MODELAGEM DA VOLATILIDADE

1. Modelos ARCH

- Objetivo: modelar a volatilidade de um retorno.

  Notação: considere uma série de retornos,

$$X_t = \ln(P_t) - \ln(P_{t-1}),$$

  e sejam

$$\mu_t = E(X_t|\mathcal{F}_{t-1}), \quad h_t = \mathrm{Var}(X_t|\mathcal{F}_{t-1}) \tag{1}$$

  a média e variância condicional de X_t, onde $\mathcal{F}_t$ denotará a informação até o instante t, que consideraremos ser

$\{X_t, X_{t-1}, \ldots, X_1\}$.

Em algumas situações iremos supor que $\mu_t = 0$, de modo que neste caso $h_t = E(X_t^2|\mathcal{F}_{t-1})$.

- Os modelos ARCH, ou modelos auto-regressivos com heteroscedasticidade condicional, foram introduzidos por Engle (1982), com o objetivo de estimar a variância da inflação. A idéia básica aqui é que o retorno X_t é não-correlacionado serialmente, mas a volatilidade (variância condicional) depende de retornos passados por meio de uma função quadrática.

- Um modelo ARCH(r) é definido por

$$X_t = \sqrt{h_t}\varepsilon_t, \tag{2}$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \ldots + \alpha_r X_{t-r}^2, \tag{3}$$

onde ε_t i.i.d. $(0,1)$, $\alpha_0 > 0, \alpha_i \geq 0, i > 0$.

Na prática, usualmente supomos $\varepsilon_t \sim \mathcal{N}(0,1)$, $\varepsilon_t \sim t_\nu$ ou uma distribuição que descreva melhor as caudas pesadas de séries financeiras.

- ARCH(1):

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, & (4) \\ h_t &= \alpha_0 + \alpha_1 X_{t-1}^2, & (5) \end{aligned}$$

com $\alpha_0 > 0, \alpha_1 \geq 0$.

Calculemos a média e variância incondicionais da série.

(i) $E(X_t) = E\{E(X_t|\mathcal{F}_{t-1})\} = 0$;

(ii) $\text{Var}(X_t) = E(X_t^2) = E\{E(X_t^2|\mathcal{F}_{t-1})\}$
$= E(\alpha_0 + \alpha_1 X_{t-1}^2) = \alpha_0 + \alpha_1 E(X_{t-1}^2)$.
Se o processo $\{X_t\}$ for estacionário de segunda ordem, então, para todo t, $E(X_t^2) = E(X_{t-1}^2) = \text{Var}(X_t)$, do que decorre

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \alpha_1}. \qquad (6)$$

Como $\text{Var}(X_t) > 0$, deveremos ter $0 \leq \alpha_1 < 1$.

(iii) $\text{Cov}(X_t, X_{t+k}) = E(X_t X_{t+k})$
$= E[X_t E(X_{t+k}|\mathcal{F}_{t+k-1})]$
$= E[X_t E(\sqrt{h_{t+k}}\varepsilon_{t+k}|\mathcal{F}_{t+k-1}) = 0$,
para $k > 0$, pois X_t está em $\mathcal{F}_{t+k-1}$ e $E(\varepsilon_{t+k}|\mathcal{F}_{t+k-1}) = 0$.

Dessa forma,

$$\gamma_X(k) = 0, \quad k \geq 1,$$

indicando que X_t é uma seqüência de variáveis não-correlacionadas (ruído branco), com média zero e variância dada por (6).

- Momento de quarta ordem:

$$\mu_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}. \qquad (7)$$

Supondo-se que momentos de quarta ordem sejam finitos e positivos, de (7) devemos ter $1 - 3\alpha_1^2 > 0$, ou seja, $0 \leq \alpha_1^2 < 1/3$. Portanto, quanto mais restrições impusermos ao processo de retornos, mais restrições teremos para os coeficientes do modelo. Isto é verdade para o modelo geral ARCH(r).

- Curtose:

$$\begin{aligned} K &= \frac{\mu_4}{[\mathrm{Var}(X_t)]^2} \\ &= 3\frac{\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}\frac{(1-\alpha_1)^2}{\alpha_0^2} \\ &= 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3. \end{aligned}$$

$X_t \sim$ ARCH $\to$ caudas serão mais pesadas do que as da normal, o que é uma propriedade vantajosa do modelo. Por outro lado, uma desvantagem do modelo é que trata retornos positivos e negativos de forma similar, já que quadrados dos retornos entram na fórmula da volatilidade. Na prática, sabe-se que a volatilidade reage de modo diferente a retornos positivos e negativos.

- Utilizando (4) e (5) e calculando $X_t^2 - h_t$, temos que

$$X_t^2 - (\alpha_0 + \alpha_1 X_{t-1}^2) = h_t(\varepsilon_t^2 - 1),$$

ou seja,

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t, \qquad (8)$$

na qual

$$v_t = h_t(\varepsilon_t^2 - 1) = h_t(X - 1), \qquad (9)$$

onde X é uma v.a. com distribuição $\chi^2(1)$, o que mostra que temos um modelo AR(1) para X_t^2, mas com erros não-gaussianos. Ainda, é fácil ver que $\{v_t\}$ é uma seqüência de v.a. de média zero, não-correlacionadas, mas com variância não-constante.

- Para um modelo ARCH(r) teremos

$$X_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i X_{t-i}^2 + v_t, \qquad (10)$$

onde os v_t são como no caso $r = 1$. Ou seja, temos um modelo AR(p) para X_t^2, com inovações não-gaussianas. Além disso, pode-se demonstrar que os retornos X_t também formam um ruído branco, com variância dada por

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{r} \alpha_i}.$$

- Identificação

  Um primeiro passo na construção de modelos ARCH é tentar ajustar modelos ARMA, para remover a correlação serial na série, se esta existir. Se este for o caso, teremos

$$\varphi(B)X_t = \theta_0 + \theta(B)\varepsilon_t,$$

  sendo que $\varepsilon_t \sim \text{ARCH}(r)$.

  Para verificarmos se a série apresenta heteroscedasticidade condicional, podemos utilizar dois testes, examinando-se a série X_t^2.

  (i) Teste de Box-Pierce-Ljung para X_t^2.

  (ii) Teste de multiplicadores de Lagrange (ML); veja Engle (1982). Queremos testar $H_0 : \alpha_i = 0$, para todo $i = 1, \ldots, r$, na regressão

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \ldots + \alpha_r X_{t-r}^2 + u_t,$$

para $t = r+1, \ldots, T$. A estatística do teste é $S = TR^2$, que tem distribuição assintótica $\chi^2(r)$ sob H_0. Aqui, R^2 é o quadrado do coeficiente de correlação múltipla da regressão acima. Um teste assintoticamente equivalente, que pode ter propriedades melhores para amostras pequenas, consiste em utilizar a estatística

$$F = \frac{(SQR_0 - SQR_1)/r}{SQR_1/(T-2r-1)} \sim F(r, T-2r-1), \tag{11}$$

na qual $SQR_0 = \sum_{t=r+1}^{T} (X_t^2 - \overline{X})^2$ e $SQR_1 = \sum_{t=r+1}^{T} \hat{u}_t^2$, com $\overline{X}$ a média amostral dos X_t^2 e $\hat{u}_t$ os resíduos de

MQ da regressão acima. Se o valor de F for significativo, dizemos que há heteroscedasticidade condicional na série.

Função de auto-correlação parcial de X_t^2 pode ser usada para encontrar a ordem r de um modelo ARCH(r).

- Estimação

  Os estimadores dos parâmetros do modelo são obtidos pelo método de máxima versossimilhança condicional. A função de verossimilhança é dada por

$$L(x_1, \ldots, x_T | \alpha)$$

$$= f(x_T | \mathcal{F}_{T-1}) f(x_{T-1} | \mathcal{F}_{T-2}) \cdots f(x_{r+1} | \mathcal{F}_r) f(x_1, \ldots, x_r | \alpha),$$

e supondo normalidade dos ε_t podemos escrever

$$L(x_1, \ldots, x_T | \alpha)$$

$$= \prod_{t=r+1}^{T} (\sigma_t \sqrt{2\pi})^{-1} \exp\{\frac{-x_t^2}{2\sigma_t^2}\} f(x_1, \ldots, x_r | \alpha).$$

Para T grande, $f(x_1, \ldots, x_r|\alpha)$ pode ser desprezado. Logo temos que maximizar a função de verossimilhança condicional

$$L(x_{r+1}, \ldots, x_T | \alpha, x_1, \ldots, x_r)$$

$$= \prod_{t=r+1}^{T} (\sigma_t \sqrt{2\pi})^{-1} \exp\{\frac{-x_t^2}{2\sigma_t^2}\},$$

onde a volatilidade $\sigma_t^2 = h_t$ é obtida recursivamente.

- Verificação

  Para um modelo ARCH(r), os resíduos padronizados

$$\tilde{X}_t = \frac{X_t}{\sqrt{h_t}}$$

são v.a. i.i.d. com distribuição normal padrão ou t- Student. Logo, uma maneira de verificar se o modelo é adequado é calcular a estatística Q de Ljung-Box, para a seqüência $\tilde{X}_t$. Além disso, podemos calcular os coeficientes de assimetria e curtose estimados e fazer um gráfico Q$\times$Q para avaliar a suposição de normalidade.

Para se verificar se ainda existe heteroscedasticidade condicional nos resíduos, pode-se aplicar o teste ML para a seqüência $\tilde{X}_t^2$.

- As previsões para a volatilidade utilizando o modelo ARCH(r) dado em (2)-(3) são obtidas recursivamente. Assim,

$$\widehat{h}_t(1) = \alpha_0 + \alpha_1 X_t^2 + \ldots + \alpha_r X_{t-r+1}^2, \tag{12}$$

é a previsão de h_{t+1}, com origem fixada no instante t. As previsões ℓ passos a frente, com origem em t, são dadas por

$$\widehat{h}_t(\ell) = \alpha_0 + \sum_{i=1}^{r} \alpha_i \widehat{h}_t(\ell - i), \tag{13}$$

em que $\widehat{h}_t(\ell - i) = X_{t+\ell-i}^2$, se $\ell - i \leq 0$.

- Exemplo: Retornos diários da Petrobraás.

  f.a.c.p. sugerem ARCH(3).

  Um primeiro modelo proposto, então, é

$$X_t = \phi_1 X_{t-1} + \phi_3 X_{t-3} + \phi_9 X_{t-9} + a_t,$$

$$a_t = \sqrt{h_t}\varepsilon_t,$$

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \alpha_3 a_{t-3}^2.$$

Assumindo $\varepsilon_t \sim \mathcal{N}(0,1)$ e utilizando o programa EVIEWS, obtemos que ϕ_3 e ϕ_9 não são significativamente diferentes de zero. Re-estimando o modelo, somente com o termo auto-regressivo de primeira ordem para representar a parte linear do modelo, obtemos o seguinte modelo ajustado:

$$\begin{aligned} X_t &= 0,1605 X_{t-1} + a_t, \\ a_t &= \sqrt{h_t}\varepsilon_t, \\ h_t &= 0,0004 + 0,1937 a_{t-1}^2 + 0,2370 a_{t-2}^2 \\ &\quad +0,2711 a_{t-3}^2, \end{aligned} \tag{14}$$

com todos os coeficientes significativamente diferentes de zero (p-valor $P = 0,0000$).

f.a.c. e f.a.c.p. dos resíduos indicam que modelo é adequado.

2. **Modelos GARCH**

- Uma generalização dos modelos ARCH foi sugerida por Bollerslev (1986, 1987, 1988), o chamado modelo GARCH ("generalized ARCH"). Vimos que um modelo ARMA pode ser mais parcimonioso, no sentido de apresentar menos parâmetros do que um modelo AR ou MA puro. Do mesmo modo, um modelo GARCH pode ser usado para descrever a volatilidade com menos parâmetros do que um modelo ARCH.

- Um modelo GARCH(r, s) é definido por

$$X_t = \sqrt{h_t}\varepsilon_t \tag{15}$$

$$h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i X_{t-i}^2 + \sum_{j=1}^{s} \beta_j h_{t-j}, \tag{16}$$

em que ε_t i.i.d. $(0,1)$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^{q}(\alpha_i+\beta_i) < 1$, $q = \max(r,s)$.

- Chamemos

$$\nu_t = X_t^2 - h_t, \tag{17}$$

de modo que, substituindo em (14) obtemos

$$X_t^2 = \alpha_0 + \sum_{i=1}^{q}(\alpha_i+\beta_i)X_{t-i}^2 + \nu_t - \sum_{j=1}^{s}\beta_j\nu_{t-j}, \tag{18}$$

ou seja, temos um modelo ARMA(q,s) para X_t^2, mas ν_t não é, em geral, um processo i.i.d. Na realidade, ν_t é uma diferença martingale, pois

$$E(\nu_t) = E(X_t^2 - h_t) = E(h_t\varepsilon_t^2 - h_t)$$

$$= E(h_t)(E(\varepsilon_t^2) - 1) = 0,$$

$$E(\nu_t|\mathcal{F}_{t-1}) = E(X_t^2|\mathcal{F}_{t-1}) - E(h_t|\mathcal{F}_{t-1})$$

$$= h_t - h_t = 0.$$

- Segue-se, em particular, que

$$E(X_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{q}(\alpha_i + \beta_i)}.$$

  A longo prazo, a volatilidade convergirá para esta média.

- GARCH (1,1)

  Um modelo bastante usado na prática é o GARCH$(1,1)$, para o qual a volatilidade é expressa como

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1}, \qquad (19)$$

com $0 \leq \alpha_1, \beta_1 < 1, \alpha_1 + \beta_1 < 1$.

- Curtose:

  Para o modelo (18) obtemos facilmente

$$\begin{aligned} K &= \frac{E(X_t^4)}{[E(X_t^2)]^2} \\ &= \frac{3[1-(\alpha_1+\beta_1)^2]}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2} > 3, \end{aligned}$$

  dado que o denominador seja positivo, o que novamente mostra que se X_t segue um modelo GARCH, as caudas de X_t serão mais longas do que as da normal.

- Identificação:

  A identificação da ordem de um modelo GARCH a ser ajustado a uma série real usualmente é difícil. Recomenda-se que se use modelos de ordem baixa, como (1,1), (1,2), (2,1) ou (2,2), e depois se escolha o modelo com base em vários critérios, como AIC ou BIC, valores da assimetria e curtose, da log-verossimilhança e de alguma função perda, como

$$\sum_{t=1}^{N} (X_t^2 - h_t)^2.$$

- Estimação:

  Os estimadores dos parâmetros do modelo (14)-(15) são obtidos pelo método de máxima verossimilhança condicional. Supondo normalidade dos ε_t, temos que a log-verossimilhança, condicional às primeiras r observações, é dada por

$$\ell(x_{r+1}, \ldots, x_T | \alpha, \beta, x_1, \ldots, x_r) \propto$$

$$-\frac{1}{2}\sum_{t=r+1}^{T}\ln(h_t)-\frac{1}{2}\sum_{t=r+1}^{T}\frac{x_t^2}{h_t}.$$

As estimativas dos parâmetros são obtidas por meio de métodos numéricos de maximização.

- Previsão:

  Previsões da volatilidade, usando um modelo GARCH, podem ser calculadas de forma similar àquelas de modelo ARMA. As previsões, com origem t, considerando um modelo GARCH$(1,1)$ da forma (18), são dadas por

$$\widehat{h}_t(1)=\alpha_0+\alpha_1 X_t^2+\beta_1 h_t,$$

e para $\ell>1$,

$$\begin{aligned}\widehat{h}_t(\ell) &= \alpha_0+\alpha_1\widehat{X}_t^2(\ell-1)+\beta_1\widehat{h}_t(\ell-1),\\ &= \alpha_0+\alpha_1\widehat{h}_t(\ell-1)\widehat{\varepsilon}_t^2(\ell-1)+\beta_1\widehat{h}_t(\ell-1),\end{aligned}$$

pois $X_t = \sqrt{h_t}\varepsilon_t$. Substituindo $\hat{\varepsilon}_t^2(\ell-1)$ por $E(\varepsilon_{t+\ell-1}^2) = 1$, temos que

$$\hat{h}_t(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\hat{h}_t(\ell-1), \quad \ell > 1. \tag{20}$$

- Exemplo: Retornos diários do IBOVESPA.

  f.a.c. e f.a.c.p. sugerem AR(10)+GARCH (1,1).

$$\begin{aligned} X_t &= 0,0020 + 0,0530X_{t-1} - 0,0518X_{t-5} \\ &\quad +0,0439X_{t-10} + a_t, \\ a_t &= \sqrt{h_t}\varepsilon_t, \\ h_t &= 0,00002 + 0,1938a_{t-1}^2 + 0,7806h_{t-1}. \end{aligned}$$

  diagnóstico: modelo adequado

3. **Modelos EGARCH**

- ARCH e GARCH tratam simetricamente os retornos, pois a volatilidade uma funço quadrática dos mesmos. Mas tambm sabido que a volatilidade reage de forma assimtrica aos retornos, tendendo a ser maior para retornos negativos.

  Tendo em vista o exposto, Nelson (1991) introduziu os modelos EGARCH ("exponential GARCH").

- Um modelo EGARCH(1,1) dado por

$$
\begin{aligned}
X_t &= \sqrt{h_t}\varepsilon_t, && (21)\\
\ln(h_t) &= \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(h_{t-1}), && (22)
\end{aligned}
$$

  em que ε_t são v.a. i.i.d. com média zero e variância um e $g(\cdot)$ a *curva de impacto de informaço*, dada por

$$
g(\varepsilon_t) = \theta\varepsilon_t + \gamma\{|\varepsilon_t| - E(|\varepsilon_t|)\}. \qquad (23)
$$

Aqui, θ e γ so parmetros reais, e $|\varepsilon_t| - E(|\varepsilon_t|)$ uma seqncia de v.a. i.i.d. com mdia zero.

- Note que podemos escrever

$$g(\varepsilon_t) = \begin{cases} (\theta + \gamma)\varepsilon_t - \gamma E(|\varepsilon_t|), & \varepsilon_t \geq 0 \\ (\theta - \gamma)\varepsilon_t - \gamma E(|\varepsilon_t|), & \varepsilon_t < 0. \end{cases} \tag{24}$$

Para que retornos negativos tenham maior impacto na volatilidade esperamos $\gamma < 0$.

- Se usarmos o EVIEWS para estimar modelos, este software considera o modelo EGARCH na forma:

$$\ln(h_t) = w + \beta \ln(h_{t-1}) + \alpha \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}. \tag{25}$$

O S+FinMetrics usa uma especificação similar. Quando $\gamma \neq 0$ o efeito assimtrico deve ser incorporado ao modelo GARCH. Observe tambm que podemos escrever $X_t/\sqrt{h_t}$ no lugar de ε_t nas frmulas acima.

- Exemplo: Retornos diários da Petrobrás

  AR(1) + EGARCH (1,1):

$$
\begin{aligned}
X_t &= 0,1576 X_{t-1} + a_t, \\
a_t &= \sqrt{h_t}\varepsilon_t, \\
\ln h_t &= -0,5231 + 0,9496 \ln h_{t-1} \\
&\quad +0,2200|\varepsilon_{t-1}| - 0,1420\varepsilon_{t-1},
\end{aligned}
$$

  usando o EVIEWS.

4. Modelos TARCH

- O modelo TARCH ("threshold ARCH") um caso particular do modelo ARCH no-linear, e a volatilidade agora segue a forma funcional

$$h_t^\gamma = \alpha_0 + \alpha_1 g^{(\gamma)}(\varepsilon_{t-1}) + \beta_1 h_{t-1}^\gamma,$$

em que

$$g^{(\gamma)}(\varepsilon_t) = \theta I_{\{\varepsilon_t > 0\}}|\varepsilon_t|^\gamma + (1-\theta) I_{\{\varepsilon_t \leq 0\}}|\varepsilon_t|^\gamma.$$

- O EVIEWS e o S+FinMetrics usam a formulação

$$h_t = w + \alpha X_{t-1}^2 + \gamma X_{t-1}^2 d_{t-1} + \beta h_{t-1},$$

com

$$d_t = \begin{cases} 1, & X_t < 0 \text{ (“bad news”)} \\ 0, & X_t \geq 0 \text{ (“good news”)}. \end{cases}$$

Se $\gamma \neq 0$, h um impacto de informaço assimtrica. Esperamos que $\gamma > 0$ para que "bad news" tenha um impacto maior.

- Exemplo: Retornos diários do Banespa.

  MA (1) + TARCH (1,1):

$$
\begin{aligned}
X_t &= a_t - 0,0686 a_{t-1}, \\
a_t &= \sqrt{h_t}\varepsilon_t, \\
h_t &= 0,00042 + 0,1672 a_{t-1}^2 \\
&\quad + 0,0999 a_{t-1}^2 d_{t-1} + 0,5893 h_{t-1},
\end{aligned}
$$

  diagnóstico: modelo adequado.

5. **Modelos de Volatilidade Estocástica**

- Os modelos da família ARCH supõem que a variância condicional depende de retornos passados. O modelo de volatilidade estocástica (MVE), primeiro proposto por Taylor (1980, 1986) não faz essa suposição. Este modelo tem como premissa o fato que a volatilidade presente depende de seus valores passados, mas é independente dos retornos passados.

- Notação:

  $\sigma_t^2 = E(X_t^2|\mathcal{F}_{t-1})$.

- Dizemos que a série X_t segue um modelo de volatilidade estocástica se

$$
\begin{aligned}
X_t &= \sigma_t \varepsilon_t, && (26)\\
\sigma_t &= e^{h_t/2}, && (27)
\end{aligned}
$$

  em que ε_t é uma seqüência estacionária, com média zero e variância um, e h_t é uma seqüência que pode se estacionária ou não, com uma densidade de probabilidade $f(h|\beta)$.

A formulação mais simples do modelo supõe que o logaritmo da volatilidade, h_t, seja dado por

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + \eta_t, \qquad (28)$$

na qual η_t é uma seqüência estacionária gaussiana, de média zero e variância σ_η^2, independente de ε_t. Segue-se que devemos ter $|\alpha_1| < 1$.

- Outras formulações do MVE foram apresentadas na literatura, dentre as quais destacamos as seguintes.

  (1) Forma canônica de Kim et al. (1998).

  Aqui, o MVE é escrito na forma

$$\begin{aligned} X_t &= \beta e^{h_t/2} \varepsilon_t, && (29) \\ h_{t+1} &= \mu + \alpha_1 (h_t - \mu) + \sigma_\eta \eta_t, && (30) \end{aligned}$$

com

$$h_t \sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\alpha_1^2}\right),$$

sendo $\varepsilon_t, \eta_t \sim \mathcal{N}(0,1)$, e se $\beta = 1$, então $\mu = 0$.

(2) Formulação de Jaquier et al. (1994), na qual

$$\begin{aligned} X_t &= \sqrt{h_t}\varepsilon_t, && (31)\\ \ln(h_t) &= \alpha_0 + \alpha_1 \ln(h_{t-1}) + \sigma_\eta \eta_t. && (32) \end{aligned}$$

- Sabemos que se $\varepsilon_t \sim \mathcal{N}(0,1)$, então $\ln(\varepsilon_t^2)$ tem uma distribuição chamada *log-qui-quadrado*, de tal sorte que

$$\begin{aligned} E(\ln(\varepsilon_t^2)) &\simeq -1,27\\ \mathrm{Var}(\ln(\varepsilon_t^2)) &= \pi^2/2. \end{aligned}$$

De (26) e (27) obtemos

$$
\begin{aligned}
\ln(X_t^2) &= \ln(\sigma_t^2) + \ln(\varepsilon_t^2), & (33)\\
h_t &= \ln(\sigma_t^2). & (34)
\end{aligned}
$$

Chamando $\xi_t = \ln(\varepsilon_t^2) - E(\ln(\varepsilon_t^2)) \simeq \ln(\varepsilon_t^2) + 1,27$, temos que $E(\xi_t) = 0$, $\text{Var}(\xi_t) = \pi^2/2$ e

$$
\begin{aligned}
\ln(X_t^2) &= -1,27 + h_t + \xi_t, & (35)\\
h_t &= \alpha_0 + \alpha_1 h_{t-1} + \eta_t, & (36)
\end{aligned}
$$

onde $\xi_t \sim$ i.i.d. $(0, \pi^2/2)$ e $\eta_t \sim$ i.i.d. $\mathcal{N}(0, \sigma_\eta^2)$.

Aqui, supomos ξ_t e η_t independentes.

- Propriedades

Vamos calcular agora alguns parâmetros associados ao MVE, considerando-se a forma (28).

(i) $E(X_t) = E(\sigma_t \varepsilon_t) = E(\sigma_t)E(\varepsilon_t) = 0$, dado que σ_t e ε_t são independentes.

(ii) $\text{Var}(X_t) = E(X_t^2) = E(\sigma_t^2 \varepsilon_t^2)$
$= E(\sigma_t^2)E(\varepsilon_t^2) = E(\sigma_t^2)$.

Dado que supusemos $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$, e h_t estacionário, com $\mu_h = E(h_t) = \alpha_0/(1-\alpha_1)$, σ_h^2=$\text{Var}(h_t) = \sigma_\eta^2/(1-\alpha_1^2)$, então obtemos

$$h_t \sim \mathcal{N}\left(\frac{\alpha_0}{1-\alpha_1}, \frac{\sigma_\eta^2}{1-\alpha_1^2}\right). \tag{37}$$

Como h_t é normal, σ_t^2 é log-normal, logo temos

$$E(X_t^2) = E(\sigma_t^2) = e^{\mu_h + \sigma_h^2/2}.$$

Não é difícil mostrar que

$$E(X_t^4) = 3e^{2\mu_h + 2\sigma_h^2},$$

da qual obtemos a curtose

$$K = \frac{3e^{2\mu_h + 2\sigma_h^2}}{e^{2\mu_h + \sigma_h^2}} = 3e^{\sigma_h^2} > 3, \qquad (38)$$

como deveríamos esperar, ou seja, caudas longas sob o MVE.

(iii) A função de auto-covariância da série X_t é dada por

$$\gamma_X(u) = E(X_t X_{t+u}) = E(\sigma_t \sigma_{t+u} \varepsilon_t \varepsilon_{t+u}) = 0,$$

pois ε_t e η_t são independentes. Logo X_t é serialmente não-correlacionada, mas não independente, pois existe correlação em $\ln(X_t^2)$. Denotando-se $Y_t = \ln(X_t^2)$, então a auto-covariância de Y_t é dada por

$$\gamma_Y(u) = E[(Y_t - E(Y_t))(Y_{t+u} - E(Y_{t+u}))].$$

Como o primeiro termo entre parênteses igual a $h_t - E(h_t) + \xi_t$ e h_t é independente de ξ_t, obtemos que

$$\begin{aligned}\gamma_Y(u) &= E[(h_t - E(h_t) + \xi_t)(h_{t+u} - E(h_{t+u}) + \xi_{t+u})] \\ &= E[(h_t - E(h_t))(h_{t+u} - E(h_{t+u})) + E(\xi_t \xi_{t+u})],\end{aligned}$$

e chamando as auto-covariâncias do segundo membro de $\gamma_h(\cdot)$ e $\gamma_\xi(\cdot)$, respectivamente, teremos

$$\gamma_Y(u) = \gamma_h(u) + \gamma_\xi(u),$$

para todo u.

Como estamos supondo (28), ou seja, um AR(1), temos que

$$\gamma_h(u) = \alpha_1^u \frac{\sigma_\eta^2}{1 - \alpha_1^2}, \;\; u > 0,$$

enquanto que $\gamma_\xi(u) = 0$, para $u > 0$. Logo, $\gamma_Y(u) = \gamma_h(u)$, para todo $u \neq 0$, e podemos escrever a função de auto-correlação de Y_t como

$$\rho_Y(u) = \frac{\gamma_Y(u)}{\gamma_Y(0)} = \frac{\alpha_1^u \sigma_\eta^2/(1-\alpha_1^2)}{\gamma_h(0) + \gamma_\xi(0)}, \;\; u > 0,$$

do que obtemos

$$\rho_Y(u) = \frac{\alpha_1^u}{1 + \pi^2/2\sigma_h^2}, \; u > 0, \qquad (39)$$

que tende a zero exponencialmente a partir do lag 2, o que indica que $Y_t = \ln(X_t^2)$ pode ser modelada por um modelo ARIMA(1, 1).

Na prática, obtemos valores de α_1 prximos de um, o que implica o aparecimento de altas correlações para volatilidades e conseqüentes grupos de volatilidades na srie.

- Estimação

Os MVE são difíceis de estimar. Podemos usar a abordagem de Durbin e Koopman (1997a, 1997b, 2000), que consiste em usar o procedimento de quase-verossimilhança, por meio do Filtro de Kalman. Aqui, o modelo (26)-(27) é reescrito na forma

$$
\begin{aligned}
X_t &= \sigma \varepsilon_t e^{h_t/2}, && (40) \\
h_t &= \alpha_1 h_{t-1} + \eta_t, && (41)
\end{aligned}
$$

em que $\sigma = \exp\{\alpha_0/2\}$. Uma forma equivalente é dada por

$$
\begin{aligned}
\ln(X_t^2) &= \kappa + h_t + u_t, && (42) \\
h_t &= \alpha_1 h_{t-1} + \eta_t, && (43)
\end{aligned}
$$

em que $u_t = \ln(\varepsilon_t^2) - E(\ln(\varepsilon_t^2))$ e $\kappa = \ln(\sigma^2) + E(\ln(\varepsilon_t^2))$.

As equações (42)-(43) estão na formulação denominada de espaço de estados: (42) é a *equação de observação* e (43) é a *equação de estado*.

Observações:

(i) Quando ϕ for próximo de 1, o ajustamento de um MVE é similar ao de um GARCH(1,1), com $\alpha_1 + \beta_1$ próximo de 1.

(ii) Quando $\phi = 1$, h_t é um passeio aleatório e o ajustamento de um MVE é similar ao de um modelo IGARCH(1,1).

(iii) Quando algumas observações forem iguais a zero, o que pode ocorrer na prática, não podemos fazer a transformação logaritmica especificada em (42). Uma solução sugerida por Fuller e analisada por Breidt e Carriquiry (1996) é fazer a seguinte transformação baseada numa expansão de Taylor:

$$\ln(X_t^2) = \ln(X_t^2 + cS_X^2) - cS_X^2/(X_t^2 + cS_X^2),$$

t=1,. . . ,N,

em que S_X^2 é a variância amostral da série X_t e c é um número pequeno.

O programa STAMP (Koopman et al., 1995) pode ser utilizado na obtenção de estimadores de quase-verossimilhança (QMV) dos parâmetros do modelo, escrito na forma de espaço de estados. Este programa incorpora a transformação acima, com um valor “default” $c = 0,02$. Uma das vantagens da utilização do procedimento de QMV é que ele pode ser aplicado sem a especificação de uma particular distribuição para ε_t.

- Shephard e Pitt (1997) propuseram o uso de amostragem ponderada (“importance sampling”) para estimar a função de verossimilhança.

  Como o MVE é um modelo hierárquico, Jaquier et al. (1994) propuseram uma análise bayesiana para o mesmo. Veja tambm Shephard e Pitt (1997) e Kim et al. (1998).

- Exemplo: Vamos re-analisar a srie de retornos diários do IBOVESPA usando o MVE e programa STAMP. O modelo ajustado é

$$
\begin{aligned}
\ln(X_t^2) &= -8,6220 + h_t + u_t, \\
h_t &= 0,9858 h_{t-1} + \eta_t, \\
\hat{\text{Var}}(u_t) &= 2,7418, \\
\hat{\text{Var}}(\eta_t) &= 0,0236.
\end{aligned}
\tag{44}
$$

diagnóstico: modelo adequado