^2 - \ldots - \delta_r \,.\, L^r}$$

Aplicando π(L) ter-se-ia:

$$h_t = \varsigma + \frac{\alpha(L)}{1-\delta(L)}\,.\,u_t^2$$

$$[1-\delta(L)]\,.\,h_t = [1-\delta(L)]\,.\,\varsigma + \alpha(L)\,.\,u_t^2$$

Assim, GARCH(r, m): $h_t - \delta(L)\,.\,h_t = K + \alpha(L)\,.\,u_t^2$ (2.3.3)

Adicionando u_t, tem-se:

$$h_t + u_t^2 - u_t^2 = K + \alpha(L)\,.\,u_t^2 + \delta(L)\,.\,h_t + \delta(L)\,.\,u_t^2 - \delta(L)\,.\,u_t^2$$

Fazendo $(u_t^2 - h_t) = w_t$, tem-se:

$$u_t^2 = K + \left[(\alpha_1 + \delta_1) \, . \, u_{t-1}^2 + \ldots + (\alpha_p + \delta_p) \, . \, u_{t-p}^2\right] + w_t - \delta_1 \, . \, w_{t-1} - \ldots - \delta_r \, . \, w_{t-r} \qquad (2.3.4)$$

Onde,

w_t – erro de previsão da variância (ruído branco) – seqüência de diferenças Martingale.

Assim, pode-se propor uma analogia entre o modelo GARCH (r, m) e os modelos ARMA (p, q), onde p = Max (r, m).

Pontos importantes:

- O requerimento de não-negatividade é satisfeito se:

$$K > 0; \alpha_j \geq 0; \delta_j \geq 0 \quad \rightarrow \quad para \ \ j = 1, \ldots, \ p$$

- De modo análogo aos processos ARMA, o modelo GARCH se mostrará estacionário de segunda ordem (covariância) se as raízes do polinômio em Z (conforme definido a seguir) apresentarem valores maiores do que 1 (um) em módulo.

$$1 - (\delta_1 + \alpha_1) \, . \, Z - (\delta_2 + \alpha_2) \, . \, Z^2 - \ldots - (\delta_p + \alpha_p) \, . \, Z^p = 0$$

Considerando as restrições de não-negatividade, o processo será estacionário se:

$$(\delta_1 + \alpha_1) + (\delta_2 + \alpha_2) + \ldots + (\delta_p + \alpha_p) < 1$$

Tomando por base o que fora apresentado até o presente momento, tem-se que:

- a média incondicional de u_t^2 será:

$$E(u_t^2) = \sigma^2 = \frac{K}{\left[1 - (\delta_1 + \alpha_1) - (\delta_2 + \alpha_2) - \ldots - (\delta_p + \alpha_p)\right]} \qquad (2.3.5)$$

- a previsão $\hat{u}_{t+s}^2$, considerando u_t^2, u_{t-1}^2, ..., u_0^2, será:

$$\hat{u}_{t+s}^2 - \sigma^2 = \begin{cases} (\delta_1 + \alpha_1)(\hat{u}_{t+s-1}^2 - \sigma^2) + \ldots + (\delta_1 + \alpha_1)(\hat{u}_{t+s-1}^2 - \sigma^2) \\ - \delta_s \hat{w}_t - \ldots - \delta_r \hat{w}_{t-r} \qquad\qquad p/\ s = 1, \ldots, r \\ \\ (\delta_1 + \alpha_1)(\hat{u}_{t+s-1}^2 - \sigma^2) + \ldots + (\delta_1 + \alpha_1)(\hat{u}_{t+s-1}^2 - \sigma^2) \qquad p/\ s > r \end{cases} \tag{2.3.6}$$

OBS.

O cálculo da seqüência de variâncias condicionais $\{h_t\}^T_{t=1}$ requer uma "amostra pré-existente" (h_{-p+1}, ..., h_0 e u_{-p+1}^2, ..., u_0^2). Para contornar tal problema, Bollerslev propôs o seguinte procedimento:

1) Fazer $h_j = u_j^2 = \hat{\sigma}^2 \quad \rightarrow \quad p/\ j = (-p+1), \ldots, 0$;
2) Definir $\hat{\sigma}^2 = T^{-1} \cdot \sum_{t=1}^{T} \left(y - X_t^T \cdot \beta \right)^2$;
3) Estimar a seqüência $\{h_t\}^T_{t=1}$ pelo método da máxima verossimilhança (otimização não-linear com restrições).

2.3.2
IGARCH

Dizemos que u_t^2 segue um processo IGARCH, caso este apresente raiz unitária (integrável). De forma análoga, tem-se:

$$\sum_{j=1}^{r} \delta_j + \sum_{l=1}^{m} \alpha_l = 1$$

Da mesma forma que nos processos ARIMA, se u_t^2 possui raiz unitária, então a variância incondicional tende para infinito (comportamento explosivo).

2.3.3
ARCH-M (ARCH "na média")

O modelo ARCH-M toma por base a teoria de finanças, a qual sugere que investidores racionais exigem um retorno maior (prêmio de risco) para negociarem ativos financeiros com maior grau de risco (risco associado à dispersão – volatilidade). Dessa forma, sendo:

$$r_t = \mu_t + u_t \tag{2.3.7}$$

Onde,

r_t – retorno de um determinado ativo financeiro;
μ_t – parcela do retorno antecipada pelos investidores em t-1;
u_t – parcela do retorno não antecipada.

Então, a teoria propõe que μ_t estaria associado com a própria variância do retorno -h_t (parâmetro de risco).

Tomando por base o que foi disposto anteriormente, em 1987 Engle propôs a seguinte formulação:

$$y_t = X_t^T \,.\, \beta + \delta \,.\, h_t + u_t \tag{2.3.8}$$

$$u_t = \sqrt{h_t} \,.\, \nu_t \qquad\qquad \nu_t \sim N(0,1) \tag{2.3.9}$$

$$h_t = \xi + \alpha_1 \,.\, u_{t-1}^2 + \alpha_2 \,.\, u_{t-2}^2 + \ldots + \alpha_m \,.\, u_{t-m}^2 \tag{2.3.10}$$

onde o efeito da volatilidade sobre o prêmio de risco (consolidado no retorno) é medido pelo hiperparâmetro δ.

2.3.4
E-GARCH (*Exponential* GARCH)

Seja:

$$u_t = \sqrt{h_t} \, . \, \nu_t \tag{2.3.11}$$

$$\log(h_t) = \xi + \sum_{j=1}^{\infty} \pi_j \, . \, \left\{ \left| \nu_{t-j} \right| - E\left| \nu_{t-j} \right| + \chi \, . \, \nu_{t-j} \right\} \tag{2.3.12}$$

então, se $\pi_j > 0$, um desvio $\left| \nu_{t-j} \right|$ ocasiona um aumento da variância condicional de u_t.

Já o termo χ, é responsável por modelar adequadamente o efeito resultante da assimetria presente na distribuição de retornos de ativos financeiros. Assim, tem-se:

- $\chi = 0$ – choques positivos e negativos geram o mesmo efeito na variância;
- $-1 < \chi < 0$ – choques positivos aumentam menos a variância do que os negativos;
- $\chi < -1$ – choques positivos reduzem a volatilidade e choques negativos produzem o efeito contrário.

Uma das vantagens da formulação do E-GARCH é o fato de o processo de estimação dos parâmetros não apresentar quaisquer tipos de restrições com relação aos valores que os mesmos poderão assumir, uma vez que estar-se-á trabalhando com $\log(h_t)$ e que $\pi_j > 0$.

Uma parametrização natural é modelar $\pi(L)$ como a razão de dois polinômios de ordem finita – procedimento análogo ao GARCH(r, m). Deste modo, ter-se-ia:

$$\log(h_t) = k + \delta_1 \,.\, \log(h_{t-1}) + \ldots + \delta_r \,.\, \log(h_{t-r}) + \alpha_1 \,.\, \{|\nu_{t-1}| - E|\nu_{t-1}| + \chi \,.\, \nu_{t-1}\}$$
$$+ \ldots + \alpha_m \,.\, \{|\nu_{t-m}| - E|\nu_{t-m}| + \chi \,.\, \nu_{t-m}\} \qquad (2.3.13)$$

2.3.4.1
Estimação dos parâmetros: máxima verossimilhança

Para o processo de estimação dos parâmetros, torna-se necessário especificar a função de distribuição de probabilidades de ν_t.

Uma alternativa seria a utilização da distribuição padronizada generalizada do erro (Nelson).

$$f(\nu_t) = \frac{\upsilon \,.\, \exp\left(-\frac{1}{2} \,.\, \left|\frac{\nu_t}{\lambda}\right|^{\upsilon}\right)}{\lambda \,.\, 2^{\left[(\upsilon+1)/\upsilon\right]} \,.\, \Gamma\left(\frac{1}{\upsilon}\right)} \qquad (2.3.14)$$

Onde,

$\Gamma(.)$ – função gama;

$$\lambda = \left\{\frac{2^{-2/\upsilon} \,.\, \Gamma\left(\frac{1}{\upsilon}\right)}{\Gamma\left(\frac{3}{\upsilon}\right)}\right\}^{1/2}$$

υ – parâmetro positivo que governa a largura das caudas.

- $\upsilon < 2$ – caudas largas;
- $\upsilon > 2$ – caudas estreitas.

OBS. Para $\upsilon = 2$, então, $\lambda = 1$ e, por conseguinte, $f(\nu_t) \sim Normal$.