3
Modelos de Volatilidade Estocástica

3.1
O Modelo Univariado

O modelo de volatilidade estocástica (VE) é composto por duas equações básicas: uma que descreve o processo dos retornos (R_t), variável observável, e outra que descreve o processo de uma variável não observável (h_t), chave para a descrição da volatilidade dos retornos. Segue abaixo um modelo particular de VE em que h_t segue um processo autoregressivo de ordem 1 com intercepto:

$$R_t = \mu_t + \varepsilon_t e^{\frac{h_t}{2}} \qquad \varepsilon_t \sim NID(0,1) \tag{1}$$

$$h_t = \gamma + \phi h_{t-1} + \eta_t \quad \text{para} \quad |\phi| < 1, \quad \eta_t \sim NID(0, \sigma_\eta^2)$$

$$com\ E(\varepsilon_t, \eta_s) = 0 \quad \forall t, s$$

Ressalta-se que μ_t representa a média dos retornos. Um fato estilizado[6] comum a séries financeiras é a presença de média nula dos retornos. Por esta razão, μ_t é comumente omitido em diversos trabalhos para efeito de simplificação da notação.

A média e variância incondicionais dos retornos podem ser calculadas da seguinte forma:

$$\begin{aligned} E(R_t) &= E(\varepsilon_t e^{\frac{h_t}{2}}) \\ &= E(\varepsilon_t)E(e^{\frac{h_t}{2}}) \\ &= 0 \end{aligned} \tag{2}$$

---

[6] Fatos estilizados são regularidades estatısticas observadas em um grande número de séries financeiras de retornos

PUC-Rio - Certificação Digital Nº 0713589/CA

$$\begin{aligned} V(R_t) &= E(R_{t_t}^2) - [E(R_t)]^2 \\ &= E(R_{t_t}^2) \\ &= E(\varepsilon_t^2 e^{h_t}) \\ &= E(\varepsilon_t^2)E(e^{h_t}) \\ &= E(e^{h_t}) \end{aligned} \tag{3}$$

Pela teoria da Distribuição Normal:

Se

$$h_t \sim N(\mu_h, \sigma_h^2) \text{ então } \quad e^{h_t} \sim Lognormal\left(e^{\mu_h + \frac{\sigma_h^2}{2}}, [(e^{2\mu_h + \sigma_h^2})(e^{\sigma_h^2} - 1)] \right) \tag{4}$$

Logo

$$V(R_t) = e^{\mu_h + \frac{\sigma_h^2}{2}} \tag{5}$$

Onde:

$$\begin{aligned} \mu_h &= E(h_t) \\ &= E(\gamma + \phi h_{t-1} + \eta_t) \\ &= E(\gamma + \phi(\gamma + \phi h_{t-2} + \eta_{t-1}) + \eta_t \\ &= E(\gamma + \phi\gamma + \phi^2\gamma + \ldots + \eta_t + \phi\eta_{t-1} + \phi^2\eta_{t-2} + \ldots) \\ &= E(\gamma + \phi\gamma + \phi^2\gamma + \ldots) + E(\eta_t + \phi\eta_{t-1} + \phi^2\eta_{t-2} + \ldots) \\ &= \frac{\gamma}{1-\phi} \end{aligned} \tag{6}$$

e

$$\begin{aligned} \sigma_h^2 &= V(h_t) \\ &= V(\gamma + \phi\gamma + \phi^2\gamma + \ldots + \eta_t + \phi\eta_{t-1} + \phi^2\eta_{t-2} + \ldots) \\ &= V(\gamma + \phi\gamma + \phi^2\gamma + \ldots) + V(\eta_t + \phi\eta_{t-1} + \phi^2\eta_{t-2} + \ldots) \\ &= \frac{\sigma_\eta^2}{1-\phi^2} \end{aligned} \tag{7}$$

Sob a hipótese de $\varepsilon_t \sim NID(0,1)$, temos E($\varepsilon_t^3$) = 0 e E($\varepsilon_t^4$)=3, o que implica em retornos simétricos e leptocúrticos, como demonstrado a seguir:

PUC-Rio - Certificação Digital Nº 0713589/CA

$$\begin{aligned}
Assimetria\,(R_t) &= \frac{E[(R_t - E(R_t)^3]}{V(R_t)^{3/2}} \\
&= \frac{E[R_t^{\;3}]}{V(R_t)^{3/2}} \\
&= \frac{E\left[\left(\varepsilon_t e^{\frac{h_t}{2}}\right)^3\right]}{V(R_t)^{3/2}} \\
&= \frac{E\left[\varepsilon_t^{\;3}\left(e^{\frac{3h_t}{2}}\right)\right]}{V(R_t)^{3/2}} \\
&= \frac{E(\varepsilon_t^{\;3})E\left(e^{\frac{3h_t}{2}}\right)}{V(R_t)^{3/2}} \\
&= 0
\end{aligned} \tag{8}$$

$$\begin{aligned}
Curtose(R_t) &= \frac{E[(R_t - E(R_t)^4]}{V(R_t)^{4/2}} \\
&= \frac{E[R_t^{\;4}]}{\left(e^{\mu_h + \frac{\sigma_h^2}{2}}\right)^2} \\
&= \frac{E\left[\left(\varepsilon_t e^{\frac{h_{tt}}{2}}\right)^4\right]}{\left(e^{\mu_h + \frac{\sigma_h^2}{2}}\right)^2} \\
&= \frac{E(\varepsilon_t^4)E(e^{2h_{tt}})}{\left(e^{\mu_h + \frac{\sigma_h^2}{2}}\right)^2} \\
&= \frac{3E(e^{2h_{tt}})}{\left(e^{2\mu_h + \sigma_h^2}\right)}
\end{aligned} \tag{9}$$

Novamente, pela teoria da Distribuição Normal:

Se $h_t \sim N(\mu_h, \sigma_h^2)$ então $2h_t \sim N(2\mu_h, 4\sigma_h^2)$ (10.a)

PUC-Rio - Certificação Digital Nº 0713589/CA

o que implica em:

$$e^{2h_t} \sim Lognormal\left(e^{2\mu_h + \frac{4\sigma_h^2}{2}}, [(e^{4\mu_h + 4\sigma_h^2})(e^{4\sigma_h^2} - 1)] \right) \tag{10.b}$$

Aplicando (10.b) em (9) temos:

$$\begin{aligned} Curtose(R_t) &= \frac{3e^{2\mu_h + \frac{4\sigma_h^2}{2}}}{\left(e^{2\mu_h + \sigma_h^2}\right)} \\ &= \frac{3e^{2\mu_h}.e^{2\sigma_h^2}}{e^{2\mu_h}.e^{\sigma_h^2}} \\ &= 3e^{\sigma_h^2} > 3 \end{aligned} \tag{11}$$

Provando, portanto, que os retornos são leptocúrticos.

A média e variância condicionais dos retornos podem ser obtidas de maneira similar:

$$\begin{aligned} E(R_t / \Im_{t-1}) &= E(\varepsilon_t e^{\frac{h_t}{2}} / \Im_{t-1}) \\ &= E(\varepsilon_t / \Im_{t-1}) E(e^{\frac{h_t}{2}} / \Im_{t-1}) \\ &= 0 \end{aligned} \tag{12}$$

$$\begin{aligned} V(R_t / \Im_{t-1}) &= E(R_{t_t}^2 / \Im_{t-1}) - [E(R_t) / \Im_{t-1}]^2 \\ &= E(R_t^2 / \Im_{t-1}) \\ &= E(\varepsilon_t^2 e^{h_t} / \Im_{t-1}) \\ &= E(\varepsilon_t^2 / \Im_{t-1}) E(e^{h_t} / \Im_{t-1}) \\ &= E(e^{h_t} / \Im_{t-1}) \end{aligned} \tag{13}$$

Pela teoria da Distribuição Normal: (14)

Se $h_t / \Im_{t-1} \sim N(\mu_{ch}, \sigma_{ch}^2)$ então $e^{h_t} / \Im_{t-1} \sim Lognormal\left(e^{\mu_{ch} + \frac{\sigma_{ch}^2}{2}}, [(e^{2\mu_{ch} + \sigma_{ch}^2})(e^{\sigma_{ch}^2} - 1)] \right)$

Desta forma, temos que

$$V(R_t / \Im_{t-1}) = e^{\mu_{ch} + \frac{\sigma_{ch}^2}{2}} \tag{15}$$

Conclui-se, portanto, que a estimação da variância condicional dos retornos passa pela estimação da média e variância condicionais de h_t,

PUC-Rio - Certificação Digital Nº 0713589/CA

representadas acima por μ_{ch} e σ^2_{ch}. Denominemos os estimadores de μ_{ch} e σ^2_{ch} , respectivamente, de a_t e P_t de forma que:

$$volatilidade\ estimada = \hat{V}(R_t / \Im_{t-1}) = e^{\hat{\mu}_h + \frac{\hat{\sigma}_h^2}{2}} = e^{at + \frac{Pt}{2}} \tag{16}$$

A estimação por *Quasi* Máxima Verossimilhança (QMV) é feita concomitantemente com a aplicação do filtro de Kalman, ferramenta que será descrita com maiores detalhes no capítulo 4 deste Trabalho. Por esta razão é preciso colocar o modelo (1) na forma de Espaço de Estado (EE), o que exige a linearização da equação dos retornos. Isto é feito através da elevação ao quadrado da primeira equação, seguida de sua logaritmização:

$$R_t^{\ 2} = \{\exp(h_t)\} * \varepsilon_t^2 \tag{17}$$

$$\log(R_t^{\ 2}) = h_t + \log(\varepsilon_t^{\ 2})$$

Em Abramovitz e Stugan (1970) mostra-se que a média e variância de $\log(\varepsilon_t^{\ 2})$ são, respectivamente, -1,27 e $\frac{\pi^2}{2}$. É conveniente, portanto, reescrever a equação (17) como:

$$\log(R_t^{\ 2}) + 1{,}27 = h_t + \log(\varepsilon_t^{\ 2}) + 1{,}27 \tag{18}$$

Denotemos $y_t = \log(R_t^{\ 2}) + 1{,}27$ e $\xi_t = \log(\varepsilon_t^{\ 2}) + 1{,}27$, de forma que $E(\xi_t) = 0$. Com isso podemos escrever um modelo em que o erro da equação das medidas tem média nula, como exige a forma em Espaço de Estado:

$$\begin{aligned} y_t &= h_t + \xi_t \qquad \xi_t \sim (0, \frac{\pi^2}{2}) \\ h_t &= \gamma + \phi h_{t-1} + \eta_t \qquad \eta_t \sim NID(0, \sigma^2_\eta) \end{aligned} \tag{19}$$

$$com\ E(\xi_t, \eta_s) = 0 \quad \forall t, s$$ [7]

É importante ressaltar que com base em análises empíricas os retornos das séries financeiras, em geral, não apresentam distribuição Normal. Desta forma, a

[7] Caso não fosse imposta a restrição de não correlação entre ε_t e η_s no modelo original (1), pode-se demonstrar que ξ_t e η_s , $\forall t, s$, são descorrelatados [cf. Harvey, Ruiz e Shephard (1994)].

PUC-Rio - Certificação Digital Nº 0713589/CA

solução via QMV por filtro de Kalman, quando se supõe normalidade dos dados, é sub-ótima.

3.2
O Modelo Multivariado

A extensão do modelo univariado para o multivariado é imediata. Supondo p séries com n observações cada, pode-se reescrever o modelo (19) da seguinte forma:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \dots \\ y_{p,t} \end{bmatrix}_{px1} = \begin{bmatrix} h_{1,t} \\ h_{2,t} \\ \dots \\ h_{p,t} \end{bmatrix}_{px1} + \begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \\ \dots \\ \xi_{p,t} \end{bmatrix}_{px1} \qquad \begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \\ \dots \\ \xi_{p,t} \end{bmatrix}_{pxn} \sim \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{px1}, H_{t_{pxp}} \right)$$

$$\begin{bmatrix} h_{1,t+1} \\ h_{2,t+1} \\ \dots \\ h_{p,t+1} \end{bmatrix}_{px1} = \begin{bmatrix} \gamma_{1,t} \\ \gamma_{2,t} \\ \dots \\ \gamma_{p,t} \end{bmatrix}_{px1} + \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1p} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2p} \\ \dots & \dots & \dots & \cdots \\ \phi_{p1} & \phi_{p2} & \dots & \phi_{pp} \end{bmatrix} \begin{bmatrix} h_{1,t} \\ h_{2,t} \\ \dots \\ h_{p,t} \end{bmatrix}_{px1} + \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \dots \\ \eta_{p,t} \end{bmatrix}_{px1} \qquad (20)$$

sendo $H_t = \dfrac{\pi^2}{2} \begin{bmatrix} 1 & \rho_{1,2}{}^* & \dots & \rho_{1,p}{}^* \\ \rho_{2,1}{}^* & 1 & \dots & \rho_{2,p}{}^* \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{p,1}{}^* & \rho_{p,2}{}^* & \dots & 1 \end{bmatrix}_{pxp}$ e $\begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \dots \\ \eta_{p,t} \end{bmatrix}_{px1} \sim \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{px1}, Q_{t_{pxp}} \right)$

Após a elevação ao quadrado da equação dos retornos seguida da logaritmização dos termos, surge uma dificuldade em recuperar os sinais originais dos erros da equação das medidas (ε_t) a partir dos elementos de H_t. Isto é percebido, quando se escrevem os coeficientes de correlação de ξ_t ($\rho_{i,j}{}^*$) em função dos elementos da matriz de correlação dos erros originais dos retornos (ε_t), representada aqui por $\wp_\varepsilon$:

PUC-Rio - Certificação Digital Nº 0713589/CA

$$\wp_{\varepsilon} = \begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,p} \\ \rho_{2,1} & 1 & \cdots & \rho_{2,p} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{p,1} & \rho_{p,2} & \cdots & 1 \end{bmatrix}_{pxp} \tag{21}$$

Em Harvey, Ruiz e Shephard (1994) demonstra-se que os elementos $\rho_{i,j}$ * podem ser escritos como:

$$\rho_{i,j}{}^* = \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(n-1)!}{(0{,}5)_n n} \rho_{i,j}{}^{2n} \tag{22}$$

Onde $(0{,}5)_n = (0{,}5) * (0{,}5+1) * .(0{,}5+2) * ... * (0{,}5+n-1)$

Para recuperar os sinais dos coeficientes de correlação $\rho_{i,j}$ sugere-se retomar as séries originais dos retornos e obter para cada instante de tempo o produto cruzado das séries duas a duas. Se mais da metade dos elementos da série "produto" for composta de termos positivos, então o coeficiente de correlação entre os erros das séries multiplicadas é positivo. Para efeito de ilustração, suponha p=3. Com base nas séries originais dos retornos $R_{1t,} R_{2t}$ *e* R_{3t} (que dão origem respectivamente a $y_{1t,} y_{2t}$ *e* y_{3t} após a linearização) devem-se obter para cada instante de tempo, as seguintes séries "produto":

$$\begin{aligned} R_{1,2} &= R_{1t} * R_{2t} \\ R_{1,3} &= R_{1t} * R_{3t} \\ R_{23} &= R_{2t} * R_{3t} \end{aligned} \tag{23}$$

para t=1,…,n

Se o número de termos positivos na série $R_{i,j}$, para i<j, i, j =1,…,3 for maior que a metade do número de termos da série $R_{i,j}$ então fixa-se o sinal de $\rho_{i,j}$ como positivo, caso contrário, negativo.

PUC-Rio - Certificação Digital Nº 0713589/CA

3.3 Modelos de Fatores Comuns para Tendência

Como o nome antecipa, os modelos de fatores comuns para tendência são aplicados a bases multivariadas na tentativa de extrair fatores não observáveis, porém comuns às tendências das séries. Utilizando a notação de Harvey (1989, p.450), onde é feita uma abordagem completa sobre o tema, temos um modelo particular com p séries e k fatores comuns, os quais seguem um passeio aleatório:

$$Y_{t\,px1} = \Theta_{pxk}\mu_{t\,kx1} + \mu^0{}_{px1} + \varepsilon_{t\,px1}\ , \quad Var(\varepsilon_t) = \Sigma_{\varepsilon\,pxp} \tag{24}$$

$$\mu_{t+1\,kx1} = \mu_{t\,kx1} + \beta_{0\,kx1} + \eta_{t\,kx1} \qquad , \ Var(\eta_t) = \Sigma_{\eta\,kxk}$$

para k≤p,

sendo μ^0 vetor com as primeiras k coordenadas nulas.

O vetor μ_t contempla os fatores comuns que regem o processo das diferentes tendências das séries temporais em Y_t.

Entretanto, o modelo como apresentado em (24) não é identificável. A comprovação pode ser feita tomando uma matriz qualquer H_{kxk} ortogonal e fazendo $\Theta^x = \Theta H^{-1}$, $\mu^x = H\ \mu$, $\beta_0^x = H\ \beta_0$, $\eta^x = H\eta$, $Var(\eta^x) = H\Sigma_\eta H'$. Com isso temos:

$$Y_t = \Theta^x \mu_t{}^x + \mu^0 + \varepsilon_t \ , \qquad Var(\varepsilon_t) = \Sigma_\varepsilon \tag{25}$$

$$\mu_{t+1}^x = \mu_t^x + \beta_0^x + \eta_t^x \qquad , \quad Var(\eta_t) = H\,\Sigma_\eta\,H'$$

para k≤p,

de forma que o modelo (25) é indistinguível de (24). Isto significa que há infinitos parâmetros que geram a mesma função densidade para as observações em Y_t. Para tornar o modelo identificável, devem-se impor algumas restrições sobre Σ_η e Θ . Harvey (1989) sugere as seguintes restrições em conjunto:

PUC-Rio - Certificação Digital Nº 0713589/CA

(i) $\sum_{\eta}$ diagonal (26)

(ii) $$\begin{cases} \Theta[i][i] = 1, \ para \ i = 1,2,...,k \\ \Theta[i][j] = 0, \ para \ j > i, \ tal \ que \ j = 1,2,...,p \ e \ i = 1,2,...,k \end{cases}$$

A primeira restrição implica que os fatores comuns são não correlacionados entre si, o que é desejável para efeitos de simplificação da interpretação dos resultados. A segunda restrição tem menos apelo intuitivo e implica que a primeira série depende do primeiro fator comum, mas independe dos outros, a segunda série depende do primeiro e segundo fatores, mas não depende dos outros e assim sucessivamente.

As principais vantagens do modelo estão na sua parcimônia, uma vez que tenta-se encontrar apenas k (k≤p) fatores para explicar o comportamento das p tendências das variáveis abarcadas em Y_t e na facilidade de interpretação do modelo. Esta advém da possibilidade do modelo ser multiplicado por uma matriz ortogonal R construída convenientemente e dita matriz de rotação dos fatores.

3.4
Modelos de Fatores Comuns para Volatilidade

Trata-se de uma simples adaptação do modelo anterior para modelos de volatilidade estocástica. Em Harvey, Ruiz e Shephard (1994) modelou-se a extração de fatores comuns entre as volatilidades das seguintes taxas de câmbio: marco/dólar, yen/dólar, libra/dólar e franco/dólar. Para tanto, supôs-se que a fator latente associado à volatilidade (h_t) seguia um processo de passeio aleatório, como mostrado a seguir:

$$\omega_{t\,px1} = -1{,}27_{px1} + \Theta_{pxk} h_{t\,kx1} + \bar{h}_{kx1} + \xi_{t\,kx1} \qquad \xi_{t\,pxp} \sim (0_{px1}, \Sigma_{\xi\,pxp}) \qquad (27)$$

$$h_{t\,kxk} = h_{t-1 kx1} + \eta_{t\,kx1}, \qquad \eta_{t\,kxk} \sim (0_{kx1}, \Sigma_{\eta\,kxk})$$

Onde:

(i) R_t são os retornos das taxas de câmbio e seguem o seguinte processo

$$R_t = \left\{ [\exp(h_t)]^{1/2} \right\} * \varepsilon_t \ , \ \ \varepsilon_t \sim (0,1)$$

PUC-Rio - Certificação Digital Nº 0713589/CA

(ii) $\omega_t = \log(R_t^2)$,

(iii) $\xi_t = \log(\varepsilon_t^2) + 1{,}27$,

(iv) $\bar{h}$ tem as primeiras k coordenadas nulas ,

(v) os elementos da diagonal da matriz Σ_ξ são $\frac{\pi^2}{2}$,

Assim como o modelo (25), o modelo (27) é não identificável, sendo necessária a imposição de restrições para estimá-lo. Em Harvey, Ruiz e Shephard (1994) as restrições sugeridas são:

$$\text{(r.i)} \quad \Theta[i][j] = 0 \ se \ j > i, para \ \mathrm{j} = 1,2,...,\mathrm{p} \quad \mathrm{i} = 1,2,...,k \tag{29}$$

$$\text{(r.ii)} \quad \Sigma_\eta = I_{kxk}$$

A grande questão, no entanto, ainda não foi abordada e trata-se de como obter o número correto de fatores comuns entre as volatilidades. A sugestão de Harvey, Ruiz e Shephard (1994) consiste em primeiramente estimar o modelo clássico de volatilidade estocástica multivariado (20), supondo que tanto a equação das medidas quanto a do estado seguem passeios aleatórios sem intercepto. Isto equivale a estimar o modelo (24) com o máximo de fatores comuns possíveis, isto é, k=p, e supondo $\Theta = I$. Observe que sob esta restrição o modelo é identificável.

Após a estimação da matriz de covariância do erro η_t, denotada por Σ_η, é feita a Análise de Componentes Principais (ACP) sobre $\hat{\Sigma}_\eta$, um método cuja finalidade é a redução da dimensionalidade do espaço dos componentes com o intuito de facilitar a interpretação do modelo. Os fatores são, em resumo, os autovetores da matriz de covariância estimada, multiplicados pelas respectivas raízes quadradas de seus autovalores. Ranqueia-se os fatores de forma decrescente com os autovalores a que estão associados. Fica a critério do analista, o número de fatores a serem escolhidos. No artigo citado, escolheu-se o menor número de fatores capaz de explicar pelo menos 95% da variação do sistema. Tal critério

PUC-Rio - Certificação Digital Nº 0713589/CA

resultou na escolha de dois fatores comuns. Uma vez determinado o número de fatores comuns (k), estima-se novamente o modelo (24), com as matrizes de coeficiente, impondo-se as devidas restrições para tornar o modelo identificável.

PUC-Rio - Certificação Digital Nº 0713589/CA