

2

Revisão Teórica: Métodos de volatilidades e covariâncias

Nesta seção serão brevemente discutidos os métodos utilizados neste trabalho para a estimação de variâncias e covariâncias. Esta seção encontra-se subdividida em três famílias de métodos: metodologia do *Riskmetrics*, modelos da família GARCH e modelos de ‘volatilidade realizada’.

2.1.

Metodologia do *Riskmetrics* (J.P. Morgan (1997))

Uma primeira tentativa para a estimação de variâncias e covariâncias seria através de uma janela móvel, onde nesta janela os estimadores tradicionais de variâncias e covariâncias seriam utilizados. Volatilidades e covariâncias no dia T seriam então estimadas por:

$$\sigma_{i,T}^2 = \frac{1}{n} \sum_{t=T-n}^{t=T-1} r_{i,t}^2 \text{ e } \sigma_{ij,T} = \frac{1}{n} \sum_{t=T-n}^{t=T-1} r_{i,t} r_{j,t} \quad (3)$$

onde: r_i e r_j representam os retornos (logarítmicos) de dois ativos, $\sigma_{i,T}^2$ a variância do ativo i no dia T , $\sigma_{ij,T}$ a covariância entre os retornos dos ativos i e j no dia T e n é o tamanho da janela.

Esses estimadores são não-viesados e consistentes em se tratando de variâncias e covariâncias incondicionais¹. No entanto, não fica claro sobre quais hipóteses estes estimadores seriam não-viesados e consistentes em se tratando de variâncias e covariâncias condicionais, o real objeto de estudo.

Em especial, note que todas as observações recebem o mesmo peso neste tipo de estimador e, apenas um retorno aberrante, irá manter as volatilidades e covariâncias altas por um período de exatamente n dias, embora a volatilidade do

¹ Conforme ressalta Alexander (1998), não existe evidência empírica de que o fato de assumir média zero para o retorno degrada a qualidade das variâncias e covariâncias estimadas.

ativo já possa ter retornado para níveis normais há algum tempo. Uma forma natural de corrigir este problema seria através da atribuição de pesos maiores para os dias mais recentes, como nas equações abaixo:

$$\sigma_{i,T}^2 = \frac{\sum_{t=T-n}^{t=T-1} \lambda^{T-1-t} r_{i,t}^2}{\sum_{t=T-n}^{t=T-1} \lambda^{T-1-t}} \text{ e } \sigma_{ij,T} = \frac{\sum_{t=T-n}^{t=T-1} \lambda^{T-1-t} r_{i,t} r_{j,t}}{\sum_{t=T-n}^{t=T-1} \lambda^{T-1-t}}, \quad (4)$$

onde λ é um parâmetro que determinará o peso das observações recentes em relação as observações mais antigas (quanto menor λ maior o peso das observações recentes). A introdução do fator de decaimento λ permite não somente que retornos aberrantes tenham seu efeito sobre a volatilidade diminuindo de forma gradativa, como também permite maior velocidade de reação de volatilidades e covariâncias a choques no mercado.

Esta metodologia representada pelas equações em (4) foi defendida pelo banco J.P.Morgan (1997) no seu influente manual do *Riskmetrics* e, devido ao decaimento exponencial nos pesos atribuídos as observações anteriores, ficou conhecida como EWMA, sigla do termo em inglês *Exponential Weighted Moving Average*.

A utilização do método EWMA difundiu-se largamente, em parte devido a sua simplicidade e em parte devido aos bons resultados obtidos por esta metodologia em comparação com as demais², e por isso servirá como *benchmark* neste trabalho. Para a utilização deste método será utilizado o valor de 0.94 para λ , conforme sugerido no manual do *Riskmetrics*³ e, quanto ao tamanho da janela n , serão utilizadas as equações presentes em (4) quando n tende ao infinito⁴:

² Para comparações entre diferentes métodos de extração de volatilidades e correlações veja, por exemplo, Lopez e Walter (2001), Lund e Hansen (2001) e Pagan e Schwert (1990).

³ Considere λ_i o valor do fator de decaimento que minimiza o erro quadrático médio de previsão da variância da série i . O fator ótimo de decaimento é obtido através da utilização de uma média ponderada dos λ_i , utilizando-se um conjunto de 480 séries temporais de ativos americanos. Ver J.P.Morgan (1997). Optou-se neste trabalho por não se utilizar tais métodos de otimização do parâmetro λ uma vez que, conforme já ressaltado, o método EWMA é utilizado apenas como *benchmark* para os demais métodos.

⁴ Existem duas motivações para se utilizar n tendendo ao infinito: a primeira devido a constatação empírica de que as volatilidades e covariâncias possuem memória muito longa (veja Ding, Granger e Engle (1993)) e a segunda, de menor importância, devido a facilidade computacional da fórmula recursiva em (6).

$$\sigma_{i,T}^2 = (1-\lambda) \sum_{t=-\infty}^{T-1} \lambda^{T-1-t} r_{i,t}^2 \text{ e } \sigma_{ij,T} = (1-\lambda) \sum_{t=-\infty}^{T-1} \lambda^{T-1-t} r_{i,t} r_{j,t} \quad (5)$$

As equações em (5) podem ser reescritas na forma recursiva como:

$$\sigma_{i,t}^2 = (1-\lambda)r_{i,t-1}^2 + \lambda\sigma_{i,t-1}^2 \text{ e } \sigma_{ij,t} = (1-\lambda)r_{i,t-1}r_{j,t-1} + \lambda\sigma_{ij,t-1} \quad (6)$$

Garante-se que as matrizes de covariância condicionais obtidas desta forma são positivas definidas uma vez que se utilizou o mesmo valor de λ para todas as séries (ver J.P.Morgan (1997)).

2.2. Modelos da Família GARCH

Desde os seminais artigos de Engle(1982) e Bollerslev(1986) os modelos da família GARCH (*Generalized Autoregressive Conditional Heteroscedasticity*) difundiram-se como uma metodologia simples e flexível para a estimação e previsão de volatilidades. Em particular, o modelo mais parcimonioso GARCH(1,1) foi amplamente utilizado na modelagem de variâncias de séries financeiras:

$$r_{i,t} = \sigma_{i,t} \varepsilon_t \quad (7)$$

$$\sigma_{i,t}^2 = \omega + \alpha r_{i,t-1}^2 + \beta \sigma_{i,t-1}^2 \quad (8)$$

onde: $r_{i,t}$ é o retorno do ativo i no instante t ⁵, $\sigma_{i,t}$ é o desvio-padrão do ativo i no instante t , ε_t é um ruído branco (geralmente supõe-se que ε_t é Gaussiano) e ω, α e β são parâmetros.

Neste modelo, impõe-se que os parâmetros α e β sejam positivos e que o parâmetro ω seja estritamente positivo. As condições de positividade dos

parâmetros garantem que as variâncias condicionais também sejam positivas e ω deve ser estritamente positivo para garantir que o processo para $r_{i,t}$ não degenere quando $t \rightarrow \infty$. Além disso, se $\alpha + \beta < 1$, a variância incondicional de $r_{i,t}$ é finita e o processo é estacionário de 2ª ordem. E mais, no caso de ε_t Gaussiano, Nelson (1990) mostra que $r_{i,t}$ é *estritamente estacionário* mesmo quando α e β somam 1⁶. Ou seja, quando $\alpha + \beta = 1$, o processo GARCH (1,1) é estritamente estacionário, embora sua variância incondicional não seja finita (note que o EWMA é um caso particular do GARCH (1,1) quando $\omega=0$ e $\alpha + \beta = 1$).

A estrutura para a variância condicional do GARCH (1,1) (equação (8)) permite que o modelo capture o já citado fato estilizado da existência de *clustering* de volatilidades, além de possibilitar a geração de distribuições leptocúrticas para as séries de retornos dos ativos. Este excesso de curtose pode ser obtido de duas diferentes formas: primeiramente devido a estrutura dinâmica da variância condicional e, além disso, diretamente através da atribuição de distribuições com caudas grossas para ε_t ⁷.

No entanto, séries de retornos financeiros são também caracterizadas por outros fatos estilizados, como uma autocorrelação de 1ª ordem nos quadrados dos retornos baixa, mas com alta persistência para as defasagens seguintes⁸. E neste quesito, como ressaltam Teräsvirta (1996) e Carnero, Peña e Ruiz (2001), os modelos GARCH não são suficientemente robustos, sendo incapazes de acomodar simultaneamente os valores típicos de excesso de curtose e de autocorrelação nos quadrados dos retornos presentes em séries financeiras (mesmo quando ε_t possui uma distribuição de cauda grossa como uma *t* de *student*).

Se por um lado o modelo GARCH (1,1) não consegue capturar completamente as regularidades empíricas de séries financeiras descritas acima, por outro lado inúmeras extensões já foram desenvolvidas para capturar outros fatos estilizados como o efeito assimétrico de choques positivos e negativos sobre

⁵ A equação para a média condicional do retorno como em (7) pode também conter variáveis explicativas, no entanto, tal prática não é comum em séries financeiras.

⁶ Nelson (1990) mostra que o GARCH (1,1) é estacionário desde que $E[\log(\beta + \alpha \varepsilon_t^2)] < 1$, condição esta que no caso de ε_t Gaussiano é satisfeita mesmo quando $\alpha + \beta = 1$.

⁷ Para modelos GARCH com distribuições condicionais com caudas grossas veja Engle e Gonzalez-Rivera (1991) e Hansen (1994).

a variância ou até mesmo mudanças de regime nas volatilidades⁹. Porém, e mais importante, no que tange ao progresso da modelagem univariada em direção a uma abordagem multivariada, os avanços foram de certa forma mais discretos. E foram mais discretos não somente devido ao agravamento das dificuldades na hora de capturar regularidades empíricas de volatilidades em conjunto com covariâncias, mas principalmente graças a problemas de ordem computacional.

Para tornar a discussão mais clara considere uma extensão natural do modelo GARCH(1,1) para um contexto de dois ativos, onde se modela a variância de cada um dos ativos e a covariância entre eles:

$$\sigma_{i,t}^2 = \omega_1 + \alpha_1 r_{i,t-1}^2 + \alpha_2 r_{j,t-1}^2 + \alpha_3 r_{i,t-1} r_{j,t-1} + \beta_1 \sigma_{i,t-1}^2 + \beta_2 \sigma_{j,t-1}^2 + \beta_3 \sigma_{ij,t-1} \quad (9)$$

$$\sigma_{j,t}^2 = \omega_2 + \alpha_4 r_{i,t-1}^2 + \alpha_5 r_{j,t-1}^2 + \alpha_6 r_{i,t-1} r_{j,t-1} + \beta_4 \sigma_{i,t-1}^2 + \beta_5 \sigma_{j,t-1}^2 + \beta_6 \sigma_{ij,t-1} \quad (10)$$

$$\sigma_{ij,t} = \omega_3 + \alpha_7 r_{i,t-1}^2 + \alpha_8 r_{j,t-1}^2 + \alpha_9 r_{i,t-1} r_{j,t-1} + \beta_7 \sigma_{i,t-1}^2 + \beta_8 \sigma_{j,t-1}^2 + \beta_9 \sigma_{ij,t-1} \quad (11)$$

O modelo representado pelas equações (9), (10) e (11) foi denominado vech e apresentado em Engle e Kroner (1993). A primeira dificuldade prática evidente é proveniente do elevado número de parâmetros: temos 21 parâmetros a serem estimados em um modelo com apenas dois ativos e, por exemplo, já com 5 ativos são 465 os parâmetros a serem estimados. Como esses modelos são usualmente estimados por máxima verossimilhança, os problemas são comuns nas rotinas numéricas de otimização e, com um número um pouco maior de ativos torna-se praticamente inviável. O segundo problema prático advém das dificuldades de se impor e verificar restrições sobre os parâmetros que garantam que a matriz de covariância seja quase certamente positiva definida.

Outras formulações para os modelos GARCH multivariados foram desenvolvidas, formulações estas que restringiam o espaço dos parâmetros de forma a diminuir a sua dimensionalidade e, além disso, simplificar as condições para a geração de matrizes de covariância positivas definidas. Exemplos destas formulações são o modelo BEKK de Engle e Kroner (1993) e o modelo vech

⁹ Ver Loudon, Watt e Yadav (2000).

diagonal de Bollerslev, Engle e Wooldridge (1988). No entanto, mesmo estes modelos sofrem de dificuldades operacionais em um ambiente com um grande número de ativos, como usualmente é relevante na prática.

Dada estas restrições serão utilizados neste trabalho dois modelos multivariados da família GARCH que possuem uma característica em comum: a estimação do modelo multivariado é quebrada em uma série de estimações univariadas. Tais modelos são o *Constant Correlation* GARCH de Bollerslev (1990) e o GARCH Ortogonal de Alexander (2000)¹⁰.

2.2.1.

Constant Correlation GARCH (Bollerslev (1990))

No *Constant Correlation* GARCH (CCOR) as covariâncias condicionais são parametrizadas para serem proporcionais aos respectivos desvios-padrão, restringindo assim o número de parâmetros e a carga computacional necessária para a estimação. Além disto, tal parametrização, facilita a imposição de condições para a geração de matrizes de covariância quase certamente positivas definidas.

Para tornar a discussão mais precisa suponha um ambiente com N ativos. Então, através da hipótese de correlação condicional constante (ou equivalentemente, covariâncias condicionais proporcionais aos respectivos desvios-padrão), o modelo CCOR pode ser escrito como:

$$r_{i,t} = \sigma_{i,t} \varepsilon_t \quad (12)$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i r_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \quad i = 1, \dots, N \quad (13)$$

$$\sigma_{ij,t} = \rho_{ij} \left(\sqrt{\sigma_{i,t}^2} \sqrt{\sigma_{j,t}^2} \right) \quad i \neq j \quad (14)$$

ou, as equações (13) e (14) podem ser escritas na forma matricial como:

⁹ Para uma revisão veja Bollerslev, Engle e Nelson (1994).

¹⁰ Note que como serão utilizados neste trabalho apenas 5 ativos, modelos como o BEKK são viáveis. No entanto, optou-se por trabalhar apenas com o *Constant Correlation* GARCH e o

$$H_t = D_t \Gamma D_t \quad (15)$$

onde: ρ_{ij} é a correlação condicional entre os ativos i e j que é suposta constante ao longo do tempo, H_t é a matriz de covariância condicional em t , D_t é uma matriz diagonal cujos elementos são os desvios-padrão condicionais ($\sigma_{i,t}$) e Γ é uma matriz com elemento típico ρ_{ij} .

O número de parâmetros a serem estimados diminui significativamente (por exemplo, quando N igual a 2 são apenas 7 parâmetros a serem estimados em contraposição ao modelo vech que totaliza 21 parâmetros) e, mais importante, com a hipótese de retornos distribuídos de acordo com uma distribuição normal multivariada, a estimação do modelo multivariado é quebrada em várias estimativas univariadas: estimam-se primeiramente os modelos GARCH(1,1) univariados em (13) obtendo-se os resíduos $\varepsilon_{i,t}$, de posse destes resíduos calculam-se as suas correlações incondicionais que serão as estimativas das correlações condicionais ρ_{ij} e, por último, através da equação (14) recuperam-se as covariâncias condicionais. Além disso, da equação (15), verifica-se que a matriz de covariância será positiva definida se, e somente se, a matriz Γ for positiva definida, mas esta condição é quase certamente garantida, dado que os elementos de Γ são por construção as correlações incondicionais dos resíduos.

Apesar das vantagens computacionais, obviamente a validade da hipótese de correlações condicionais constantes é uma questão empírica, que pode ser inadequada em alguns casos. Para um teste formal desta hipótese veja Bera e Roh (1991).

2.2.2. GARCH Ortogonal (Alexander (2000))

Uma outra possibilidade para contornar o excessivo número de parâmetros envolvidos em modelos GARCH multivariados seria através do uso de fatores que seriam responsáveis pelos movimentos de toda matriz de covariância. Tal abordagem ficou conhecida como *factor* GARCH e foi introduzida por Engle, Ng

GARCH Ortogonal, devido a efetiva possibilidade de utilização destes modelos em um contexto de dimensionalidade muito maior do que a presente.

e Rothschild (1990), que utilizaram o retorno de mercado como fator que origina os movimentos de toda a matriz de covariância.

O modelo GARCH Ortogonal de Alexander (2000) insere-se neste contexto de modelos *factor* GARCH, tendo como idéia principal o uso de fatores ortogonais que norteiam os movimentos da matriz de covariância. Estes fatores são obtidos através do uso da metodologia de componentes principais e, por serem ortogonais, podem ser modelados individualmente através de modelos GARCH univariados.

Sejam n ativos cada qual com t observações e considere os retornos (logarítmicos) destes ativos armazenados na matriz R_t ($t \times n$). Então a matriz de componentes principais P_t ($t \times n$) é definida como:

$$P_t = R_t W_t \quad (16)$$

onde: W_t ($n \times n$) é a matriz de autovetores de $R_t^T R_t$ e como tal é uma matriz ortonormal.

É imediato modelar as variações da matriz de retornos originais R_t a partir de variações da matriz de componentes principais P_t :

$$R_t = P_t W_t^T, \quad \text{pois } W_t^{-1} = W_t^T \text{ (} W_t \text{ é ortonormal)} \quad (17)$$

A representação da matriz de retornos a partir da equação (17) é vantajosa, pois a matriz de componentes principais é uma matriz ortogonal¹¹. Desta forma, a matriz de covariância dos retornos pode ser escrita através da matriz de pesos W_t ¹² e da matriz de covariância de P_t , que por ser uma matriz ortogonal, envolve apenas a modelagem de variâncias de forma univariada:

$$\text{Var}(R_t) = \text{Var}(P_t W_t^T) = W_t \text{Var}(P_t) W_t^T = W_t D_t W_t^T \quad (18)$$

¹¹ Para verificar que a matriz P_t é ortogonal parta da tradicional fórmula de diagonalização de matrizes simétricas:

$R_t^T R_t = W_t \Lambda_t W_t^T \Rightarrow R_t^T R_t W_t = W_t \Lambda_t$, onde Λ_t é a matriz de autovalores de $R_t^T R_t$.

Logo, $P_t^T P_t = W_t^T R_t^T R_t W_t = W_t^T W_t \Lambda_t = \Lambda_t$, pois W_t é ortonormal.

¹² Na prática, conforme ressalta Alexander (2000), a matriz de pesos W_t não varia muito ao longo do tempo, podendo ser considerada como constante e diminuindo ainda mais a carga computacional.

onde D_t é uma matriz diagonal contendo as variâncias em t da matriz de componentes principais P_t .

Em outras palavras, a matriz de covariância dos retornos está sendo modelada a partir dos fatores representados pelas colunas da matriz de componentes principais P_t . Como tais fatores são ortogonais, para a estimação de suas matrizes de covariância são necessários apenas métodos univariados da família GARCH¹³. A partir desta estimação das variâncias condicionais dos fatores, retorna-se para a estimação de variâncias e covariâncias de toda a matriz de retornos, através da utilização da equação (18).

O GARCH Ortogonal obtém simplicidade computacional ao reduzir a estimação multivariada em uma série de estimações univariadas e, além disso, contorna o problema de geração de matrizes de covariância positivas definidas (como D_t é uma matriz diagonal com elementos positivos, por (18) resulta que $\text{Var}(R_t)$ será uma matriz positiva definida). No entanto, assim como os demais modelos GARCH multivariados, é baseado em algumas hipóteses simplificadoras: assumir em (18) que a matriz de covariância *condicional* dos componentes principais é uma matriz diagonal, constitui-se em uma hipótese, pois por construção, os componentes principais são apenas *incondicionalmente* não correlacionados.

2.3.

Volatilidade Realizada (Anderson et al. (2003))

A metodologia de ‘volatilidade realizada’ é computacionalmente trivial e possui um mecanismo direto: estimam-se as volatilidades (covariâncias) diárias através da soma de retornos ao quadrado intradiários (produto cruzado dos retornos intradiários). Apesar da simplicidade, trata-se de um estimador não-paramétrico que dispensa uma dinâmica específica sobre a matriz de covariância (como, por exemplo, a dinâmica em (6)), e mais importante, sob condições gerais, trata-se de um estimador não-viesado e consistente de volatilidades e covariâncias.

¹³ Neste trabalho serão utilizados dois métodos univariados para a estimação de variâncias das componentes principais: o GARCH (1,1) e o GARCH Exponencial (EGARCH) de Nelson (1991). Quando o GARCH(1,1) for utilizado em P_t , será mantido o nome da metodologia desta seção em GARCH Ortogonal (OGARCH), mas quando o EGARCH for utilizado em P_t , a metodologia desta seção será denominada EGARCH Ortogonal (OEGARCH).

Baseado em Oomen (2001) será apresentado um modelo simples em tempo discreto que possui duplo propósito: fixar a intuição e demonstrar as dificuldades práticas relacionadas a implementação do estimador de “volatilidade realizada”. Para uma versão em tempo contínuo onde é suposto que o processo de preços segue um semi-martingal especial (se o processo de preços não admite arbitragem ele encontra-se nesta classe) veja Anderson, Bollerslev, Diebold e Labys (2003).

Seja $R_{i,D}$ ¹⁴ o retorno (logarítmico) do ativo i no dia D e $R_D = [R_{1,D}, R_{2,D}, \dots, R_{N,D}]$ o vetor de retornos dos N ativos no dia D . Será suposto que o vetor de retornos pode ser caracterizado por:

$$R_D \sim MVN(0, \Sigma_D)$$

$$\Sigma_D = \begin{pmatrix} \sigma_{1,D}^2 & \sigma_{12,D} & \cdots & \sigma_{1N,D} \\ \vdots & \ddots & & \vdots \\ \sigma_{1N,D} & \cdots & \sigma_{N,D}^2 \end{pmatrix} \quad (19)$$

onde: $MVN(0, \Sigma_D)$ indica a distribuição normal multivariada de vetor de média nulo e matriz de covariância Σ_D .

Um primeiro estimador para variâncias e covariâncias diárias comumente utilizado na literatura consiste simplesmente em utilizar o retorno diário ao quadrado como um estimador da variância diária e o produto dos retornos diários como um estimador das covariâncias diárias. Apesar de não-viesado, tal estimador apresenta claramente muito ruído, por utilizar apenas um dado para a estimação:

$$E(R_{i,D}^2) = \sigma_{i,D}^2 \text{ e } \text{Var}(R_{i,D}^2) = 2\sigma_{i,D}^4$$

$$E(R_{i,D}R_{j,D}) = \sigma_{ij,D} \text{ e } \text{Var}(R_{i,D}R_{j,D}) = \sigma_{ij,D}^2 + \sigma_{i,D}^2\sigma_{j,D}^2 \quad (20)$$

O ruído inerente ao estimador acima poderia ser contornado através da utilização das informações contidas em dados intradiários. Para tanto, seguindo o modelo delineado para os retornos diários, considere m retornos intradiário $r_{i,t}$ espaçados igualmente ao longo do dia (para simplificar a notação normalize o tamanho de um dia para 1):

¹⁴ Para evitar considerações sobre a média que complicariam a notação e não alterariam a essência da discussão abaixo, considere o retorno como o retorno *em excesso*.

$$r_t \sim \text{MVN}(0, \Sigma_t), \quad r_t = [r_{1,t}, r_{2,t}, \dots, r_{N,t}] \quad t = 1/m, 2/m, \dots, 1$$

$$\Sigma_t = \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} & \cdots & \sigma_{1N,t} \\ \vdots & & \ddots & \vdots \\ \sigma_{1N,t} & \cdots & & \sigma_{N,t}^2 \end{pmatrix} \quad (21)$$

Com a suposição que os retornos intradiários são temporalmente desconcorrelatados, e dado que o retorno diário é a soma dos retornos intradiários (lembre-se que os retornos são logarítmicos), é direto obter as seguintes relações que ligam variâncias e covariâncias diárias a variâncias e covariâncias intradiárias:

$$R_{i,D} = \sum_{t=1/m}^1 r_{i,t} \Rightarrow \sigma_{i,D}^2 = \sum_{t=1/m}^1 \sigma_{i,t}^2 \text{ e } \sigma_{ij,D} = \sum_{t=1/m}^1 \sigma_{ij,t} \quad (22)$$

Da expressão (22) vemos que a variância diária é a soma das variâncias intradiárias e, de forma similar, a covariância diária é a soma das covariâncias intradiárias. Cada variância intradiária poderia ser estimada através do respectivo retorno intradiário ao quadrado e então obteríamos o seguinte estimador para a variância diária: $\sum_{t=1/m}^1 r_{i,t}^2$. Analogamente cada covariância intradiária poderia ser estimada pelo produto dos respectivos produtos intradiários e o estimador da covariância diária seria dado por $\sum_{t=1/m}^1 r_{i,t} r_{j,t}$. A estes estimadores atribuiu-se o nome de estimadores de ‘volatilidade realizada’.

Dado que os retornos intradiários são temporalmente não correlacionados, o estimador de “volatilidade realizada” será não-viesado. A demonstração é imediata:

$$R_{i,D}^2 = \left(\sum_{t=1/m}^1 r_{i,t} \right)^2 = \sum_{t=1/m}^1 r_{i,t}^2 + 2 \sum_{t_1=1/m}^{1-1/m} \sum_{t_2=t_1+1/m}^1 r_{i,t_1} r_{i,t_2} \quad (23)$$

$\sigma_{i,D}^2 = E(R_{i,D}^2) = E\left(\sum_{t=1/m}^1 r_{i,t}^2\right)$, dado que os retornos são temporalmente não correlacionados.

(a demonstração é análoga para o caso da covariância, sendo que enquanto na variância é necessária a hipótese de que o retorno intradiário do ativo i em um instante do tempo é não correlacionado com os retornos intradiários do ativo i em outros instantes, na covariância usa-se que o retorno do ativo i é não correlacionado com os retornos dos ativos $j \neq i$ em outros instantes do tempo).

Apresentam-se então dois estimadores não-tendenciosos para variâncias diárias, a saber, o quadrado do retorno diário e a soma dos quadrados dos retornos intradiários. Da mesma forma, apresentam-se dois estimadores para covariâncias diárias: o produto dos retornos diários e a soma dos produtos dos retornos intradiários. No entanto, ao explorar a informação contida nos dados intradiários, o estimador de “volatilidade realizada” consegue diminuir o ruído da estimação. Veja o caso da estimação das variâncias:

$$Var\left(\sum_{t=1/m}^1 r_{i,t}^2\right) = 2 \sum_{t=1/m}^1 \sigma_{i,t}^4 < 2 \left(\sum_{t=1/m}^1 \sigma_{i,t}^2\right)^2 = 2(\sigma_{i,D}^2)^2 = Var(R_{i,D}^2) \quad (24)$$

A intuição para o resultado acima é a seguinte: o estimador de “volatilidade realizada” obtém a variância diária através da estimação de cada variância intradiária pelo respectivo retorno intradiário ao quadrado e cada uma destas estimações possui ruído, no entanto, como estes estimadores são independentes, estes ruídos tenderão a se cancelar. No limite, quando o número de retornos intradiários $m \rightarrow \infty$, não existiria mais ruído na estimação da variância e o estimador de “volatilidade realizada” seria consistente. Para fixar idéias dentro do contexto desta seção¹⁵ suponha que as variâncias e covariâncias intradiárias são proporcionais às variâncias e covariâncias diárias, ou seja:

$$\sigma_{i,t}^2 = \sigma_{i,D}^2 / m \text{ e } \sigma_{ij,t} = \sigma_{ij,D} / m \quad (25)$$

Então teríamos o seguinte:

¹⁵ Para uma demonstração formal da consistência do estimador de “volatilidade realizada” veja Anderson, Bollerslen, Diebold e Labys (2003)

$$\begin{aligned} \text{Var}\left(\sum_{t=1/m}^1 r_{i,t}^2\right) &= 2 \sum_{t=1/m}^1 \sigma_{i,t}^4 = 2 \sigma_{i,D}^4 / m \\ \text{Var}\left(\sum_{t=1/m}^1 r_{i,t} r_{j,t}\right) &= \sum_{t=1/m}^1 \sigma_{ij,t}^2 + \sum_{t=1/m}^1 \sigma_{i,t}^2 \sigma_{j,t}^2 = \sigma_{ij,D}^2 / m + \sigma_{i,D}^2 \sigma_{j,D}^2 / m \end{aligned} \quad (26)$$

E quando $m \rightarrow \infty$, as variâncias dos estimadores de ‘volatilidade realizada’ iriam a zero, ou seja, os estimadores seriam consistentes em erro quadrático médio.

O estimador de “volatilidade realizada” apresenta-se então como um estimador não-viesado de variâncias e covariâncias e, teoricamente, na medida que retornos intradiários fossem amostrados em frequências cada vez maiores, variâncias e covariâncias poderiam ser estimadas com um grau de precisão arbitrariamente grande¹⁶. No entanto, na prática existe um *trade-off* entre eficiência e viés ao se determinar a frequência de amostragem dos retornos. Enquanto amostragens em frequências maiores tornam o estimador mais eficiente, elas ao mesmo tempo introduzem viés.

A origem do viés fica clara através da equação (23). Quando os retornos intradiários são correlacionados, o segundo termo da direita da equação não é zero em esperança e, conseqüentemente, o estimador de “volatilidade realizada” torna-se tendencioso. Se o termo $\sum_{t_1=1/m}^{1-1/m} \sum_{t_2=t_1+1/m}^1 r_{i,t_1} r_{i,t_2}$ for tipicamente negativo, o estimador tenderá a superestimar a volatilidade do retorno diário.

Apesar de que no contexto de mercados eficientes possa parecer surpreendente que os retornos intradiários sejam negativamente correlacionados, este é um efeito típico da presença de fricções de microestruturas de mercado¹⁷. Fatores como preços discretos, períodos com ausência de transações e efeitos advindos de limites de “bid-ask”, tendem a gerar correlações nas cotações quando estas são amostradas em janelas muito pequenas. Neste *trade-off* entre a utilização de janelas suficientemente pequenas de forma a não comprometer a precisão das estimativas de volatilidade realizada e, no entanto, grandes o suficiente para mitigar os efeitos advindos de fricções de microestruturas de mercado, optou-se por amostrar os retornos intradiários a cada 15 minutos. Na próxima seção esta escolha será discutida.

¹⁶ Para uma discussão sobre construção de intervalos de confiança para os estimadores de “volatilidade realizada” veja Barndorff-Nielsen e Shephard (2002)

¹⁷ Veja, por exemplo, o tratamento de Campbell, Lo e MacKinlay (1998).