



Instituto Nacional de Matemática Pura e Aplicada

Mestrado em Métodos Matemáticos em Finanças

Estimação de Volatilidade Realizada em
Alta Frequência na Presença de
Microestrutura: Uma Abordagem
Multiescala

por

Henrique Fernandes Macedo

Agosto 2011



Instituto Nacional de Matemática Pura e Aplicada

Estimação de Volatilidade Realizada em Alta
Frequência na Presença de Microestrutura: Uma
Abordagem Multiescala

Henrique Fernandes Macedo

Dissertação apresentada ao Instituto Nacional de Matemática Pura e Aplicada como parte dos requisitos necessários para a conclusão do Mestrado Profissional em Métodos Matemáticos em Finanças.

Orientador: Hugo Alexander de La Cruz Cansino

Co-Orientador: Jorge Passamani Zubelli

Agosto 2011

Dedico esse trabalho aos meus pais, Denise e Sergio, que sempre acreditaram no meu potencial, me incentivaram e apoiaram em todas minhas decisões. Sempre me amaram e não mediram esforços em me dar a melhor educação que estava ao seu conhecimento.

Agradecimentos

Agradeço primeiramente a minha amada Deylla que me acompanhou durante todo o curso, que entendeu toda a minha ausência, ouviu todas minhas queixas e vibrou com cada conquista.

A toda minha família, em especial, meu irmão Fernando Macedo que apesar de odiar matemática sempre me incentivou e torceu a distância. Ao meus avós Geraldo e Geralda Macedo e José Délio e Inalde Fernandes, que apesar de nunca terem cursado uma faculdade souberam reconhecer o valor da educação.

Ao meu co-orientador Jorge Zubelli por ter sugerido o tema, apontado imperfeições e me incentivado incessantemente a me dedicar ao máximo em meu trabalho. Ao meu orientador Hugo de La Cruz, pelas sugestões e toques finais e pela ajuda na preparação da defesa. Ao Professor Max Souza, membro da banca, pelo grande trabalho em Derivativos e pelas sugestões e observações para engrandecer o trabalho.

Agradeço também a todos Professores do mestrado, em especial, Professor Paulo Cezar Carvalho, exemplo de competência técnica e perfeita didática e meu maior incentivador no ambiente acadêmico.

Aos meus amigos do mestrado, Maristela Carvalho, Daniel Tonholo, Felipe Santiago, Bruno Agrelino, Diogo Duarte, Leonardo Marotta, Rafael Gonçalves, Fabio Marco, Artur Xavier pelas frequentes discussões sobre listas e conceitos de finanças.

Aos meus amigos, irmãos da vida, Guilherme Ferreira, Luiz Flávio Moraes, Filipe Gonçalves, Jair Neto, Raphael Guarilha, Arthur Resende, Antonio Miranda, Flavio

Brandão, Igor Grobério, Mussa Haikal, José Kuiaski, Felipe Rufino, Rodrigo Octavio, Carlos Eduardo Dias, Yuri Duarte, Leonardo Spinelli, Felipe Sapucahy, Thiago Nikyel, Gabriela Cavalcante pelo incentivo e por entenderem toda a minha ausência a inúmeros eventos sociais.

Aos meus colegas do BNDES que entenderam minha flexibilidade de horário pelas noites viradas trabalhando em minha dissertação.

A todos os funcionários do IMPA que fazem esta casa ser a maior referência acadêmica brasileira no mundo.

Peço perdão pelo esquecimento de algum nome que tenha sido fundamental para o desenvolvimento e conclusão do mestrado.

Resumo

A recente disponibilidade de dados em alta frequência culminou numa explosão de trabalhos sobre as características de dados em alta frequência. Uma das grandes questões que surgiram destes estudos foi a presença de ruídos originados da própria estrutura das transações que, na literatura, é chamada de microestrutura.

O principal objetivo do trabalho é aplicar o método proposto por (Aït-Sahalia, Mykland & Zhang 2010) para estimar a volatilidade realizada em alta frequência e comparar com a metodologia usual para este tipo de abordagem. No entanto, antes de atingirmos esse ponto, desenvolvemos importantes resultados para o comportamento de estimadores em alta frequência. Além disso, introduzimos a microestrutura no modelo dos retornos observado para entendermos como se comporta este ruído e diminuir seus efeitos na estimação da volatilidade. Isto leva ao uso de um estimador de volatilidade realizada em duas escalas.

Finalmente, realizamos a aplicação desse estimador para as quatro ações mais líquidas da Bovespa no momento em que os dados foram coletados e o comparamos com o método mais usualmente praticado na literatura. Mostramos também uma aplicação do método na estimação do parâmetro do modelo de um fator para hedging, no qual a volatilidade implícita impacta o delta hedging da opção, verificando empiricamente um *leverage effect*.

Palavras-chave: Volatilidade; Dados em Alta Frequência; Microestrutura.

Abstract

The recent availability of high frequency data resulted in boom of papers on the properties of high frequency data. One of the major issues that emerged from such studies is the presence of noise arising from the very structure of transactions. In the literature, it is called microstructure.

The main objective of this study is to explain and to apply the method proposed by (Aït-Sahalia, Mykland & Zhang 2010) to the estimation of the realized volatility in high frequency compared with the usual methodology for this type of approach. However, before we attain this goal, we develop important results for the behavior of high frequency estimators. Moreover, we introduce microstructure in the model of observed returns to understand how this noise behaves and diminishes its effect on the estimation of the volatility. This leads to a two scales realized volatility estimator.

Finally, we apply this estimator to the four most liquid stocks on the Bovespa index at the time the data were collected and compared with the method most commonly practiced in the literature. We also show an application of the method for the estimation of one-factor model used in hedging and for which the implicit volatility impacts the option hedging. This verifies empirically a leverage effect.

Keywords: Volatility; High Frequency Data; Microstructure.

Sumário

Introdução	4
1 Ferramentas Matemáticas	7
1.1 Movimento Browniano	7
1.2 Integral Estocástica	8
1.3 Processo de Itô	11
1.4 Integral Estocástica como Modelo	12
1.5 Tempo de Parada e Martingal Local	13
1.6 Semimartingais	15
1.7 Variação Quadrática de um Semimartingal	16
1.8 Teorema de Lévy	17
1.9 Fórmula de Itô	17
1.10 Continuidade Absoluta	18
1.11 Teorema de Radon-Nikodym	19
1.11.1 Propriedades da Derivada de Radon-Nikodym	20
1.12 A Fórmula de Black & Scholes Clássica	20
1.13 Derivadas Parciais de Black & Scholes: Gregas	21
1.14 O Delta-Hedge	22
1.15 Hipóteses do Modelo de Black-Scholes	23

2	O Comportamento dos Estimadores	25
2.1	Hipótese de Martingal Local	25
2.2	O Processo de Erro	26
2.3	Símbolos de Ordem Estocáticos	26
2.4	Variação Quadrática do Processo de Erro	27
2.4.1	Norma L^p	28
2.4.2	Desigualdade de Burkholder-Davis-Gundy	29
2.4.3	Demonstração de Proposição 2.1	30
2.5	Variação Quadrática do Erro: Observações Independentes	31
2.6	Variação Quadrática do Tempo	37
2.7	Variação Quadrática no Caso Geral	38
2.8	Convergência Estável	39
2.9	Normalidade Assintótica	39
2.10	Aplicação para Volatilidade Realizada	41
2.11	Estimadores Não Centrados	43
2.11.1	Estimação do Modelo Browniano	43
2.11.2	Comportamento dos Estimadores Não Centrados	44
2.12	Livrando-se do <i>drift</i>	45
3	Análise de Volatilidade Realizada sob o Ruído da Microestrutura	47
3.1	O Modelo dos Retornos Observados	48
3.2	A Volatilidade Realizada	52
3.3	Possíveis Estimadores	54
3.3.1	Amostrando Esparsamente	54
3.3.2	Subamostrando e Calculando a Média	57
3.3.3	O Erro devido ao Ruído	58
3.3.4	O Erro devido à Discretização	60

3.4	Combinando os Dois Erros	62
3.5	O Estimador Final: Volatilidade em Duas Escalas de Tempo	63
4	Estimação da Volatilidade Realizada em Duas Escalas em Alta Frequência	66
4.1	Os Dados	67
4.2	Estimação	68
4.2.1	VRDE para PETR4	68
4.2.2	VRDE para VALE5	76
4.2.3	VRDE para OGXP3	83
4.2.4	VRDE para ITUB4	90
5	Aplicação do Estimador de VRDE para Delta Hedge em Modelo de Um Fator	98
5.1	Volatilidade Implícita e Realizada	98
5.2	Modelo de Zero Fator	101
5.2.1	Identidades de Não Arbitragem	103
5.3	Hedgeando a Volatilidade Implícita para Modelo de Um Fator	104
5.4	Estimação do Delta no Modelo de um Fator	107
	Conclusão	124
	Bibliografia	129

Introdução

O aumento do poder computacional somado com a migração das negociações em bolsa para a Internet, permitiu um fácil acesso e processamento de dados que tinham sido pouco utilizados até então. A partir dessa disponibilidade ocorreu um *boom* de trabalhos sobre as aplicações e características de dados em alta frequência. Uma das grandes questões que emergiram destes estudos foi a presença de ruídos advindos da própria estrutura de negociação como a formação do livro de ofertas de compra e venda, falta de liquidez, a própria discretização das transções. Essas especificidades alteram o caminho que o preço seguiria na ausência destas, que na literatura, são chamadas de microestrutura.

Dessa forma, pesquisas sobre a aplicação de estimadores de volatilidade realizada em alta frequência que desconsideram a microestrutura se mostraram extremamente ineficientes em estimar a volatilidade realizada de forma robusta com dados de mercado, indicando uma significância do erro causado pela microestrutura inserida no preço observado. A solução mais utilizada na literatura foi usar janelas de retornos em tempo maiores, escolhidas *ad hoc*, como 5, 10 e 30 minutos, para estimação da volatilidade realizada. No entanto, este tipo de abordagem, ainda que melhor que estimadores com dados diários, acaba por desperdiçar a grande parte dos dados, uma vez que se tem pelo menos uma negociação a cada 5 segundos. Além disso, quando variamos a janela do retorno na estimação esparsa, obtemos valores discrepantes de volatilidade para os mesmos dados, indicando pouca robustez desse tipo de estimador.

A proposta principal do trabalho é explicitar e utilizar o método proposto por (Aït-Sahalia, Mykland & Zhang 2010) para estimar a volatilidade realizada em alta frequência para comparar com a metodologia usual para este tipo de abordagem. Esse estimador utiliza todas observações, não desperdiçando informação, e obtém um viés e uma variância menor que o estimador mais utilizada na literatura, possuindo pouca dependência da janela de tempo utilizada para a estimação. Além disto, é de fácil implementação, podendo rapidamente ser utilizado em condições práticas.

No trabalho mostramos teoricamente porque o estimador proposto tende a ser melhor que o método esparsos. Depois, comprovamos o resultado teórico aplicando ambos estimadores para dados de ações da BOVESPA em condições distintas. Mostramos também o desempenho do método para estimar covariância realizada em alta frequência, em que variáveis aleatórias possuem características diferentes. Por fim, essa covariância é utilizada no modelo proposto por (Zhang 2009) para o cálculo de taxa de delta hedge.

Este trabalho está dividido da seguinte forma: no Capítulo 1 descrevemos os principais resultados da teoria de finanças que serão usados ao longo do trabalho.

No Capítulo 2 desenvolvemos importantes teoremas para o comportamento de estimadores em alta frequência destacando as particularidades ao se trabalhar com este tipo de dados, especialmente, a convergência em distribuição de estimadores de volatilidade não centrados (em que não se estima a média amostral) é a mesma os estimadores centrados, o que permite utilizar o primeiro, na presença de muitas observações como em dados em alta frequência.

No Capítulo 3 iniciamos a abordagem de microestruturas, que não são consideradas no Capítulo 2. Ao introduzir a microestrutura no modelo, mostramos que uma tentativa de estimar a volatilidade com modelo em que o retorno observado do ativo segue um processo de Itô, sem a presença de microestrutura, leva uma estimação viesada e pouca robusta. A partir da identificação do erro associado à estimação, buscamos

construir um estimador melhor que resulta na volatilidade realizada em duas escalas (ver (Zhang, Mykland & Ait-Sahalia 2005)).

No Capítulo 4 aplicamos a volatilidade realizada em duas escalas para as quatro ações mais líquidas da Bovespa no momento da amostragem, PETR4, VALE5, OGXP3 e ITUB4, em um período em que as ações da PETR4 e OGXP3 sofreram um choque de volatilidade, enquanto a VALE5 e a ITUB4 não tiveram nenhum evento incomum, e comparamos com o método de estimação esparsa.

Por fim, no Capítulo 5, utilizamos o estimador de volatilidade realizada em duas escalas para estimar o fator de *leverage effect* proposto por (Zhang 2009), contemplado no modelo de um fator estrutural. Este capítulo tem por objetivo finalizar o trabalho mostrando uma aplicação do estimador em um contexto diferente do seu original de volatilidade, evidenciando sua importância na estimação do parâmetro do modelo de um fator e na verificação empírica de um *leverage effect*.

Capítulo 1

Ferramentas Matemáticas

Neste capítulo abordamos conceitos gerais da teoria de finanças matemáticas que são de extrema importância para o estudo e compreensão das definições e resultados matemáticos percorridos nos capítulos posteriores. Seguiremos de perto as seguintes obras que abrangem o conteúdo: (Karatzas & Shreve 1991), (Jacod & Shiryaev 1987) e (Korn & Korn 2001). A maioria das demonstrações dos resultados deste capítulo não serão apresentadas, mas daremos referências onde podem ser encontradas.

1.1 Movimento Browniano

Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade isto é, Ω é um espaço amostral, \mathcal{F} uma σ -álgebra e \mathbb{P} uma medida de probabilidade definida sobre \mathcal{F} .

Definição 1.1. *Uma filtração é uma família de σ -álgebras $\{\mathcal{F}_t\}_{t \in I}$, onde I é um conjunto ordenado e tal que $\mathcal{F}_s \subset \mathcal{F}_t$ para $s < t$, com $s, t \in I$.*

Definição 1.2. *Um processo estocástico $\{X_t\}_{t \in I}$ adaptado a filtração $\{\mathcal{F}_t\}_{t \in I}$ é uma família de variáveis aleatórias com valores em \mathbb{R}^n tal que $\{X_t\}$ é \mathcal{F}_t -mensurável. Neste caso diz-se que $\{X_t\}$ é um processo adaptado a $\{\mathcal{F}_t\}_{t \in I}$.*

Definição 1.3. Um processo estocástico $\{W_t\}_{t \geq 0}$ com trajetórias contínuas e satisfazendo

i) $W_0 = 0$ \mathbb{P} -q.c.

ii) $W_t - W_s \sim N(0, t - s)$ para $0 \leq s \leq t$ (Incrementos Estacionários),

iii) $W_t - W_s$ independente de $W_u - W_r$ para $0 \leq r \leq u \leq s \leq t$ (Incrementos

Independentes), é chamado de Movimento Browniano unidimensional.

Definição 1.4. Um Movimento Browniano d -dimensional é um processo estocástico com valores \mathbb{R}^d , $\mathbf{W}(t) = (W_1(t), W_2(t), \dots, W_d(t))$, tal que cada componente é um Movimento Browniano unidimensional e W_i é independente de W_j se $i \neq j$.

Definição 1.5. Um processo estocástico $\{X_t$ adaptado a filtração $\mathcal{F}_t\}_{t \in I}$, com $\mathbb{E}|X_t| < \infty$ para $t \in I$, onde I é um conjunto ordenado, é chamado um martingal se $\forall s \leq t$, com $s, t \in I$,

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s \tag{1.1}$$

Teorema 1.1. O movimento Browniano $\{W_t\}_{t \geq 0}$ é um martingal em relação ao seu filtro natural $\mathcal{F}_t^W = \sigma\{W_s | 0 \leq s \leq t\}$, $t \in [0, \infty)$, onde a σ é a σ -álgebra gerada pelos processos $\{W_s\}_{s \geq 0}$, $s \leq t$.

A demonstração deste resultado pode ser encontrada em (Korn & Korn 2001).

1.2 Integral Estocástica

A primeira abordagem para podermos construir uma integral estocástica seria realizar uma integral sob a densidade do movimento Browniano. No entanto, sua distribuição é não diferenciável em todos os caminhos em P -q.c. (ver (Karatzas &

Shreve 1991)). Dessa forma, poderíamos tentar calculá-la no sentido de Lebesgue, se o integrando for uma função contínua. Mas o movimento Browniano admite variação infinita no intervalo $[0, 1]$ P -quase certamente (ver (Karatzas & Shreve 1991)). Dado esses dois resultados negativos, temos que definir um novo tipo de integral. Antes disso, precisamos construir um conceito chamado processo simples X_t .

Definição 1.6. Um processo estocástico $\{X_t\}_{t \in [0, T]}$ é chamado de processo simples se existem número reais $0 = t_0 < \dots < t_p = T, p \in \mathbb{N}$, e variáveis aleatórias $H_i : \Omega \rightarrow \mathbb{R}, i = 1, \dots, p$ com H_0 \mathcal{F}_0 -mensurável e H_i $\mathcal{F}_{t_{i-1}}$ -mensurável com a seguinte representação

$$X_t(\omega) = H_0(\omega)1_{\{0\}}(t) + \sum_{i=1}^p H_i(\omega)1_{(t_{i-1}, t_i]}(t) \quad (1.2)$$

para cada $\omega \in \Omega$.

Definição 1.7. Para um processo simples $X = \{X_t\}_{t \in [0, T]}$ a integral estocástica $I_t(X)$ para $t \in [0, T]$ é dada por

$$I_t(X) := \int_0^t X_s dW_s := \sum_{1 \leq i \leq n} H_i(W_{t_i \wedge t} - W_{t_{i-1} \wedge t}) \quad (1.3)$$

onde $u \wedge t = \min\{u, t\}$.

Definição 1.8. Um processo estocástico é dito progressivamente mensurável se $\forall t \geq 0$ a função

$$\begin{aligned} [0, T] \times \Omega &\longrightarrow \mathbb{R}^n \\ (s, \omega) &\longrightarrow X_s(\omega) \end{aligned} \quad (1.4)$$

é $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ mensurável.

Definição 1.9. Definimos o espaço $L^2[0, T]$ como o conjunto dos processos estocásticos $\{X_t, \mathcal{F}_t\}_{t \in [0, T]}$ tais que $\{X_t\}_{t \in [0, T]}$ é progressivamente mensurável e

$$\mathbb{E} \left(\int_0^T X^2 dt \right) < \infty. \quad (1.5)$$

Definimos ainda a norma de $X = \{X_t\}_{t \in [0, T]}$ em $L^2[0, T]$ como o valor finito dado em (1.5) e denotamos este número por $\|X\|_{L^2[0, T]}^2$.

Teorema 1.2. *Um processo estocástico arbitrário $X \in L^2[0, T]$ pode ser aproximado por uma sequência de processos simples $X^{(n)}$. De forma mais precisa, existe uma sequência $X^{(n)}$ de processos simples com*

$$\lim_{n \rightarrow \infty} \mathbb{E} \int_0^T (X_s - X_s^{(n)})^2 ds = 0. \quad (1.6)$$

A demonstração deste resultado pode ser encontrada em (Korn & Korn 2001).

Teorema 1.3. *Existe uma única aplicação linear J definida em $L^2[0, T]$ e com valores no espaço dos martingais contínuos definidas em $[0, T]$ com respeito a $\{\mathcal{F}_t\}_{t \in [0, T]}$ satisfazendo as seguintes condições:*

1. *se $X = \{X_t\}_{t \in [0, T]}$ é um processo simples então $P(J_t(X) = I_t(X), \forall t \in [0, T]) = 1$,*
2. *vale a isometria de Itô*

$$\mathbb{E}(J_t(X)^2) = \mathbb{E} \left(\int_0^T X_s^2 ds \right). \quad (1.7)$$

Esta aplicação linear é única no sentido que se duas aplicações J e J' satisfazem as condições anteriores então para todo $X \in L^2[0, T]$ os processos $J(X)$ e $J'(X)$ são iguais com probabilidade 1.

A demonstração deste resultado pode ser encontrada em (Korn & Korn 2001).

Definição 1.10. *Para $X \in L^2[0, T]$ e J como no teorema anterior, definimos a integral estocástica do processo X com respeito a $\{W_t\}_{t \in [0, T]}$ como sendo*

$$\int_0^t X_s dW_s := J_t(X) \quad (1.8)$$

Teorema 1.4. Para qualquer $X \in L^2[0, T]$ a integral de Itô

$$\int_0^t X_s dW_s \quad (1.9)$$

é um \mathcal{F}_t -martingal em $[0, T]$. Em particular, a integral dada em (1.8) tem esperança igual a zero.

A demonstração deste resultado pode ser encontrada em (Korn & Korn 2001).

1.3 Processo de Itô

Definição 1.11. X_t é um processo de Itô relativo a (\mathcal{F}_t) , X_t adaptado a (\mathcal{F}_t) ; se há um processo (\mathcal{F}_t) -adaptado (W_t) , e processos (μ_t) e (σ_t) (\mathcal{F}_t) -adaptados

$$\int_0^T |\mu_t| dt < \infty, \text{ e} \quad (1.10)$$

$$\int_0^T \sigma_t^2 dt < \infty, \quad (1.11)$$

então

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s. \quad (1.12)$$

onde W_t é um Browniano padrão, μ_t é o drift e σ_t^2 é a variância instantânea dos retornos do processo X_t .

O processo é frequentemente escrito na forma diferencial:

$$dX_t = \mu_t dt + \sigma_t dW_t. \quad (1.13)$$

Note que a propriedade do processo de Itô é preservada sob a integral estocástica.

Se H_t é limitado e adaptado, então

$$\int_0^t H_s dX_s = \int_0^t H_s \mu_s ds + \int_0^t H_s \sigma_s dW_s \quad (1.14)$$

com

$$\int_0^T |H_t \mu_t| dt < \infty \quad \text{e} \quad (1.15)$$

$$\int_0^T (H_t \sigma_t)^2 dW_t < \infty. \quad (1.16)$$

1.4 Integral Estocástica como Modelo

Ao utilizarmos a Equação (1.12) como modelo, esperamos que esta seja uma ferramenta suficientemente geral para capturar o processo mais relevante. A vantagem de usarmos integrandos adaptados vem da simplicidade de se conectar o modelo com os ganhos de um portfólio.

Suponhamos que X_t é o valor de uma ação. Seja H_t o número de ações que o portfólio possui em t . No caso de um processo simples como a Equação (1.2) significa que carregaremos $H^{(i)}$ unidades de X do instante s_i ao instante t_i . O lucro/perda do portfólio é dado pela integral estocástica da Equação (1.12).

Para μ_t e σ_t^2 simples, a integral

$$\sum_i \mu^{(i)}(t_i - s_i) + \sum_i \sigma^{(i)}(W_{t_i} - W_{s_i}) \quad (1.17)$$

é uma soma de variáveis aleatórias normais, com média $\mu^{(i)}(t_i - s_i)$ e variância $(\sigma^{(i)})^2 (t_i - s_i)$. A soma não precisa ser normal, uma vez que μ e σ^2 podem ser aleatórios.

Vale notar que neste modelo, $\int_0^T \mu_t dt$ é soma das médias instantâneas (*drift*), e $\int_0^T \sigma_t^2 dt$ é soma das variâncias instantâneas. De fato, no modelo da Equação (1.12),

podemos mostrar que: seja $\text{Var}(\cdot|\mathcal{F}_t)$ a variância condicional dada a informação em t . Se X_t é um processo de Itô, e se $0 = t_{n,0} < t_{n,1} < \dots < t_{n,n} = T$, então

$$\sum_i \text{Var}(X_{t_{n,i+1}} - X_{t_{n,i}}|\mathcal{F}_{t_{n,i}}) \xrightarrow{p} \int_0^T \sigma_t^2 dt \quad (1.18)$$

quando

$$\max_i |t_{n,i+1} - t_{n,i}| \rightarrow 0. \quad (1.19)$$

Se μ_t e σ_t^2 são processos não aleatórios, então X_t é um processo Gaussiano, e X_T é normal com média $X_0 + \int_0^T \mu_t dt$ e variância $\int_0^T \sigma_t^2 dt$.

1.5 Tempo de Parada e Martingal Local

O conceito de martingal é melhor entendido considerando esta integral com respeito a um processo de Wiener(veja (Duffie 2001)):

$$X_t = \int_0^t \frac{1}{\sqrt{T-s}} dW_s \quad (1.20)$$

Note que para $0 \leq t \leq T$, X_t é um processo Gaussiano de média zero e incrementos independentes. A integral possui variância

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \int_0^t \frac{1}{T-s} ds \\ &= \int_{T-t}^T \frac{1}{u} du \\ &= \log \frac{T}{T-t}. \end{aligned} \quad (1.21)$$

Como a variância de X_t vai para infinito à medida que nos aproximamos de T , X_t não está definido em T . No entanto, podemos parar o processo em um tempo τ conveniente. Seja $A > 0$,

$$\tau = \inf\{t \geq 0 : X_t = A\}. \quad (1.22)$$

Definindo a integral modificada por

$$Y_t = \int_0^t \frac{1}{\sqrt{T-s}} I\{s < \tau\} dW_s = X_{\tau \wedge t}. \quad (1.23)$$

O processo (1.23) tem a seguinte interpretação em finanças. Suponhamos que W_t tenha o valor de um ativo financeiro no tempo t (este valor pode ser negativo, como por exemplo, contratos futuros). Consideramos também o valor da taxa de juros de curto prazo igual a zero. O processo X_t possui o valor de um portfólio com $\frac{1}{\sqrt{T-s}}$ unidades deste ativo no tempo t . O processo Y_t é obtido carregando o portfólio até $X_t = A$ para liquidar a posição.

Em, outras palavras, montamos uma estratégia de negociação que inicia com riqueza $Y_0 = 0$ em $t = 0$, e termina com a riqueza $Y_T = A > 0$ em $t = T$, ou seja, realizamos uma arbitragem. A abordagem mais comum para esse tipo de problema é impor que a riqueza do agente em qualquer instante não pode ser menor que o montante de $-K$, que é uma restrição de crédito. Apesar disso prevenir a arbitragem, não garante que a integral é um martingal.

Definição 1.12. *Um tempo de parada é uma variável aleatória τ satisfazendo $\{\tau < t\} \in \mathcal{F}_t$ para todo t .*

O que queremos dessa definição é que estejamos aptos a saber no tempo t se τ ocorreu ou não. O tempo (1.22) é um tempo de parada. Por outro lado, a variável $\tau = \inf\{t : W_t = \max_{0 \leq s \leq T} W_s\}$ não é tempo de parada, caso contrário existiria uma ótima estratégia de investimento.

Definição 1.13. O processo M_t é chamado de *martingal local* para $0 \leq t \leq T$ se existe uma sequência de tempos de parada τ_n tal que

1. $M_{\tau_n \wedge t}$ é um martingal para cada n .
2. $\mathbb{P}(\tau_n = T) \rightarrow 1$ quando $n \rightarrow \infty$.

O resultado básico para integrais estocásticas é que a integral com respeito a um martingal local é um martingal local (Jacod & Shiryaev 1987).

1.6 Semimartingais

X_t é um semimartingal se puder ser escrito da forma

$$X_t = X_0 + M_t + A_t, \quad 0 \leq t \leq T, \quad (1.24)$$

onde X_0 é \mathcal{F}_0 -mensurável, M_t é martingal local, e A_t é um processo com variação finita, *i.e.*,

$$\sup \sum_i |X_{t_{i+1}} - X_{t_i}| < \infty, \quad (1.25)$$

onde o supremo é sobre todos os grids $0 = t_0 < \dots < t_n = T$, para todo n .

Em particular, um processo de Itô é um semimartingal, com

$$M_t = \int_0^t \sigma_s dW_s \text{ e}$$

$$A_t = \int_0^t \mu_s ds.$$

Um *supermartingal* é um semimartingal para A_t não crescente. Um *submartingal* é um semimartingal para A_t não decrescente.

1.7 Variação Quadrática de um Semimartingal

Um grid com observações no tempo é dado por

$$\mathcal{G} = \{t_0, t_1, \dots, t_n\}, \quad (1.26)$$

onde supomos que

$$0 = t_0 < t_1 < \dots < t_n = T. \quad (1.27)$$

Seja

$$\Delta(\mathcal{G}) = \max_{1 \leq i \leq n} (t_i - t_{i-1}). \quad (1.28)$$

Para qualquer processo X , definimos *variação quadrática* relativa ao grid \mathcal{G} por

$$[X, X]_t^{\mathcal{G}} = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^2. \quad (1.29)$$

Um importante teorema do cálculo estocástico diz que

Teorema 1.5. *Para qualquer semimartingal, existe um processo $[X, X]_t$ tal que*

$$[X, X]_t^{\mathcal{G}} \xrightarrow{p} [X, X]_t \quad \forall t \in [0, T], \quad \text{quando } \Delta(\mathcal{G}) \rightarrow 0. \quad (1.30)$$

O limite é independente da sequência de grids \mathcal{G} .

Este segue do Teorema I.4.47(p.52) em (Jacod & Shiryaev 1987). Para um processo de Itô,

$$[X, X]_t = \int_0^t \sigma_s^2 ds. \quad (1.31)$$

Ver demonstração em (Jacod & Shiryaev 1987)(p.48 e p.55). O processo $[X, X]_t$ é geralmente referido como a variação quadrática de um semimartingal (X_t) .

1.8 Teorema de Lévy

Um resultado importante na construção que se segue é:

Teorema 1.6. *Seja $(M_t)_{t \geq 0}$ um processo contínuo e adaptado a uma filtração \mathcal{F}_t , com $M_0 = 0$. Suponha que M_t e $(M_t^2 - t)_{t \geq 0}$ são martingais locais, tal que $[M, M]_t = t$. Então M_t é um Browniano com respeito a \mathcal{F}_t . Mais precisamente, se $0 < s < t$, então $M_t - M_s$ é independente de \mathcal{F}_s e possui distribuição normal com média 0 e variância $t - s$.*

A demonstração deste resultado se encontra em (Jacod & Shiryaev 1987)(p.102).

1.9 Fórmula de Itô

Teorema 1.7. *Suponha f uma função duas vezes diferenciáveis, e que X_t é um processo de Itô. Então,*

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d[X, X]_t. \quad (1.32)$$

Similarmente, no caso multivariado, para $X_t = (X_t^{(1)}, \dots, X_t^{(p)})$,

$$df(X_t) = \sum_{i=1}^p \frac{\partial f}{\partial x^{(i)}}(X_t)dX_t^{(i)} + \frac{1}{2} \sum_{i,j=1}^p \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(X_t)d[X^{(i)}, X^{(j)}]_t. \quad (1.33)$$

A referência deste teorema é (Jacod & Shiryaev 1987).

É importante notar que na Equação (1.32) é o mesmo que dizer

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2} \int_0^t f''(X_s)d[X, X]_s. \quad (1.34)$$

Se substituirmos $dX_t = \mu_t dt + \sigma_t dW_t$ e $d[X, X]_t = \sigma_t^2 dt$, a Equação (1.32) se torna

$$\begin{aligned} df(X_t) &= f'(X_t)(\mu_t dt + \sigma_t dW_t) + \frac{1}{2} f''(X_t) \sigma_t^2 dt \\ &= (f'(X_t) + \frac{1}{2} f''(X_t) \sigma_t^2) dt + f'(X_t) \sigma_t dW_t. \end{aligned} \quad (1.35)$$

Notamos, em particular, que se X_t é um processo de Itô, então $f(X_t)$ também o é.

1.10 Continuidade Absoluta

Podemos pensar em ter duas medidas de probabilidade distintas sobre as mesmas observações. Por exemplo, seja P a medida correspondente ao sistema

$$dX_t = \sigma_t dW_t, \quad X_0 = x_0, \quad (1.36)$$

equanto Q corresponde ao sistema

$$dX_t = \mu_t dt + \sigma_t dW_t^Q, \quad X_0 = x_0. \quad (1.37)$$

Neste caso, W_t é um Browniano sob P , e W_t^Q é um Browniano sob Q . Notemos que desde que estejamos modelando o processo X_t , este processo é a quantidade observável da distribuição que buscamos. Então, o processo X_t não se altera de P pra Q , mas

suas distribuições sim. Utilizando as Equações (1.36) e (1.37) temos

$$\mu_t dt + \sigma_t dW_t^Q = \sigma_t dW_t, \quad (1.38)$$

ou

$$\frac{\mu_t}{\sigma_t} dt + dW_t^Q = dW_t. \quad (1.39)$$

No caso em que μ e σ são constantes, quando possuímos observações em um intervalo fixo $[0, T]$, o processo μ_t não pode ser consistentemente estimado, discutiremos este caso particular mais a frente. De fato, μ não pode ser a média observável, pois não podemos distinguir completamente entre P e Q , mesmo com infinitos dados. Este conceito é capturado pela seguinte definição:

Definição 1.14. *Para uma σ -álgebra dada, duas medidas de probabilidades são mutuamente absolutamente contínuas (ou equivalentes) se, para todo $A \in \mathcal{A}$, $P(A) = 0 \iff Q(A) = 0$. De forma mais geral, Q é absolutamente contínua com respeito a P se, para todo $A \in \mathcal{A}$, $P(A) = 0 \implies Q(A) = 0$.*

1.11 Teorema de Radon-Nikodym

Teorema 1.8. *(Radon-Nikodym) Suponha que Q é absolutamente contínua sob P em uma σ -álgebra \mathcal{A} . Então existe uma variável aleatória (\mathcal{A} -mensurável) dQ/dP tal que para todo $A \in \mathcal{A}$,*

$$Q(A) = \mathbb{E}_P \left(\frac{dQ}{dP} I_A \right). \quad (1.40)$$

A demonstração e um teorema mais geral ver (Billingsley 2008).

A quantidade dQ/dP é geralmente chamada de derivada de Radon-Nikodym.

1.11.1 Propriedades da Derivada de Radon-Nikodym

- $P\left(\frac{dQ}{dP} \geq 0\right) = 1$
- Se Q é equivalente a P : $P\left(\frac{dQ}{dP} > 0\right) = 1$
- $\mathbb{E}_P\left(\frac{dQ}{dP}\right) = 1$
- Para todo \mathcal{A} -mensurável Y : $\mathbb{E}_P(Y) = \mathbb{E}_P\left(Y \frac{dQ}{dP}\right)$
- Se Q é equivalente a P : $\frac{dP}{dQ} = \left(\frac{dQ}{dP}\right)^{-1}$

1.12 A Fórmula de Black & Scholes Clássica

Vamos apresentar a fórmula encontrada nos trabalhos de (Black & Scholes 1973) e (Merton 1973) para apreçar o derivativo chamado opção de compra europeia. Uma *opção de compra europeia* é um contrato que permite ao seu detentor o direito, mas não a obrigação de comprar um ativo S por um preço previamente fixado K em uma data especificada T . É fácil notar que o valor da opção em T é $(S_T - K)^+$, onde $x^+ = x$ se $x > 0$, e $x^+ = 0$ caso contrário.

Se assumirmos que S_t é um Browniano que segue a Equação (1.13), tal que $X_t = \log S_t$. Assumimos também que a taxa de juros de curto prazo r é constante (no tempo), então o preço da opção de compra no tempo t , tal que $0 \leq t \leq T$, é

$$\text{preço} = C(S_t, r(t - T), \sigma^2(T - t)), \quad (1.41)$$

onde

$$\begin{aligned} C(S, R, \Xi) &= S\Phi(d_1) - K \exp(-R)\Phi(d_2), \quad \text{onde} \\ d_{1,2} &= (\log(S/K) + R \pm \Xi/2)/\sqrt{\Xi} \quad \text{e} \\ \Phi(x) &= P(N(0, 1) \leq x) \quad \text{a distribuição normal padrão.} \end{aligned} \quad (1.42)$$

Esta é a fórmula de Black-Scholes-Merton.

1.13 Derivadas Parciais de Black & Scholes: Gregas

Uma vez que são definidos as variáveis que influenciam o preço de uma opção, podemos querer calcular como esse preço se altera com a variação nesses parâmetros. Tecnicamente, podemos calcular derivadas parciais desse preço em relações a essas variáveis. Essas derivadas são tão conhecidas que foi denominada uma letra grega para cada derivada. As derivadas mais importantes são: em relação ao ativo objeto, o Delta (símbolo Δ); em relação à volatilidade, o Vega e a segunda derivada em relação ao ativo objeto, o Gamma (Γ).

Como a fórmula de Black & Scholes (**BS**) é fechada, podemos calcular essas derivadas parciais e obter funções fechadas para cada uma delas. Seja C o preço de uma opção de compra européia, vamos definir:

$$C_\alpha = \frac{\partial C}{\partial \alpha} \quad (1.43)$$

Se assumirmos que o preço de C é dado por BS, temos então que seu delta é:

$$C_S^{BS} = \Delta^{BS} = \Phi(d_1), \quad (1.44)$$

onde R e $\Phi(d_1)$ são dados por (1.42).

O gamma é:

$$C_{SS}^{BS} = \Gamma^{BS} = \frac{\Phi'(d_1)}{S\sigma\sqrt{T-t}}. \quad (1.45)$$

E o Vega:

$$C_{\sigma}^{BS} = S\Phi'(d_1)\sqrt{T-t}, \quad (1.46)$$

onde $\Phi'(x)$ representa a densidade normal padrão.

1.14 O Delta-Hedge

Suponha que X_t ativo que segue a Equação (1.13), tal que $X_t = \log S_t$ e que exista uma taxa livre de risco r . Suponha também que $V(S, t)$ seja uma opção sobre S_t , com vencimento em T . Pela fórmula de Itô uma variação infinitesimal no valor dessa opção é dado por:

$$dV = \left(\mu SV_S + V_t + \frac{1}{2} S^2 \sigma^2 V_{SS} \right) dt + \sigma SV_S dW. \quad (1.47)$$

Considere agora um portfólio que tem como objetivo hedgear a posição comprada em uma unidade de $V(S, t)$, chamado de portfólio Delta-Hedge Π , e vende Δ unidades do ativo S_t ,

$$\Pi = V - \Delta S. \quad (1.48)$$

Uma variação infinitesimal nesse portfólio é dada por:

$$d\Pi = dV + \Delta dS, \quad (1.49)$$

Para que o portfólio seja livre de risco temos que zerar o termo do Browniano, logo devemos fazer $\Delta = \partial V / \partial S$. Então substituindo as Equações (1.13) e (1.47) em (1.49) temos

$$\begin{aligned} d\Pi &= \left(\mu SV_S + V_t + \frac{1}{2} \sigma^2 S^2 V_{SS} \right) dt + \sigma SV_S dW - \mu SV_S dt - \sigma SV_S dW \\ d\Pi &= \left(V_t + \frac{1}{2} \sigma^2 S^2 V_{SS} \right) dt \end{aligned} \quad (1.50)$$

A taxa de retorno desse portfólio deve ser igual a taxa livre de risco, caso contrário geraria uma oportunidade de arbitragem. Então:

$$\begin{aligned} r\Pi dt &= d\Pi \\ rV dt - rSV_S dt &= \left(V_t + \frac{1}{2} \sigma^2 S^2 V_{SS} \right) dt \\ \left(V_t + \frac{1}{2} \sigma^2 S^2 V_{SS} - rV dt + rSV_S \right) dt &= 0 \end{aligned} \quad (1.51)$$

Dessa forma obtemos a equação parcial diferencial de Black & Scholes:

$$V_t + \frac{1}{2} \sigma^2 S^2 V_{SS} + rSV_S - rV = 0 \quad (1.52)$$

1.15 Hipóteses do Modelo de Black-Scholes

A fórmula de Black-Scholes possui algumas hipóteses fortes e distantes da realidade. Primeiramente, a hipótese da existência de um ativo livre de risco. Em seguida, o modelo supõe que os retornos do ativo de risco possuem distribuição normal e volati-

lidade constante. Pelo argumento de não-arbitragem, podemos replicar a opção com o ativo objeto arriscado e o ativo livre de risco. A não-arbitragem só ocorre pela hipótese de completude do mercado. Tecnicamente,

Definição 1.15. (*Mercado Completo*) *Um mercado é completo se, e só se, todo ativo negociável neste mercado pode ser replicado por um portfólio autofinanciado composto somente com ativos desse mercado.*

Certamente, nenhuma dessas hipóteses se verificam na realidade. Uma melhoria no modelo BS a ser realizada é considerarmos que a volatilidade segue um processo estocástico. No entanto, ao introduzirmos uma nova fonte de incerteza o mercado deixa de ser completo, uma vez que a volatilidade instântanea é uma ativo não negociável e não pode ser replicado no mercado. A consequência disso é que um prêmio de risco é gerado para compensar a incompletude do mercado (ver mais em (Fouque, Papanicolaou & Sircar 2000)).

Capítulo 2

O Comportamento dos Estimadores

Neste capítulo desenvolvemos os principais resultados sobre convergência com dados em alta frequência. O problema clássico é a estimação de $\int_0^t \sigma_s^2 ds$ (neste caso, na ausência de microestrutura). O estimador padrão, a *Volatilidade Realizada*, é $[X, X]_t^{\mathcal{G}}$. Este estimador é consistente à medida que $\Delta(\mathcal{G}) \rightarrow 0$, resultado da definição de variância quadrática. No entanto, outras questões importantes emergem, como por exemplo, se o comportamento normal assintótico continua preservado.

2.1 Hipótese de Martingal Local

Vamos considerar o seguinte caso

$$X_t = X_0 + \int_0^t \sigma_s dW_s. \quad (2.1)$$

i.e., X_t é martingal local. Vamos assumir por enquanto que σ_t é limitada, *i.e.*, existe um σ_+ não aleatório que

$$\sigma_t^2 \leq \sigma_+^2 \quad \forall t. \quad (2.2)$$

Isto torna X_t um martingal.

2.2 O Processo de Erro

Em um grid $\mathcal{G} = \{t_0, t_1, \dots, t_n\}$ nós obtemos da fórmula de Itô que

$$(X_{t_{i+1}} - X_{t_i})^2 = 2 \int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) dX_s + \int_{t_i}^{t_{i+1}} \sigma_s^2 ds. \quad (2.3)$$

Definindo

$$t_* = \max\{t_i \in \mathcal{G} : t_i \leq t\} \quad (2.4)$$

analogamente, vale a mesma equação com (t_*, t) no lugar de (t_i, t_{i+1}) . Então

$$M_t = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^2 + (X_t - X_{t_*})^2 - \int_0^t \sigma_s^2 ds \quad (2.5)$$

é um martingal local da forma

$$M_t = 2 \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) dX_s + 2 \int_{t_*}^t (X_s - X_{t_*}) dX_s. \quad (2.6)$$

Devemos estudar o comportamento de martingais como M_t . No entanto, vale lembrar, que só observamos $[X, X]_t^{\mathcal{G}} = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^2$ mas veremos que os mesmos resultados se aplicam a este estimador

2.3 Símbolos de Ordem Estocásticos

Definição 2.1. *Suponha que Z_n seja uma sequência de variáveis aleatórias. Dizemos que $Z_n = o_p(1)$ se $Z_n \rightarrow 0$ em probabilidade, e que $Z_n = o_p(u_n)$ se $Z_n/u_n = o_p(1)$. Similarmente, dizemos que $Z_n = O_p(1)$ se para todo $\epsilon > 0$, existe um M tal que $\sup_n P(|Z_n| > M) \geq \epsilon$. Isto é o mesmo que dizer, que para cada subsequência n_{K_i} , existe outra subsequência n_{K_l} tal que Z_{n_l} converge em distribuição. Analogamente,*

$Z_n = O_p(u_n)$ se $Z_n/u_n = O_p(1)$.

Para ilustrar a Definição 2.1, temos que sob a Equação (2.2)

$$\begin{aligned}\mathbb{E}(X_t - X_{t_*})^2 &= \mathbb{E}([X, X]_t - [X, X]_{t_*}) \\ &= \mathbb{E} \int_{t_*}^t \sigma_s^2 ds \\ &\leq \mathbb{E}(t - t_*)\sigma_+^2 \\ &\leq \mathbb{E}\Delta(\mathcal{G})\sigma_+^2\end{aligned}$$

então $(X_t - X_{t_*})^2 = O_p(\Delta(\mathcal{G}))$.

2.4 Variação Quadrática do Processo de Erro

Para encontrar a variância do estimador, iniciamos computando sua variação quadrática

$$[M, M]_t^{\mathcal{G}} = 4 \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})^2 d[X, X]_s + 4 \int_{t_*}^t (X_s - X_{t_*})^2 d[X, X]_s. \quad (2.7)$$

Um resultado, originalmente realizado por (Barndorff-Nielsen & Shephard 2002), consiste na estimação desta variação. Definindo a quarticidade por

$$[X, X, X, X]_t^{\mathcal{G}} = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^4 + (X_t - X_{t_*})^4 \quad (2.8)$$

Utilizando a fórmula de Itô podemos ver que

$$\begin{aligned}d(X_t - X_{t_i})^4 &= 4(X_t - X_{t_i})^3 dX_t + 6(X_t - X_{t_i})^2 d[X, X]_t \\ &= 4(X_t - X_{t_i})^3 dX_t + \frac{6}{4} d[M, M]_t,\end{aligned} \quad (2.9)$$

Se definirmos

$$M_t^{(2)} = \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})^3 dX_s + \int_{t_*}^t (X_s - X_{t_*})^3 dX_s \quad (2.10)$$

obtemos

$$[X, X, X, X]_t^{\mathcal{G}} = \frac{3}{2}[M, M]_t + 4M_t^{(2)}. \quad (2.11)$$

Como o termo $M_t^{(2)}$ possui ordem $o_p(n^{-1})$, temos que $(2/3)n[X, X, X, X]_t^{\mathcal{G}}$ é um estimador consistente da variação quadrática (2.7).

Proposição 2.1. *Assumindo a Equação (2.2), supomos que, com $n \rightarrow 0$, $\Delta(\mathcal{G}) = o_p(1)$ e,*

$$\sum_{i=0}^{n-1} (t_{i+1} - t_i)^3 = O_p(n^{-2}). \quad (2.12)$$

Então

$$\sup_{0 \leq t \leq T} |[M, M]_t - \frac{2}{3}[X, X, X, X]_t^{\mathcal{G}}| = o_p(n^{-1}) \text{ quando } n \rightarrow 0. \quad (2.13)$$

2.4.1 Norma L^p

Para $1 \leq p < \infty$, definimos a *norma* L^p :

$$\|X\|_p = (\mathbb{E}|X|^p)^{\frac{1}{p}}, \quad (2.14)$$

as desigualdades de Minkowski e Hölder dizem que, respectivamente,

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p \quad (2.15)$$

$$\|XY\|_1 \leq \|X\|_p \|Y\|_q \text{ para } \frac{1}{p} + \frac{1}{q} = 1. \quad (2.16)$$

Um caso especial da desigualdade de Hölder é $\|X\|_1 \leq \|X\|_p$ (suponha $Y = 1$).

Em particular sob (2.12), para $1 \leq v \leq 3$:

$$\begin{aligned} \left(\frac{1}{n} \sum_{i=0}^n (t_{i+1} + t_i)^v \right)^{\frac{1}{v}} &\leq \left(\frac{1}{n} \sum_{i=0}^n (t_{i+1} + t_i)^3 \right)^{\frac{1}{3}} \\ &= \left(\frac{1}{n} \times O_p(n^{-2}) \right)^{\frac{1}{3}} = (O_p(n^{-3}))^{\frac{1}{3}} = O_p(n^{-1}), \end{aligned} \quad (2.17)$$

tal que

$$\sum_{i=0}^n (t_{i+1} - t_i)^v = O_p(n^{1-v}). \quad (2.18)$$

Para demonstrarmos a *Proposição 1*, precisamos da desigualdade de Burkholder-Davis-Gundy.

2.4.2 Desigualdade de Burkholder-Davis-Gundy

Para $1 \leq p < \infty$, existem constantes universais c_p e C_p tal que para todo martingal contínuo N_t ,

$$c_p \| [N, N]_T \|_{p/2}^{1/2} \leq \left\| \sup_{0 \leq t \leq T} |N_t| \right\| \leq C_p \| [N, N]_T \|_{p/2}^{1/2}. \quad (2.19)$$

Notamos que, em particular, para que $1 < p < \infty$,

$$C_p^2 = q^p \left(\frac{p(p-1)}{2} \right) \quad (2.20)$$

onde q é dado por $p^{-1} + q^{-1} = 1$.

2.4.3 Demonstração de Proposição 2.1

Demonstração. Da fórmula de Itô temos:

$$\begin{aligned} [M^{(2)}, M^{(2)}]_t &= \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i})^6 d[X, X]_s + \int_{t_*}^t (X_s - X_{t_*})^6 d[X, X]_s \\ &= \frac{1}{28} [X; 8]_t^{\mathcal{G}} + \text{termo martingal}, \end{aligned} \quad (2.21)$$

onde $[X; 8]_t^{\mathcal{G}} = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^8 + \frac{1}{28} (X_t - X_{t_*})^8$ é a *octicidade*.

Note que para um tempo de parada $\tau < T$, $[X; 8]_t^{\mathcal{G}} = \sum_i (X_{t_{i+1} \wedge \tau} - X_{t_i \wedge \tau})^8$. Então pela desigualdade de Burkholder-Davis-Gundy (com $p = 8$)

$$\begin{aligned} \mathbb{E}[M^{(2)}, M^{(2)}]_{\tau} &= \frac{1}{28} \mathbb{E}[X; 8]_{\tau}^{\mathcal{G}} \\ &\leq \frac{1}{28} C_8^8 \mathbb{E} \sum_i ([X, X]_{t_{i+1} \wedge \tau} - [X, X]_{t_i \wedge \tau})^4 \\ &\leq \frac{1}{28} C_8^8 \sigma_+^8 \mathbb{E} \sum_i (t_{i+1} \wedge \tau - t_i \wedge \tau)^4. \end{aligned} \quad (2.22)$$

Suponha $\epsilon > 0$, e seja

$$\tau_n = \inf \{ t \in [0, T] : n^{-2} \sum_i (t_{i+1} \wedge \tau - t_i \wedge \tau)^4 > \epsilon \}. \quad (2.23)$$

Então

$$\mathbb{E}[M^{(2)}, M^{(2)}]_{\tau_n} \leq n^{-2} \frac{1}{28} C_8^8 \sigma_+^8 \epsilon. \quad (2.24)$$

Por hipótese, $n^{-2} \sum_i (t_{i+1} \wedge \tau - t_i \wedge \tau)^4 \leq \Delta(\mathcal{G}) n^{-2} \sum_i (t_{i+1} - t_i)^3 \xrightarrow{p} 0$, e assim

$$P(\tau_n \neq T) \rightarrow 0 \text{ quando } n \rightarrow \infty. \quad (2.25)$$

Portanto, para qualquer $\delta > 0$,

$$\begin{aligned} P(n \sup_{0 \leq t \leq T} |M_t^{(2)}| > \delta) &\leq P(n \sup_{0 \leq t \leq T} |M_t^{(2)}| > \delta) + P(\tau_n \neq T) \\ &\leq \frac{1}{\delta^2} \mathbb{E} \left(n \sup_{0 \leq t \leq T} |M_t^{(2)}| \right)^2 + P(\tau_n \neq T) \text{ (Chebychev)} \\ &\leq \frac{1}{\delta^2} C_2^2 n^2 \mathbb{E}[M^{(2)}, M^{(2)}]_{\tau_n} + P(\tau_n \neq T) \text{ (Burkholder-Davis-Gundy)} \\ &\leq \frac{1}{\delta^2} C_2^2 \frac{1}{28} C_8^8 \sigma_+^8 \epsilon + P(\tau_n \neq T) \text{ (de (2.24))} \\ &\rightarrow \frac{1}{\delta^2} C_2^2 \frac{1}{28} C_8^8 \sigma_+^8 \epsilon \text{ quando } n \rightarrow \infty \text{ (de (2.25))} \end{aligned} \quad (2.26)$$

Assim segue a *Proposição 2.1*.

□

2.5 Variação Quadrática do Erro: Observações Independentes

Vamos assumir agora que as observações no tempo são processos independentes do processo X . A idéia básica é computar o estimador sobre pequenos intervalos de forma que $(X_t - X_*)^2 \approx [X, X]_t - [X, X]_{t_*}$. Para mostrarmos que esta aproximação é

válida, segue de (2.7) que

$$\begin{aligned}
[M, M]_t &= 4 \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} ([X, X]_s - [X, X]_{t_i}) d[X, X]_s \\
&\quad + \int_{t_*}^t ([X, X]_s - [X, X]_{t_*}) d[X, X]_s \\
&= 2 \sum_{t_{i+1} \leq t} ([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 + 2([X, X]_t - [X, X]_{t_*})^2 + o_p(n^{-1}). \quad (2.27)
\end{aligned}$$

Mais rigorosamente,

Proposição 2.2. *Assumindo (2.2), e que σ_t^2 é contínua em média quadrática:*

$$\sup_{0 \leq t-s \leq \delta} \mathbb{E}(\sigma_t^2 - \sigma_s^2)^2 \rightarrow 0 \text{ quando } \delta \rightarrow 0. \quad (2.28)$$

Também supomos que os grids \mathcal{G}_n são não-aleatórios, ou independentes do processo X_t . Também supomos que, $n \rightarrow 0$ quando $\Delta(\mathcal{G}) = o_p(n^{-1})$, e assumimos a Equação (2.12). Então

$$[M, M]_t = 2 \sum_{t_{i+1} \leq t} ([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 + 2([X, X]_t - [X, X]_{t_*})^2 + o_p(n^{-1}). \quad (2.29)$$

Se σ_t é contínua, sua média quadrática também é (por causa de (2.2)). De forma mais geral, σ_t pode, por exemplo, ter saltos.

Para mostrar a Proposição 2.2, precisamos de algumas notações e do seguinte lema:

Lema 2.1. *Suponha que N_t seja um processo de Itô martingal, com (para $a, b > 0$), que para todo t ,*

$$\frac{d}{dt} \mathbb{E}[N, N]_t \leq a(t - t_*)^b. \quad (2.30)$$

Assuma que H_t seja um processo adaptado, satisfazendo $|H_t| \leq H_+$ para alguma cons-

tante H_+ . Suponha também

$$R(\mathcal{G})_v = \left(\sum_{i=0}^n (t_{i+1} - t_i)^v \right). \quad (2.31)$$

Então

$$\begin{aligned} & \left\| \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_s ds + \int_{t_*}^t (N_s - N_{t_*}) H_s ds \right\|_1 \\ & \leq \left(H_+^2 \frac{a}{b+3} R_{b+3}(\mathcal{G}) \right)^{1/2} + R_{(b+3)/2}(\mathcal{G}) \frac{2}{b+3} \left(\frac{a}{b+1} \right)^{1/2} \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\|_2. \end{aligned} \quad (2.32)$$

Demonstração. (da Proposição 2.2) Suponha que $N_t = M_t$ e $H_t = \sigma_t^2$. Portanto

$$\begin{aligned} d[M, M]_t &= 4(X_t - X_{t_i})^2 d[X, X]_t \\ &= 4([X, X]_t - [X, X]_{t_i}) d[X, X]_t + 4 \left((X_t - X_{t_i})^2 - ([X, X]_t - [X, X]_{t_i}) \right) d[X, X]_t \\ &= 4([X, X]_t - [X, X]_{t_i}) d[X, X]_t + 2(N_t - N_{t_i}) \sigma_t^2 dt. \end{aligned} \quad (2.33)$$

Dessa forma, a aproximação do erro na Equação (2.29) é exatamente da forma do lado esquerdo da Equação (2.32). Notamos que

$$\begin{aligned} \mathbb{E}d[N, N]_t &= \mathbb{E}(X_t - X_{t_i}) d[X, X]_t \\ &= \mathbb{E}(X_t - X_{t_i}) \sigma_+^2 dt \\ &= (t - t_i) \sigma_+^4 dt \end{aligned} \quad (2.34)$$

segue que as condições do Lema 2.1 são satisfeitas com $a = \sigma_+^4$ e $b = 1$. \square

Demonstração. (do Lema 2.1) Decompondo o problema em:

$$\int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_s ds = \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_{t_i} ds + \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) (H_s - H_{t_i}) ds. \quad (2.35)$$

No primeiro termo, da fórmula de Itô, $d((t_{i+1} - s)(N_s - N_{t_i})) = -(N_s - N_{t_i}) ds + (t_{i+1} - s) dN_s$, tal que

$$\int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_{t_i} ds = H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - s) dN_s \quad (2.36)$$

segue

$$\begin{aligned} \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_s ds &= \sum_{t_{i+1} \leq t} H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - s) dN_s \\ &+ \sum_{t_{i+1} \leq t} H_{t_i} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) (H_s - H_{t_i}) ds. \end{aligned} \quad (2.37)$$

O primeiro termo é o tempo que marca o final do martingal. Para cada incremento,

$$\begin{aligned} \mathbb{E} \left(\int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_{t_i} ds \right)^2 &= \mathbb{E} \left(H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - s) dN_s \right)^2 \\ &\leq H_+^2 \mathbb{E} \left(\int_{t_i}^{t_{i+1}} (t_{i+1} - s) dN_s \right)^2 \\ &= H_+^2 \mathbb{E} \left(\int_{t_i}^{t_{i+1}} (t_{i+1} - s)^2 d[N, N]_s \right) \\ &= H_+^2 \int_{t_i}^{t_{i+1}} (t_{i+1} - s)^2 d\mathbb{E}[N, N]_s \\ &= H_+^2 \int_{t_i}^{t_{i+1}} (t_{i+1} - s)^2 \frac{d}{ds} \mathbb{E}[N, N]_s ds \\ &= H_+^2 \int_{t_i}^{t_{i+1}} (t_{i+1} - s)^2 a(s - t_i)^b ds \\ &= H_+^2 \frac{a}{b+3} (t_{i+1} - t_i)^{b+3} \end{aligned} \quad (2.38)$$

e pela independência dos incrementos dos martingais,

$$\begin{aligned} \mathbb{E} \left(\sum_{t_{i+1} \leq t} H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - s) dN_s \right)^2 &\leq H_+^2 \frac{a}{b+3} \left(\sum_{t_{i+1} \leq t} (t_{i+1} - t_i)^3 \right) \\ &\leq H_+^2 \frac{a}{b+3} R_{b+3}(\mathcal{G}). \end{aligned} \quad (2.39)$$

Por outro lado, para o segundo termo de Equação (2.37),

$$\begin{aligned} \|(N_s - N_{t_i})(H_s - H_{t_i})\|_1 &\leq \|(N_s - N_{t_i})\|_2 \|(H_s - H_{t_i})\|_2 \\ &\leq (\mathbb{E}(N_s - N_{t_i})^2)^{1/2} \|(H_s - H_{t_i})\|_2 \\ &= (\mathbb{E}([N, N]_s - [N, N]_{t_i}))^{1/2} \|(H_s - H_{t_i})\|_2 \\ &= \left(\int_{t_i}^s \frac{d}{du} \mathbb{E}([N, N]_u) du \right)^{1/2} \|(H_s - H_{t_i})\|_2 \\ &\leq \left(\int_{t_i}^s a(u - t_i)^b du \right)^{1/2} \|(H_s - H_{t_i})\|_2 \\ &= \left(\frac{a}{b+1} (s - t_i)^{b+1} \right)^{1/2} \|(H_s - H_{t_i})\|_2 \\ &= (s - t_i)^{(b+1)/2} \left(\frac{a}{b+1} \right)^{1/2} \|(H_s - H_{t_i})\|_2 \end{aligned} \quad (2.40)$$

e disso temos

$$\begin{aligned} \left\| \int_{t_i}^{t_{i+1}} (N_s - N_{t_i})(H_s - H_{t_i}) ds \right\|_1 &\leq \int_{t_i}^{t_{i+1}} \|(N_s - N_{t_i})(H_s - H_{t_i})\|_1 ds \\ &\leq \int_{t_i}^{t_{i+1}} (s - t_i)^{(b+1)/2} ds \left(\frac{a}{b+1} \right)^{1/2} \sup_{t_i \leq s \leq t_{i+1}} \|(H_s - H_{t_i})\|_2 \\ &= (t_{i+1} - t_i)^{(b+3)/2} \frac{2}{b+3} \left(\frac{a}{b+3} \right)^{1/2} \sup_{t_i \leq s \leq t_{i+1}} \|(H_s - H_{t_i})\|_2. \end{aligned} \quad (2.41)$$

Então, finalmente, para o segundo termo da Equação (2.37),

$$\begin{aligned}
& \left\| \sum_{t_{i+1} \leq t} H_{t_i} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i})(H_s - H_{t_i}) ds \right\|_1 \leq \\
& \leq \sum_{t_{i+1} \leq t} H_{t_i} ((t_{i+1} - t_i)^{(b+3)/2}) \frac{2}{b+3} \left(\frac{a}{b+3} \right)^{1/2} \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|(H_s - H_{t_i})\|_2 = \\
& = R_{(b+3)/2}(\mathcal{G}) \frac{2}{b+3} \left(\frac{a}{b+1} \right)^{1/2} \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\|_2. \quad (2.42)
\end{aligned}$$

Então, a soma completa das Equações (2.37), (2.39) e (2.42) temos

$$\begin{aligned}
& \left\| \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_s ds \right\|_1 \leq \left\| \sum_{t_{i+1} \leq s} H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - t) dN_s \right\|_1 \\
& + \left\| \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i})(H_s - H_{t_i}) ds \right\|_1 \\
& \leq \left\| \sum_{t_{i+1} \leq s} H_{t_i} \int_{t_i}^{t_{i+1}} (t_{i+1} - t) dN_s \right\|_2 \\
& + \left\| \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (N_s - N_{t_i})(H_s - H_{t_i}) ds \right\|_1 \\
& \leq \left(H_+^2 \frac{a}{b+3} R_{b+3}(\mathcal{G}) \right)^{1/2} \\
& + R_{(b+3)/2}(\mathcal{G}) \frac{2}{b+3} \left(\frac{a}{b+1} \right)^{1/2} \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\|_2. \quad (2.43)
\end{aligned}$$

A parte com t_* para t é análoga, demonstrando o resultado. \square

2.6 Variação Quadrática do Tempo

Para obtermos uma forma final da variação quadrática, definimos a “Variação Quadrática Assintótica do Tempo” (VQAT), que é dado por

$$H(t) = \lim_{n \rightarrow \infty} \frac{n}{T} \sum_{t_{n,j+1} \leq t} (t_{n,j+1} - t_{n,j})^2, \quad (2.44)$$

supondo que o limite existe. Então temos:

Proposição 2.3. *Assumindo as condições da Proposição 2.2, e que VQAT existe, temos*

$$n[M, M]_t \xrightarrow{p} 2T \int_0^t \sigma_s^4 dH_s. \quad (2.45)$$

Demonstração. Da Equação (2.29),

$$\begin{aligned} [M, M]_t &= 2 \sum_{t_{i+1} \leq t} ([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 + 2([X, X]_s - [X, X]_{t_*})^2 + o_p(n^{-1}) \\ &= 2 \sum_{t_{i+1} \leq t} \left(\int_{t_i}^{t_{i+1}} \sigma_s^2 ds \right)^2 + 2 \left(\int_{t_*}^t \sigma_s^2 ds \right)^2 + o_p(n^{-1}) \\ &= 2 \sum_{t_{i+1} \leq t} ((t_{i+1} - t_i) \sigma_{t_i}^2)^2 + 2((t - t_*) \sigma_{t_*}^2)^2 + o_p(n^{-1}) \\ &= 2 \frac{T}{n} \int_0^t \sigma_s^4 dH_s + o_p(n^{-1}). \end{aligned} \quad (2.46)$$

□

Vamos dar um exemplos do VQAT:

Quando os instantes são equidistantes: $t_{i+1} - t_i = T/n$, então

$$\begin{aligned}
H(t) &\approx \frac{n}{T} \sum_{t_{n,i+1} \leq t} \left(\frac{T}{n}\right)^2 \\
&= \frac{T}{n} \#\{t_{i+1} \leq t\} \\
&= T \times \text{fração de } t_{i+1} \in [0, t] \\
&\approx T \times \frac{t}{T} = t.
\end{aligned} \tag{2.47}$$

2.7 Variação Quadrática no Caso Geral

Agora relaxando a hipótese de independência do tempo em relação ao processo X . Utilizando a inequação de Burkholder-Davis-Gundy, temos que

$$c_4^4 \mathbb{E}((X_{t_{i+1}} - X_{t_i})^4 | \mathcal{F}_{t_i}) \leq \mathbb{E}([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 | \mathcal{F}_{t_i}) \leq C_4^4 \mathbb{E}((X_{t_{i+1}} - X_{t_i})^4 | \mathcal{F}_{t_i}), \tag{2.48}$$

onde c_4 e C_4 são constantes universais. Pela lei dos grandes números, $[X, X, X, X]_t - \sum_i \mathbb{E}((X_{t_{i+1}} - X_{t_i})^4 | \mathcal{F}_{t_i})$ é um martingal de ordem menor que o próprio $[X, X, X, X]_t$. O mesmo ocorre para $\sum_i [([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 - \mathbb{E}([X, X]_{t_{i+1}} - [X, X]_{t_i})^2 | \mathcal{F}_{t_i})]$. Pela *Proposição 2.3*, segue que sob as condições necessárias, se $n[X, X, X, X]_t \xrightarrow{p} U_t$ quando $n \rightarrow \infty$, e se o VQAT H_t é absolutamente contínuo em t , então U_t também é absolutamente contínua e

$$c_4^4 2T \sigma_t^4 H'_t \leq U'_t \leq C_4^4 2T \sigma_t^4 H'_t \tag{2.49}$$

2.8 Convergência Estável

Para conceber este conceito precisamos de uma filtração \mathcal{F}_t com todos os processos relevantes (X_t, σ_t , etc) adaptados. Devemos assumir que Z_n (a quantidade que converge em distribuição) é mensurável com respeito a σ -álgebra \mathcal{X} , $\mathcal{F}_t \subseteq \mathcal{X}$. Essa condição é importante para excluir o ruído da microestrutura da filtração \mathcal{F}_t .

Definição 2.2. *Suponha que Z_n seja uma sequência de variáveis aleatórias \mathcal{X} -mensurável, $\mathcal{F}_t \subseteq \mathcal{X}$. Dizemos que Z_n converge \mathcal{F}_t -estável em distribuição para Z quando $n \rightarrow \infty$ se Z é mensurável com respeito a \mathcal{X} , então para todo $A \in \mathcal{F}_t$ e para toda função g contínua e limitada, $\mathbb{E}[I_A g(Z_n)] \rightarrow \mathbb{E}[I_A g(Z)]$ quando $n \rightarrow \infty$.*

Essa definição significa que, sob condições regulares, Z_n converge conjuntamente em lei com todas variáveis aleatórias \mathcal{F} -mensurável.

2.9 Normalidade Assintótica

Devemos nos preocupar com uma sequência de martingais M_t^n , $0 \leq t \leq T, n = 1, 2, \dots$, e como ela converge para um limite M_t . Consideramos apenas martingais contínuos, que são imaginados como variáveis aleatórias tomando valores no conjunto \mathbb{C} das funções contínuas $[0, T] \rightarrow \mathbb{R}$.

Para definir convergência, fraca e estável, precisamos do conceito de continuidade. Dizemos que g é uma função contínua $\mathbb{C} \rightarrow \mathbb{R}$ se:

$$\sup_{0 \leq t \leq T} |x_n(t) - x(t)| \rightarrow 0 \text{ implica que } g(x_n) \rightarrow g(x). \quad (2.50)$$

Notamos que se $(M_t^n) \xrightarrow{\mathcal{L}} (M_t)$ no senso de processo, como por exemplo, $M_t^n \xrightarrow{\mathcal{L}} M_T$ como variável aleatória. Isto acontece pois a função $x \rightarrow g(x) - x(T)$ é contínua.

Para mostrarmos resultados sobre martingais contínuos devemos utilizar a seguinte hipótese.

Hipótese 2.1. *Existem Brownianos $W_t^{(1)}, \dots, W_t^{(p)}$ (para algum p) que geram (\mathcal{F}_t) .*

Sob a *Hipótese 2.1*, segue que do *Lema 2.1* em (Jacod & Protter 1998) que convergência estável em lei de um martingal local M^n para um processo M é equivalente a convergência em lei do proceso $W^{(1)}, \dots, W^{(p)}, M^n$ para o proceso $W^{(1)}, \dots, W^{(p)}, M$.

Outro fato importante sobre convergência estável é que limite e variância quadrática estão unicamente conectados, no sentido da proposição a seguir.

Proposição 2.4. *Assumindo que M^n é uma sequência de martingais locais contínuos que convergem para um processo estável M . Então: $(M_n, [M_n, M_n])$ convergem estavelmente para $(M, [M, M])$.*

Para a prova dessa proposição ver Corolário VI.6.30 em (Jacod & Shiryaev 1987).

Teorema 2.1. *Assumindo a Hipótese 2.1. Seja (M_t^n) uma sequência de martingais locais em $[0, T]$, cada um adaptado para (\mathcal{F}_t) , com $M_0^n = 0$. Suponha que existe um processo (\mathcal{F}_t) adaptado f_t tal que*

$$[M^n, M^n]_t \xrightarrow{p} \int_0^t f_s^2 ds \text{ para cada } t \in [0, T]. \quad (2.51)$$

Suponha também que, para cada $i = 1, \dots, p$,

$$[M^n, W^{(i)}]_t \xrightarrow{p} 0 \text{ para cada } t \in [0, T]. \quad (2.52)$$

Portanto existe uma extensão (\mathcal{F}'_t) de (\mathcal{F}_t) , e um (\mathcal{F}'_t) -martingal M_t tal que (M_t^n) converge estavelmente para (M_t) . Existe também um Browniano (W'_t) tal que $(W_t^{(1)}, \dots, W_t^{(p)}, W'_t)$

é um processo de Wiener (\mathcal{F}'_t), e tal que

$$M = \int_0^t f_s dW'_s. \quad (2.53)$$

A demonstração deste teorema se encontra em (Zhang 2001) (p. 65-67).

2.10 Aplicação para Volatilidade Realizada

Mostraremos, agora, uma aplicação do estimador proposto nesta seção. Considere um processo de erro normal (por \sqrt{n})

$$M_t^n = 2n^{1/2} \sum_{t_{i+1} \leq t} \int_{t_i}^{t_{i+1}} (X_s - X_{t_i}) dX_s + 2n^{1/2} \int_{t_*}^t (X_s - X_{t_*}) dX_s. \quad (2.54)$$

Da Subseção 2.6, temos que a condição (2.51) do *Teorema 2.1* é satisfeita, com

$$f_t^2 = 2T\sigma^4 H'_t. \quad (2.55)$$

Basta, então, checarmos a condição da Equação (2.52). Notamos que

$$d[M^n, W^{(i)}]_t = 2n^{1/2}(X_t - X_{t_*})d[X, W^{(i)}]_t. \quad (2.56)$$

Podemos agora aplicar o *Lema 2.1* com $N_t = X_t$ e $H_t = (d/dt)[X, W^{(i)}]_t$. Da inequação de Cauchy-Shwartz (conhecida, neste caso, como desigualdade de Kunita-Watanabe)

$$\begin{aligned} |[X, W^{(i)}]_{t+h} - [X, W^{(i)}]_t| &\leq \sqrt{[X, X]_{t+h} - [X, X]_t} \sqrt{[W^{(i)}, W^{(i)}]_{t+h} - [W^{(i)}, W^{(i)}]_t} \\ &\leq \sqrt{\sigma_+^2 h} \sqrt{h} = \sigma_+ h \end{aligned} \quad (2.57)$$

(relembrando que a variação quadrática é um limite de uma soma de quadrados), então podemos tomar $H_+ = \sigma_+$. Por outro lado, $(d/dt)\mathbb{E}[N, N]_t \leq \sigma_+^2 = a(t - t_*)^b$ com $a = \sigma_+^2$ e $b = 0$.

Então, do *Lema 2.1*,

$$\begin{aligned}
\| [M^n, W^{(i)}]_t \|_1 &= 2n^{1/2} \left\| \int_{t_i}^{t_{i+1}} (N_s - N_{t_i}) H_s ds + \int_{t_*}^t (N_s - N_{t_*}) H_s ds \right\|_1 \\
&\leq 2n^{1/2} \left(H_+^2 \frac{a}{b+3} R_{b+3}(\mathcal{G}) \right)^{1/2} \\
&\quad + R_{(b+3)/2}(\mathcal{G}) \frac{2}{b+3} \left(\frac{a}{b+1} \right)^{1/2} \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\|_2 \\
&= O_p(n^{1/2} R_3(\mathcal{G})^{1/2}) + O_p(n^{1/2} R_{3/2}(\mathcal{G}) \sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\|) \\
&= o_p(1)
\end{aligned} \tag{2.58}$$

sob as condições da *Proposição 2.2*, desde que $R_v(\mathcal{G}) = O_p(n^{1-v})$ da Equação (2.18), e desde que $\sup_{0 \leq t-s \leq \Delta(\mathcal{G})} \|H_s - H_t\| = o_p(1)$. Com os resultados citados podemos mostramos que:

Teorema 2.2. *Assumindo a Hipótese 2.1 e as condições da Proposição 2.2, e que a VQAT H_t existe e é absolutamente contínua. Seja M_t^n dado por (2.54). Então (M_t^n) converge estavelmente em lei para M_t , dado por*

$$M_t = 2T \int_0^t \sigma_s^2 \sqrt{H'_s} dW'_s. \tag{2.59}$$

Um caso especial:

Corolário 2.1. *Sob as condições do Teorema acima, para t fixo,*

$$\sqrt{n} ([X, X]_t^{\mathcal{G}^n} - [X, X]_t) \xrightarrow{\mathcal{L}} U \times \left(2T \int_0^t \sigma_s^4 dH_s \right)^{1/2} \tag{2.60}$$

onde U é uma variável aleatória normal padrão independente de \mathcal{F}_T .

2.11 Estimadores Não Centrados

2.11.1 Estimação do Modelo Browniano

Vamos considerar a estimação do modelo Browniano. Tomamos o tempo $t = 0$ como o início do dia de negociação e $t = T$ como o final do dia.

Vamos assumir que existem n observações do processo (transações) neste período. Supomos que as transações são igualmente espaçadas no tempo, ou seja, que acontece uma observação a cada $\Delta t_n = T/n$ unidades de tempo. Essa hipótese não é realista, no entanto, não é o problema que estamos preocupados no momento e simplifica consideravelmente os cálculos.

As observações (log dos preços transacionados) são $X_{t_{n,i}}$, onde $t_{n,i} = i\Delta t_n$. Se tomarmos as diferenças, temos as seguintes observações

$$\Delta X_{t_{n,i+1}} = X_{t_{n,i+1}} - X_{t_{n,i}}, \quad i = 0, \dots, n-1. \quad (2.61)$$

Os $\Delta X_{t_{n,i+1}}$ são independentes e identicamente distribuídos (i.i.d.) com distribuição $N(\mu\Delta t_n, \sigma^2\Delta t_n)$. Os estimadores naturais de μ e σ^2 são:

$$\hat{\mu}_n = \frac{1}{n\Delta t_n} \sum_{i=0}^{n-1} \Delta X_{t_{n,i+1}} = \frac{(X_T - X_0)}{T} \text{ que é o EMV e o ENVVUM, e} \quad (2.62)$$

$$\hat{\sigma}_{n,EMV}^2 = \frac{1}{n\Delta t_n} \sum_{i=0}^{n-1} (\Delta X_{t_{n,i+1}} - \overline{\Delta X_{t_n}})^2 \text{ que é o EMV ou} \quad (2.63)$$

$$\hat{\sigma}_{n,ENVVUM}^2 = \frac{1}{(n-1)\Delta t_n} \sum_{i=0}^{n-1} (\Delta X_{t_{n,i+1}} - \overline{\Delta X_{t_n}})^2 \text{ que é o ENVVUM,} \quad (2.64)$$

onde, EMV é o estimador de máxima verossimilhança, e ENVVUM o estimador não viesado de variância uniformemente mínima.

Estes estimadores evidenciam algumas propriedades. Primeiramente, temos que destacar que μ não pode ser consistentemente estimado para T fixo. De fato, o $\hat{\mu}_n$ não depende de n , mas somente do processo em T e no início do período.

Mais surpreendentemente é que σ^2 pode ser estimado consistentemente para T fixo, quando $n \rightarrow \infty$. Em outras palavras $\hat{\sigma}_n^2 \rightarrow \sigma^2$ quando $n \rightarrow \infty$.

2.11.2 Comportamento dos Estimadores Não Centrados

Apesar de ser menos usual, para dados em alta frequência geralmente não se remove a média na estimação. A razão para isso é, definindo

$$\hat{\sigma}_{n,\text{n\~ao centrado}}^2 = \frac{1}{n\Delta t_n} \sum_{i=0}^{n-1} (\Delta X_{t_n,i+1})^2. \quad (2.65)$$

Notamos que para o EMV de $\hat{\sigma}_n^2$,

$$\begin{aligned} \hat{\sigma}_{n,EMV}^2 &= \frac{1}{n\Delta t_n} \sum_{i=0}^{n-1} (\Delta X_{t_n,i+1} - \overline{\Delta X_{t_n}})^2 \\ &= \frac{1}{n\Delta t_n} \left(\sum_{i=0}^{n-1} (\Delta X_{t_n,i+1})^2 - n(\overline{\Delta X_{t_n}})^2 \right) \\ &= \hat{\sigma}_{n,\text{n\~ao centrado}}^2 - \Delta t_n \hat{\mu}_n^2 \\ &= \hat{\sigma}_{n,\text{n\~ao centrado}}^2 - \frac{T}{n} \hat{\mu}_n^2. \end{aligned}$$

Como $\hat{\mu}_n^2$ não depende de n , temos que

$$n^{1/2}(\hat{\sigma}_{n,EMV}^2 - \hat{\sigma}_{n,\text{n\~ao centrado}}^2) \xrightarrow{p} 0.$$

Então, $\hat{\sigma}_{n,\text{n\~ao centrado}}^2$ é consistente e possui a mesma distribuição assintótica que $\hat{\sigma}_{n,EMV}^2$ e $\hat{\sigma}_{n,ENVVUM}^2$. Isto é comum nos estimadores com dados em alta frequência.

2.12 Livrando-se do *drift*

A idéia é emprestada da teoria de apreçamento de ativos. Pensamos que existe a medida de probabilidade real Q , mas preferimos trabalhar com a medida neutra ao risco P , que possui cálculos bem mais simples.

O plano é sempre realizar a análise sob P e ajustar os resultados de volta para Q utilizando a derivada de Radon-Nikodym dP/dQ . Especificamente, suponha que θ é a quantidade a ser estimada (como, por exemplo, $\int_0^T \sigma_t^2 dt$). Um estimador $\hat{\theta}_n$ é encontrado com a ajuda de P , e o resultado assintótico é estabelecido por,

$$n^{1/2}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} N(b, a^2) \quad (2.66)$$

estavelmente sob P . Este é um resultado direto da equivalência da teoria da medida que $n^{1/2}(\hat{\theta}_n - \theta)$ também converge em lei sob Q . Em particular, a consistência e a taxa de convergência não são afetados pela mudança de medida, para um horizonte fixo T .

A distribuição assintótica deve ser diferente sob P e Q . Enquanto a distribuição normal permanece, seus parâmetros b e a^2 podem ser distintos para cada medida. O resultado técnico é

Proposição 2.5. *Suponha que Z_n é uma sequência de variáveis aleatórias que convergem estavelmente para $N(b, a^2)$ sob P . Se quer dizer que $N(b, a^2) = b + aN(0, 1)$, onde $N(0, 1)$ é uma variável normal padrão independente de \mathcal{F}_t , com a e b \mathcal{F}_t -mensuráveis. Então Z_n converge estavelmente em lei para $b + aN(0, 1)$ sob P , onde $N(0, 1)$ permanece independente de \mathcal{F}_t sob Q .*

Demonstração.

$$\mathbb{E}_Q I_{Ag}(Z_n) = \mathbb{E}_P \frac{dQ}{dP} I_{Ag}(Z_n) \rightarrow \mathbb{E}_P \frac{dQ}{dP} I_{Ag}(Z) = \mathbb{E}_Q I_{Ag}(Z) \quad (2.67)$$

pela integrabilidade uniforme de $\frac{dQ}{dP}I_{Ag}(Z_n)$. □

A *Proposição 2.5* simplifica substancialmente os cálculos e resultados. No caso aqui exposto, podemos sempre tomar o *drift* como zero, enquanto no caso de apreçamento de opções, isto só pode ser feito para processos de preço descontados de ativos. Em ambos os casos, a intenção é se livrar do “termo dt ”. A idéia de combinar convergência estável com mudança de medida apareceu primeiramente em (Rootzen 1980).

Capítulo 3

Análise de Volatilidade Realizada sob o Ruído da Microestrutura

O modelo citado até este capítulo, faz referência a um mundo ideal, onde o preço de compra é o mesmo do preço de venda e todas as operações são pequenas a ponto de não influenciar o mercado. No entanto, observações do mundo real são bem diferentes. Por isto vamos introduzir ao modelo (1.13) um termo de erro que tem como objetivo incorporar as imperfeições de mercado, que são denominadas microestruturas. A introdução deste termo se justifica pela observação empírica de que a estimação utilizando o modelo para o mundo ideal diverge do esperado teoricamente, o que significa que o modelo não consegue estimar robustamente o parâmetro requerido. O modelo proposto por (Zhang, Mykland & Aït-Sahalia 2005) encontra uma solução para que se possa utilizar todas observações e obter, simultaneamente, um estimador não-viesado da integral da volatilidade. No processo de construirmos o melhor estimador para volatilidade integrada, analisamos alguns prováveis estimadores.

A microestrutura, pode ser visto com uma incompletude do mercado que impede a replicação do ativo arriscado por não podermos ter disponível um só preço para compra ou venda ou por termos que truncar o preço em duas casas decimais. Esse mercado

incompleto gera um prêmio de risco associado à σ_t tornando a volatilidade integrada apenas como uma estimação ruim da volatilidade realizada. Tendo ciência desse fato, para melhorarmos essa estimação da volatilidade realizada temos que adicionar termos à volatilidade integrada que contemplem esse prêmio de risco, que no nosso modelo é simploriamente representado pelo termo de erro da microestrutura.

3.1 O Modelo dos Retornos Observados

Considere um processo de preço de um ativo S_t e suponha que o processo $X_t = \log S_t$ segue um processo de Itô:

$$dX_t = \mu_t dt + \sigma_t dW_t \quad (3.1)$$

onde W_t é um processo de Wiener padrão, μ_t é o *drift* e σ_t^2 é a variância instantânea dos retornos do processo X_t .

A primeira dificuldade originada pelo mercado advém do que estatisticamente é chamado de “erro observado”. Dessa forma, uma transação é vista como uma observação do processo de preço. Assim, incorpora-se o erro da observação para estimar a volatilidade integrada, ou seja, se supõe que cada retorno da amostra observado é da forma

$$Y_{t_i} = X_{t_i} + \epsilon_{t_i}, \quad (3.2)$$

onde X_t é o processo de retorno verdadeiro. Os ϵ_{t_i} 's são os erros independentes do processo verdadeiro em relação ao processo observado Y_t .

Então Y será o logaritmo do preço de transação observado nos tempos $0 = t_0, t_1, \dots, t_n = T$. Vamos considerar que em cada um destes momentos, Y é relacionado ao preço verdadeiro X , também em escala logarítmica, de acordo com (3.2). O ruído

ϵ_{t_i} satisfaz as seguintes hipóteses:

1. ϵ_{t_i} é independente e identicamente distribuído(iid),
2. $\mathbb{E}(\epsilon_{t_i}) = 0$,
3. $\text{Var}(\epsilon_{t_i}) = \mathbb{E}(\epsilon)^2$,
4. $\epsilon \perp X$;

onde \perp denota a independência entre os dois processos aleatórios. O modelo descrito na Equação (3.2) não requer que ϵ_t exista para todo t .

O parâmetro que se quer estimar é a volatilidade integrada sobre vários períodos, $\int_0^{T_1} \sigma_t^2, \int_{T_1}^{T_2} \sigma_t^2, \dots$. Uma forma natural de estimar este objeto seria realizar a soma dos retornos ao quadrado de 0 a T,

$$[X, X]_T = \sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2 \quad (3.3)$$

onde os X_{t_i} 's são todas observações do processo de retorno de $[0, T]$. O estimador $\sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2$ é normalmente utilizado e geralmente chamado de “variância realizada”. De acordo com a equação (3.2), a aproximação teórica de (3.3) é justificado pelo resultado teórico em processos estocásticos (Isometria de Itô) de forma que

$$p - \lim \sum_{t_i} (X_{t_{i+1}} - X_{t_i})^2 = \int_0^T \sigma_t^2 dt \quad (3.4)$$

à medida que a frequência amostral cresce, ou seja, a estimação do erro da volatilidade realizada diminui. Então, de acordo com (3.3) quanto maior frequência computada da amostra, melhor deveria ser a estimação da integral da volatilidade estocástica.

No entanto, isto não ocorre na prática, pelo menos na literatura de finanças empírica. O que usualmente é feito é se amostrar com frequências menores, pois se mostrou empiricamente que o estimador da volatilidade realizada não é robusto quando

a amostra possui um pequeno intervalo de tempo entre cada observação, mesmo quando o ativo objeto pode ser observado em extrema alta frequência, como vários negócios por segundo. Esse problema advém do grande viés na estimação e da ausência de robustez para intervalos pequenos como descrito em (Aït-Sahalia, Mykland & Zhang 2010) quando se utiliza observações em alta frequência. Em termos matemáticos, a volatilidade realizada não é nem um semimartingal; em outras palavras, a volatilidade realizada não converge com o crescimento da frequência de observações em cada amostra.

A principal explicação para este fenômeno se dá pelas imperfeições do mercado, chamadas de microestrutura, incluindo a existência do *spread* de compra-venda das ordens. A literatura mostra que, quando a amostra é de alta frequência, estas imperfeições se acentuam, criando um viés de microestrutura, tornando este tipo de dado sem utilidade para esse fim, de forma que a maioria dos autores procuram amostrar com horizontes mais longos para obterem resultados consistentes. O tamanho da janela é escolhido *ad hoc* entre 5 e 30 minutos. Como exemplo, uma janela de 5 minutos, amostrada a cada segundo, acaba por descartar, em torno 299 de 300 observações.

Nosso foco é determinar a volatilidade integrada X no período de $[0, T]$, que é também conhecida como

$$\langle X, X \rangle = \int_0^T \sigma_t^2 dt. \quad (3.5)$$

Antes de começar a análise do modelo, é necessário definir alguns conceitos.

Definição 3.1. *O grid completo que contém todas as observações é dado por*

$$\mathcal{G} = \{t_0, \dots, t_n\}. \quad (3.6)$$

Consideraremos também *grids* arbitrários, $\mathcal{H} \subseteq \mathcal{G}$.

Definição 3.2. *Para fazermos referência aos elementos destes grids temos que se*

$t_i \in \mathcal{H}$ então $t_{i,+}$ e $t_{i,-}$ serão os elementos sucessor e antecessor, respectivamente, em \mathcal{H} .

Obs.: t_i será sempre um i -ésimo elemento do grid completo. Quando $\mathcal{H} = \mathcal{G}$ temos $t_{i,-} = t_{i-1}$ e $t_{i,+} = t_{i+1}$. No entanto, quando \mathcal{H} é um subgrid de \mathcal{G} , geralmente teremos $t_{i,-} < t_{i-1}$ e $t_{i,+} > t_{i+1}$.

Por fim, definimos

$$|\mathcal{H}| = (\# \text{ número de pontos em } \mathcal{H}) - 1. \quad (3.7)$$

Essa definição é para elucidar que $|\mathcal{H}|$ é o numero de incrementos $(t_i, t_{i,+}]$ e que os extremos estão contidas em \mathcal{H} . Em particular, $|\mathcal{G}| = n$.

Definição 3.3. A variação quadrática $[\cdot, \cdot]$ para um processo genérico Z (como X ou Y) em grid arbitrário $\mathcal{H} \subseteq \mathcal{G}$ é dado por

$$[Z, Z]_t^{\mathcal{H}} = \sum_{t_i, t_{i,+} \in \mathcal{H}, t_{i,+} \leq t} (Z_{t_{i,+}} - Z_{t_i})^2. \quad (3.8)$$

Quando fizermos referência a um *grid* como $[Z, Z]_t^{\mathcal{H}}$ na continuação do texto, a representação será apenas $[Z, Z]_t$. Quando utilizarmos a totalidade do *grid* \mathcal{G} , a variação quadrática será denotada por

$$[Z, Z]_t^{\text{tudo}} = [Z, Z]_t^{\mathcal{G}} = \sum_{t_i, t_{i+1} \in \mathcal{G}, t_{i+1} \leq t} (\Delta Z_{t_i})^2, \quad (3.9)$$

onde $\Delta Z_{t_i} = Z_{t_{i+1}} - Z_{t_i}$. Note que esta notação adotada para ΔZ_{t_i} é diferente da convenção usualmente adotada em cálculo estocástico $\Delta Z_{t_i} = Z_{t_i} - Z_{t_{i-1}}$.

Como última observação vale ressaltar que quando $n \rightarrow \infty$ estaremos lidando com uma sequência de *grids* $\mathcal{G}_n = \{t_{0,n}, \dots, t_{n,n}\}$, assim como para *subgrids*. No entanto, não utilizaremos o subíndice duplo para não carregar a notação.

3.2 A Volatilidade Realizada

A primeira idéia natural corresponde ignorar totalmente o ruído e utilizar todos os dados para calcular a variação quadrática $[Y, Y]_T^{(tudo)}$ na esperança de se diminuir o erro. Para analisar se volatilidade realizada é um bom estimador da volatilidade integrada, devemos encontrar o comportamento da sua média para sabermos o viés e sua variância para investigar sua consistência. Sob o modelo (3.2), a volatilidade realizada baseada nos retornos observados Y_{t_i} tem a forma

$$[Y, Y]_T^{tudo} = [X, X]_T^{tudo} + 2[X, \epsilon]_T^{tudo} + [\epsilon, \epsilon]_T^{tudo}. \quad (3.10)$$

Dessa forma podemos calcular a esperança e a variância condicional de $[Y, Y]_T^{tudo}$,

$$\mathbb{E}([Y, Y]_T^{tudo} | X) = [X, X]_T^{tudo} + 2n\mathbb{E}\epsilon^2, \quad (3.11)$$

sob as hipóteses de (3.2). Analogamente,

$$\text{Var}([Y, Y]_T^{tudo} | X) = 4n\mathbb{E}\epsilon^4 + O_p(1), \quad (3.12)$$

sujeito as condições de (3.2) e $E\epsilon_{t_i}^4 = E\epsilon^4 < \infty, \forall i$. Sob condições um pouco mais fortes, como $|\mu_t|$ e σ_t serem limitados superiormente por uma constante, temos,

$$\text{Var}([Y, Y]_T^{tudo} | X) = 4n\mathbb{E}\epsilon^4 + (8[X, X]_T^{tudo}\mathbb{E}\epsilon^2 - 2\text{Var}(\epsilon^2)) + O_p(n^{-1/2}). \quad (3.13)$$

Quando $n \rightarrow \infty$, condicionado ao processo X , temos normalidade assintótica,

$$n^{1/2}([Y, Y]_T^{tudo} - 2n\mathbb{E}\epsilon^2) \xrightarrow{\mathcal{L}} 2(\mathbb{E}\epsilon^4)^{1/2} Z_{\text{ruído}}. \quad (3.14)$$

Onde $Z_{\text{ruído}}$ é uma normal padrão e o subscrito “ruído” indica a aleatoriedade vinda

do ruído ϵ , que são os desvios das observações Y do processo verdadeiro X .

As Equações (3.11) e (3.12) sugerem que no mundo discreto, na presença de microestruturas, a volatilidade realizada $[Y, Y]_T^{tudo}$ não seria um bom estimador de a verdadeira variação $[X, X]_T^{tudo}$ dos retornos. Para n grande, a volatilidade realizada é relacionada com o termo do ruído, $\mathbb{E}\epsilon^2$ em primeira ordem e $\mathbb{E}\epsilon^2$ em segunda ordem. Podemos notar em (3.11) que $[Y, Y]_T^{tudo}$ possui viés positivo com o crescimento da magnitude da amostra. Podemos então deduzir um estimador consistente da variância do ruído,

$$\widehat{\mathbb{E}\epsilon^2} = \frac{1}{2n}[Y, Y]_T^{tudo}. \quad (3.15)$$

Para um processo fixo X ,

$$n^{1/2}(\widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2) \rightarrow N(0, \mathbb{E}\epsilon^4) \text{ quando } n \rightarrow \infty. \quad (3.16)$$

O estimador consistente da variância assintótica de $\widehat{\mathbb{E}\epsilon^2}$ é dado por

$$\widehat{\mathbb{E}\epsilon^4} = \frac{1}{2n} \sum_i (\Delta Y_{t_i})^4 - 3(\widehat{\mathbb{E}\epsilon^2})^2. \quad (3.17)$$

A demonstração de (3.16) e (3.17) se encontra em (Zhang, Mykland & Ait-Sahalia 2005).

O questão que emerge naturalmente para quem está interessado em calcular a variação quadrática de X (mesmo $\langle X, X \rangle$) no lugar na variação quadrática de Y . Alguém poderia argumentar que $[Y, Y]$ é de fato a volatilidade que está acontecendo.

O principal motivo de estarmos interessado em estimar a variação quadrática de X é que a variação causada por ϵ 's é relacionada com cada transação e não com o caminho seguido pelo ativo objeto. Os ϵ 's, do ponto de vista da operação, representam os custos de transação, que são diferentes dos custos causados pela volatilidade, pois cada agente, por exemplo, pode “enxergar” um custo particular dada a natureza da sua

relação com o mercado.

É notório também que aplicar finanças em tempo contínuo para Y é de difícil implementação pois estaríamos utilizando quantidades dependentes da frequência dos dados. Finalmente, independente das aplicações, é interessante separar o processo de log-retornos dos efeitos do mecanismo de transação para podermos indentificar esse processo do ativo puramente.

3.3 Possíveis Estimadores

3.3.1 Amostrando Esparsamente

Como mostramos na seção anterior, a volatilidade realizada estima uma quantidade errada. Este problema piora quando as observações são amostrada mais frequentemente e a interpretação de Finanças da Equação (3.2) não se aplica na presença de microestrutura. Quando a frequência de amostragem é muito alta, a flutuação do processo de retorno é pesadamente contaminado pelo ruído da microestrutura e se torna menos representativa da verdadeira variação $\langle X, X \rangle_T$ dos retornos. Pelo problema apresentado, é consenso na aplicação em finanças, amostrar com menos frequência quando se utiliza volatilidade realizada. Vamos discutir, como isto pode ser visto no modelo (3.2).

Formalmente, a volatilidade realizada amostrando esparsamente é implementado amostrando no subgrid \mathcal{H} de \mathcal{G} ,

$$[Y, Y]^{esparsa} = [Y, Y]^{\mathcal{H}} = \sum_{t_i, t_{i,+} \in \mathcal{H}} (Y_{t_{i,+}} - Y_{t_i})^2. \quad (3.18)$$

Por enquanto, supomos o subgrid como dado. Vamos chamar $n_{esparsa} = |\mathcal{H}|$ e supor que a distribuição de ϵ , $\mathcal{L}(\epsilon)$, é um elemento de um conjunto \mathcal{D} de todas as distribuições

tal que $\mathbb{E}(\epsilon) = 0$ e onde $\mathbb{E}(\epsilon^2)$ e $\mathbb{E}(\epsilon^4)/\mathbb{E}(\epsilon^2)^2$ são limitadas por constantes arbitrárias.

Lema 3.1. *Suponha X um processo de Itô da forma (3.1), onde $|\mu_t|$ é limitada superiormente por uma constante. Suponha que para dado n , o grid \mathcal{H}_n é dado por, com $n_{\text{esparso}} \rightarrow \infty$ quando $n \rightarrow \infty$, e para cada n , Y é relacionado a X pelo modelo (3.2) e suas hipóteses de ϵ_t , onde $\mathcal{L}(\epsilon) \in \mathcal{D}$. Assim como o grid \mathcal{H}_n , a distribuição $\mathcal{L}(\epsilon)$ pode depender de n (onde o processo X é fixo). Suponha também que $\max_i \Delta t_i \rightarrow 0$ quando $n \rightarrow \infty$ satisfeito pela sequência de grids \mathcal{H}_n . Então*

$$\begin{aligned} [Y, Y]_T^{\mathcal{H}} &= [X, X]_T^{\mathcal{H}} + 2n_{\text{esparso}}\mathbb{E}\epsilon^2 + \left(4n_{\text{esparso}}\mathbb{E}\epsilon^4 + (8[X, X]_T^{\mathcal{H}}\mathbb{E}\epsilon^2 - 2\text{Var}(\epsilon^2))\right)^{1/2} Z_{\text{noise}} \\ &\quad + O_p\left(n_{\text{esparso}}^{1/4}(\mathbb{E}\epsilon^2)^{1/2}\right), \end{aligned} \quad (3.19)$$

onde Z_{noise} é a quantidade que é assintoticamente normal padrão.

A demonstração do *Lema 3.1* pode ser encontrada em (Zhang, Mykland & Aït-Sahalia 2005).

Note que agora a ordem relativa nos termos de (3.19) depende da quantidade n_{esparso} e $\mathbb{E}\epsilon^2$. Logo, se $\mathbb{E}\epsilon^2$ é pequeno em relação n_{esparso} , então $[Y, Y]_T^{\mathcal{H}}$ pode ser um bom substituto para $[X, X]_T^{\mathcal{H}}$.

Alguém poderia sugerir que a escolha ótima para n_{esparso} seria fazê-lo o menor possível. No entanto, deve-se atentar que quanto maior o n_{esparso} , mais próximo $[X, X]_T^{\mathcal{H}}$ estará da volatilidade integrada $\langle X, X \rangle_T$, que é o principal objetivo.

Para quantificarmos o erro total, combinamos o *Lema 3.1* com os resultados do erro de discretização para estudar o erro de $[Y, Y]_T^{\mathcal{H}} - \langle X, X \rangle_T$. Sob as condições impostas nos trabalhos (Barndorff-Nielsen & Shephard 2002) e (Mykland & Zhang 2006), pode-se mostrar que

$$\left(\frac{n_{\text{esparso}}}{T}\right)^{1/2} ([X, X]_T^{\mathcal{H}} - \langle X, X \rangle_T) \xrightarrow{\mathcal{L}} \left(\int_0^T 2H'(t)\sigma_t^4 dt\right)^{1/2} \times Z_{\text{discreto}}. \quad (3.20)$$

Onde $Z_{discreto}$ é uma variável aleatória normal padrão, com o subscrito “discreto” devido a aleatoriedade advir do efeito da discretização de $[X, X]_T^H$ quando calculamos $\langle X, X \rangle_T$. H_t é a variância assintótica quadrática do tempo, como discutido em (Mykland & Zhang 2006),

$$H(t) = \lim_{n \rightarrow \infty} \frac{n_{esparso}}{T} \sum_{t_i, t_{i,+} \in \mathcal{H}, t_{i,+} \leq t} (t_{i,+} - t_i)^2. \quad (3.21)$$

Em caso de observações equidistantes, $\Delta t_0 = \dots = \Delta t_{n-1} = \Delta t = T/n_{esparso}$ e $H'(t) = 1$. Pelo fato de ϵ 's ser independente do processo X , $Z_{ruído}$ é independente de $Z_{discreto}$.

Para um $\mathbb{E}\epsilon^2$ pequeno podemos agora estimar $\langle X, X \rangle_T$. Segue do *Lema 1* e da *Proposição 1* de (Mykland & Zhang 2006) que

Proposição 3.1. *Assumindo as condições do Lema 3.1 e que $\max_{t_i, t_{i,+} \in \mathcal{H}} (t_{i,+} - t_i) = O(1/n_{esparso})$. Sejam X_t e μ_t processos adaptados com relação a uma filtração $(\mathcal{F}_t)_{0 \leq t \leq T}$. Então H_t , da Equação (3.21) está bem definido e*

$$[Y, Y]_T^H = \langle X, X \rangle_T + 2\mathbb{E}\epsilon^2 n_{esparso} + \Gamma Z_{total} + O_p(n_{esparso}^{-1/4} (\mathbb{E}\epsilon^2)^{1/2}) + o_p(n_{esparso}^{-1/2}), \quad (3.22)$$

no sentido de convergência estável, onde Z_{total} é assintoticamente normal padrão e onde a variância Γ^2 tem a forma

$$\Gamma^2 = \underbrace{4n_{esparso}\mathbb{E}\epsilon^2 + (8[X, X]_T^H\mathbb{E}\epsilon^2 - 2\text{Var}(\epsilon^2))}_{\text{devido ao ruído}} + \underbrace{\frac{T}{n_{esparso}} \int_0^T 2H'(t)\sigma_t^4 dt}_{\text{devido à discretização}}. \quad (3.23)$$

A proposição mostra que existe um viés, $2\mathbb{E}\epsilon^2 n_{esparso}$, no entanto este viés diminui se utilizarmos poucas observações. Então, a prática em finanças empíricas de utilizar o estimador $[Y, Y]^{esparso}$ no lugar do candidato natural $[Y, Y]^{tudo}$ faz todo sentido. Pesquisadores empíricos geralmente utilizam retornos ao quadrado, amostrados em 5,

15 ou 30 minutos, mas deve-se atentar que não se pode amostrar exageradamente distante ao utilizar este estimador pois a medida que n_{esparsa} decresce a variância Γ^2 do estimador aumenta devido a discretização, como (3.23) mostra que a variância é proporcional a n_{esparsa}^{-1} , evidenciando o custo de oportunidade entre se amostrar mais frequentemente e se amostrar mais espalhadamente.

3.3.2 Subamostrando e Calculando a Média

Na subseção anterior, mostramos a vantagem de se amostrar com menos frequência. No entanto, fica claro que este tipo de abordagem desperdiça a maior parte das observações, o que é estatisticamente muito ruim, especialmente quando temos dados de ações muito líquidas, com mais de uma transação por segundo, em média. Pensando estatisticamente, a primeira idéia que se pode ter para resolver este problema, é dividir o grid original de observações em subgrids, $\mathcal{G} = \{t_0, \dots, t_n\}$ e tomar a média do estimador nestes subgrids. Um outro benefício imediato desta abordagem é diminuição da variância do estimador. Primeiramente vamos introduzir mais alguns conceitos importantes para estudarmos a volatilidade em múltiplos grids.

Vamos supor que o grid completo $\mathcal{G} = \{t_0, \dots, t_n\}$ é particionado em K subgrids não sobrepostos $\mathcal{G}^{(k)}$, onde $k = 1, \dots, K$; em outras palavras.

$$\mathcal{G} = \bigcup_{k=1}^K \mathcal{G}^{(k)}, \quad \text{onde } \mathcal{G}^{(k)} \cap \mathcal{G}^{(l)} = \emptyset \text{ quando } k \neq l. \quad (3.24)$$

Para nossos propósitos, o K -ésimo subgrid $\mathcal{G}^{(k)}$ se inicia em t_{k-1} e então pegamos todos os K -ésimos pontos seguintes até T . Ou seja,

$$\mathcal{G}^{(k)} = \{t_{k-1}, t_{k-1+K}, t_{k-1+2K}, \dots, t_{k-1+n_k K}\} \quad (3.25)$$

para $k = 1, \dots, K$, onde n_k é um inteiro que faz $t_{k-1+n_k K}$ o último elemento de $\mathcal{G}^{(k)}$.

Vamos chamar isto de uma **alocação regular** dos pontos nos subgrids.

Independente da alocação ser regular, temos $n_k = |\mathcal{G}^{(k)}|$ como definido em (3.7). Lembre-se que a volatilidade realizada sobre todos os pontos em \mathcal{G} é representada por $[Y, Y]_T^{tudo}$. Portanto, se utilizarmos uma subamostra Y_t tal que $t \in \mathcal{G}^{(k)}$ então a volatilidade realizada denotada por $[Y, Y]_T^k$, é da forma

$$[Y, Y]_T^k = \sum_{t_i, t_{i,+} \in \mathcal{G}^{(k)}} (Y_{t_{i,+}} - Y_{t_i})^2, \quad (3.26)$$

onde, para $t_i \in \mathcal{G}^{(k)}$, $t_{i,+}$ representa o próximo elemento em $\mathcal{G}^{(k)}$.

Dessa forma, um competidor natural a $[Y, Y]_T^{tudo}$ e $[Y, Y]_T^{esparso}$ é dado por

$$[Y, Y]_T^{\text{média}} = \frac{1}{K} \sum_{k=1}^K [Y, Y]_T^k. \quad (3.27)$$

Como anteriormente, temos T fixo e utilizamos apenas as observações no período de $[0, T]$. Assintoticamente consideramos,

$$\begin{aligned} \max_i \Delta t_i &\rightarrow 0 \text{ quando } n \rightarrow \infty \text{ e} \\ n \rightarrow \infty \text{ então } \frac{n}{K} &\rightarrow \infty. \end{aligned} \quad (3.28)$$

Por fim, vamos definir

$$\bar{n} = \sum_{k=1}^K n_k = \frac{n - K - 1}{K}. \quad (3.29)$$

3.3.3 O Erro devido ao Ruído

Relembrando que nosso objeto de desejo é estimar a volatilidade integrada $\langle X, X \rangle_T$, ou seja, a variação quadrática dos retornos verdadeira, mas não observável. Para tal, primeiro vamos descobrir quão bom o estimador $[Y, Y]_T^{\text{média}}$ se aproxima de $[X, X]_T^{\text{média}}$, onde o último é a volatilidade integrada considerada apenas em uma escala discreta.

De (3.11) e (3.27) temos

$$\mathbb{E}([Y, Y]_T^{\text{média}} | X) = [X, X]_T^{\text{média}} + 2\bar{n}\mathbb{E}\epsilon^2. \quad (3.30)$$

Também, pelo fato de $\{\epsilon_t, t \in \mathcal{G}^{(k)}\}$ são independentes para k diferentes,

$$\text{Var}([Y, Y]_T^{\text{média}} | X) = \frac{1}{K^2} \sum_{k=1}^K \text{Var}([Y, Y]_T^k | X) = 4\frac{\bar{n}}{K}\mathbb{E}\epsilon^4 + O_p\left(\frac{1}{K}\right), \quad (3.31)$$

assim como em (3.12). Incorporando o termo de próxima ordem da variância, como em (3.13) temos

$$\text{Var}([Y, Y]_T^{\text{média}} | X) = 4\frac{\bar{n}}{K}\mathbb{E}\epsilon^4 + \frac{1}{K} (8[X, X]_T^{\text{média}}\mathbb{E}\epsilon^2 - 2\text{Var}(\epsilon^2)) + O_p\left(\frac{1}{K}\right). \quad (3.32)$$

As condições assintóticas do estimar são

Teorema 3.1. *Suponha que X seja um processo de Itô da forma (3.1), que esteja relacionado a Y pelo modelo (3.2) e que ϵ satisfaz as hipóteses descritas na (Seção 2.1) com $\mathbb{E}\epsilon^4 < \infty$. Assuma também que t_i e t_{i+1} não estão no mesmo subgrid para qualquer i . Sob a hipótese (3.28), com $n \rightarrow \infty$,*

$$\sqrt{\frac{K}{\bar{n}}}([Y, Y]_T^{\text{média}} - [X, X]_T^{\text{média}} - 2\bar{n}\mathbb{E}\epsilon^2) \xrightarrow{\mathcal{L}} 2(\mathbb{E}\epsilon^4)^{1/2} Z_{\text{ruído}}^{\text{média}}, \quad (3.33)$$

condicional ao processo X , onde $Z_{\text{ruído}}^{\text{média}}$ é uma normal padrão.

A demonstração deste teorema se encontra em (Zhang, Mykland & Aït-Sahalia 2005).

Este teorema pode ser comparado ao resultado de (3.14), apesar de $Z_{\text{ruído}}^{\text{média}}$ quase nunca ser $Z_{\text{ruído}}$. Comparando com estimador que utiliza o grid completo, $Z_{\text{ruído}}^{\text{média}}$ traz melhoras com viés e variância com menor ordem em relação a n se comparado a (3.11) e (3.12).

3.3.4 O Erro devido à Discretização

Nesta seção mostraremos qual é o impacto da discretização do tempo no estimador, ou seja, queremos saber o quanto $[X, X]_T^{\text{média}}$ se desvia da volatilidade integrada $\langle X, X \rangle_T$ do processo verdadeiro. Vamos chamar o efeito da discretização de D_T , onde

$$D_T = [X, X]_T^{\text{média}} - \langle X, X \rangle_T = \frac{1}{K} \sum_{k=1}^K ([X, X]_T^k - \langle X, X \rangle_T), \quad (3.34)$$

onde

$$[X, X]_T^k = \sum_{t_i, t_{i,+} \in \mathcal{G}, t_{i,+} \leq t} (X_{t_{i,+}} - X_{t_i})^2 \quad (3.35)$$

Queremos o resultado assintótico para D_T . O resultado é semelhante à (3.20), no entanto é mais complexo pela existência dos múltiplos grids.

Vamos considerar que os pontos estão regularmente alocados nos subgrids como em (3.25) e assumir também que

$$\max_i |\Delta t_i| = O\left(\frac{1}{n}\right) \quad \text{e que} \\ K/n \rightarrow 0. \quad (3.36)$$

Definindo a função peso:

$$h_i = \frac{4}{K\Delta t} \sum_{j=1}^{(K-1)\wedge i} \left(1 - \frac{j}{K}\right)^2 \Delta t_{i-j}. \quad (3.37)$$

No caso em que os t_i 's são equidistantes, como em uma alocação, regular nos subgrids, $\Delta t_i = \Delta t$, e todo h_i (exceto o primeiro $K - 1$) é igual a

$$h_i = \frac{4}{K} \sum_{j=1}^{(K-1)\wedge i} \left(1 - \frac{j}{K}\right)^2. \quad (3.38)$$

De forma mais geral, a Equação (3.38) mais a hipótese (3.36) implica que

$$\sup_i h_i = O(1). \quad (3.39)$$

Denominando $\langle D, D \rangle_T$ como a variação quadrática de D_t quando visto como um processo contínuo definido em (3.34), obtemos a melhor aproximação da variância de D_T .

Teorema 3.2. *Suponha que X_t seja um processo de Itô da forma (3.1), com média μ_t e variância σ_t , ambos contínuos quase certamente. Suponha, também, que σ_t é limitado inferiormente por 0. Assumindo (3.36) e que as observações da amostra são regularmente alocadas nos grids. Então a variância quadrática de D_T é aproximadamente*

$$\langle D, D \rangle_T = \frac{TK}{n} \nu_n^2 + o_p\left(\frac{K}{n}\right), \quad (3.40)$$

$$\nu_n^2 = \sum_i h_i \sigma_{t_i}^4 \Delta t_i. \quad (3.41)$$

Em particular, $D_T = O_p((\frac{K}{n})^{1/2})$. Disto derivamos o custo de oportunidade entre os dois efeitos citados: ruído e discretização. No entanto, temos que primeiro discutir a distribuição assintótica de D_T .

Teorema 3.3. *Assumindo as hipóteses do Teorema 3.2 temos*

$$\nu_n^2 \xrightarrow{P} \nu^2 \quad (3.42)$$

onde ν é aleatório. Assumimos que o processo D_T é adaptado.

Então

$$\frac{D_T}{(K/n)^{1/2}} \xrightarrow{\mathcal{L}} \nu \sqrt{T} Z_{discreto}, \quad (3.43)$$

onde $Z_{discreto}$ é uma normal padrão e independente do processo X . A convergência é

estável em distribuição.

Dito de outra forma, $D_T/(K/n)^{1/2}$ pode ser considerada assintoticamente normal com “ $N(0, \nu^2 T)$ ”. É conveniente supor (3.41), pois esta é satisfeita na maioria dos casos. No entanto, quando t_i 's são equidistantes, como numa alocação regular dos subgrids, temos,

$$\nu^2 = \frac{4}{3} \int_0^T \sigma_t^4 dt, \quad (3.44)$$

implicada por (3.39) (ver mais em (Zhang, Mykland & Aït-Sahalia 2005)).

3.4 Combinando os Dois Erros

Nesta seção combinaremos as duas fontes de erro citadas nas últimas duas seções: ruído e discretização. Do *Teorema 3.1* e *Teorema 3.3* temos

$$[Y, Y]_T^{\text{média}} - \langle X, X \rangle_T - 2\bar{n}\mathbb{E}\epsilon^2 = \xi Z_{total} + o_p(1), \quad (3.45)$$

onde Z_{total} é a variância assintótica normal padrão independente do processo X e

$$\xi^2 = \underbrace{4 \frac{\bar{n}}{K} \mathbb{E}\epsilon_4}_{\text{devido ao ruído}} + \underbrace{T \frac{1}{\bar{n}} \nu^2}_{\text{devido à discretização}}. \quad (3.46)$$

Uma vez que o ruído é desprezível assintoticamente, poderíamos procurar por um \bar{n} ótimo para balancear o custo de oportunidade entre viés e variância em (3.45).

Dessa forma, ao utilizar o estimador $[Y, Y]_T^{\text{média}}$, com dados alocados regularmente, podemos usar todos os dados da amostra e obter um desempenho melhor que $[Y, Y]_T^{\text{esparsa}}$.

3.5 O Estimador Final: Volatilidade em Duas Escalas de Tempo

Vimos na última seção que $[Y, Y]^{\text{média}}$ é um estimador viesado da volatilidade integrada $\langle X, X \rangle$. Nesta seção buscamos ajustar este viés para obter o melhor estimador. Da Equação (3.15), $\mathbb{E}\epsilon^2$ pode ser consistentemente aproximado por:

$$\widehat{\mathbb{E}\epsilon^2} = \frac{1}{2n} [Y, Y]_T^{\text{tudo}}. \quad (3.47)$$

Então o viés de $[Y, Y]^{\text{média}}$ pode ser estimado por $2\bar{n}\mathbb{E}\epsilon^2$. Então um estimador não-viesado para $\langle X, X \rangle$ pode ser obtido com

$$\widehat{\langle X, X \rangle}_T = [Y, Y]_T^{\text{média}} - \frac{\bar{n}}{n} [Y, Y]_T^{\text{tudo}}, \quad (3.48)$$

Para o resultado assintótico da Equação (3.47) considere a Equação (3.1) e as hipóteses sobre ϵ da Seção 3.1,

$$\begin{aligned} & \left(\frac{K}{\bar{n}}\right)^{1/2} \left(\widehat{\langle X, X \rangle}_T - [X, X]_T^{\text{média}} \right) = \\ & \left(\frac{K}{\bar{n}}\right)^{1/2} \left([Y, Y]_T^{\text{média}} - [X, X]_T^{\text{média}} - 2\bar{n}\mathbb{E}\epsilon^2 \right) - (K\bar{n})^{1/2} \left(\widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2 \right) \xrightarrow{\mathcal{L}} N\left(0, 8(\mathbb{E}\epsilon^2)^2\right), \end{aligned} \quad (3.49)$$

onde a convergência em distribuição é condicional a X .

Combinando este resultado com D_t da Subseção 3.3.4 podemos determinar a escolha ótima para K quando $n \rightarrow \infty$,

$$\begin{aligned} \widehat{\langle X, X \rangle}_T - \langle X, X \rangle_T &= \left(\widehat{\langle X, X \rangle}_T - [X, X]_T^{\text{média}} \right) + \left([X, X]_T^{\text{média}} - \langle X, X \rangle_T \right) \\ &= O_p\left(\frac{\bar{n}^{1/2}}{K^{1/2}}\right) + O_p\left(\bar{n}^{-1/2}\right) \end{aligned} \quad (3.50)$$

O erro é minimizado igualando os dois termos no lado direito de (3.50), lembrando que $\bar{n} \approx n/K$ concluímos que o passo ótimo para $[Y, Y]_T^{\text{média}}$ é $K = O(n^{2/3})$, assim o lado direito de (3.50) possui ordem de $O_p(n^{1/6})$.

Considerando o modelo (3.1) e tomando $K = cn^{2/3}$, a variância condicional de segunda ordem de $\widehat{\langle X, X \rangle}_T$, dado o processo X , é dada por:

$$\begin{aligned} \text{Var} \left(\widehat{\langle X, X \rangle}_T | X \right) &= n^{-1/3} c^{-2} 8(\mathbb{E}\epsilon^2)^2 + \\ &+ n^{-2/3} c^{-1} [8[X, X]_T^{\text{média}} \mathbb{E}\epsilon^2 - 2\text{Var}(\epsilon^2)] + o_p(n^{-2/3}). \end{aligned} \quad (3.51)$$

Se tivermos uma amostra grande, podemos ampliar $\widehat{\langle X, X \rangle}_T$ para sub-amostras menores, ou seja, calculadas em um intervalo menor que o tempo $[0, T]$, que vamos chamar de $\widehat{\langle X, X \rangle}_T^{\text{aju}}$ pois devemos ajustar seu viés. Para (a, b) arbitrários, consideremos todos estimadores da forma:

$$\widehat{\langle X, X \rangle}_T^{\text{aju}} = a[Y, Y]_T^{\text{média}} - b \frac{\bar{n}}{n} [Y, Y]_T^{\text{tudo}}, \quad (3.52)$$

De (3.11) e (3.30) temos,

$$\begin{aligned} \mathbb{E} \left(\widehat{\langle X, X \rangle}_T^{\text{aju}} | X \right) &= a ([X, X]_T^{\text{média}} + 2\bar{n}\mathbb{E}\epsilon^2) - b \frac{\bar{n}}{n} ([X, X]_T^{\text{tudo}} + 2n\mathbb{E}\epsilon^2) \\ &= a[X, X]_T^{\text{média}} - b \frac{\bar{n}}{n} ([X, X]_T^{\text{tudo}}) + 2(a - b)\bar{n}\mathbb{E}\epsilon^2. \end{aligned} \quad (3.53)$$

Naturalmente, se escolhermos $a = b$, removermos completamente o efeito de $\bar{n}\mathbb{E}\epsilon^2$ e $[X, X]_T^{\text{média}}$ $[X, X]_T^{\text{tudo}}$ são assintoticamente estimadores não-viesados de $\langle X, X \rangle_T$. Se fizermos $a(1 - \bar{n}/n) = 1$, então temos

$$\widehat{\langle X, X \rangle}_T^{\text{aju}} = \left(1 - \frac{\bar{n}}{n} \right) \widehat{\langle X, X \rangle}_T, \quad (3.54)$$

Analogamente, um estimador ajustado $\mathbb{E}\epsilon^2$'s dado por

$$\begin{aligned}\widehat{\mathbb{E}\epsilon^{2}{}^{aju}} - \mathbb{E}\epsilon^2 &= (\widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2) (1 + O(K^{-1})) + O_p(Kn^{-3/2}) \\ &= \widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2 (O_p(n^{-1/2}K^{-1})) + O_p(Kn^{-3/2}) = \widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2 + O_p(n^{-5/6})\end{aligned}\quad (3.55)$$

de (3.50). Segue que $n^{1/2}(\widehat{\mathbb{E}\epsilon^2} - \mathbb{E}\epsilon^2)$ e $n^{1/2}(\widehat{\mathbb{E}\epsilon^{2}{}^{aju}} - \mathbb{E}\epsilon^2)$ possuem, então, a mesma distribuição assintótica.

Capítulo 4

Estimação da Volatilidade

Realizada em Duas Escalas em Alta Frequência

Neste capítulo realizamos a estimação do modelo proposto no capítulo anterior, a volatilidade realizada em duas escalas (**VRDE**). O estimador usualmente utilizado na literatura para estimação volatilidade em alta frequência é a volatilidade realizada amostrada esparsamente, que neste capítulo, vamos denominar de volatilidade realizada (**VR**). As vantagens da VRDE em relação a este estimador é, primeiramente, a utilização de todas as observações, aproveitando, assim, toda informação disponível. Esta característica sugere que deveríamos ter uma maior variância na VRDE em relação a VR, para cada janela de tempo escolhida, pela presença da microestrutura, no entanto, assim como teoria prevê, a VRDE é bem mais robusta que a VR fornecendo um nível de volatilidade bastante confiável e pouco depende do intervalo de tempo utilizado na amostragem do estimador.

Recordando que ao denominarmos a volatilidade realizada em duas escalas (VRDE)

estamos nos referindo à

$$\widehat{\langle X, X \rangle}_T^{aju} = \left(1 - \frac{\bar{n}}{n}\right) \widehat{\langle X, X \rangle}_T, \quad (4.1)$$

onde,

$$\widehat{\langle X, X \rangle}_T = [Y, Y]_T^{\text{média}} - \frac{\bar{n}}{n} [Y, Y]_T^{\text{tudo}}, \quad (4.2)$$

e

$$[Y, Y]_T^{\text{média}} = \frac{1}{K} \sum_{k=1}^K [Y, Y]_T^k. \quad (4.3)$$

Quando nos referimos a volatilidade realizada (VR), queremos dizer:

$$[Y, Y]^{\text{esparso}} = [Y, Y]^{\mathcal{H}} = \sum_{t_i, t_{i,+} \in \mathcal{H}} (Y_{t_{i,+}} - Y_{t_i})^2, \quad (4.4)$$

onde \mathcal{H} é descrito na *Definição 3.2*.

Estamos supondo como hipótese do modelo, alocação regular, ou seja que as transações são igualmente espaçadas no tempo. Isto quer dizer, que quando definimos a janela de tempo que queremos utilizar, automaticamente definimos o número de grids K e seus respectivos tamanhos \bar{n} .

4.1 Os Dados

Os dados utilizados são todas as transações de PETR4, VALE5, OGXP3 E ITUB4, as quatro ações mais líquidas da Bovespa neste período, realizadas nos pregões do dia 27/09/2010 à 08/10/2010, totalizando dez dias negociação (duas semanas). É importante salientar, que o preço dos leilões de abertura e fechamento foram excluídos,

pois são momentos em que as negociações são suspensas por um grande intervalo de tempo comparado ao tempo médio de transação das ações, as vezes atingindo mais de 15 minutos de duração.

4.2 Estimação

Primeiramente, mostramos uma comparação da VRDE e a VR nos dez dias de negociação. Para o cálculo, tanto da VRDE e da VR, utilizamos um *range* de 1 a 9 minutos, com um passo de 15 segundos, ou seja, calculamos, 33 volatilidades para cada dia, para cada estimador. Podemos notar que a VRDE é mais robusta que VR independente da ação escolhida e mesmo para aquelas que possuem menos liquidez.

Nas Figuras 4.11, 4.22, 4.33 e 4.44, estimamos a superfície de volatilidade da VRDE para os 10 dias de transação. Novamente, podemos notar a estabilidade da VRDE, independente da janela escolhida para sua estimação. Nesses gráficos fica evidente que a VRDE consegue capturar a variação na volatilidade dos papéis entre os dias de negociação.

4.2.1 VRDE para PETR4

Nas Tabelas 4.1 e 4.2 mostramos a estatística descritiva da PETR4 dos dez dias de negociação. Nas Figuras 4.1 a 4.10 representamos os gráficos da volatilidade realizada em duas escalas (VRDE) e da volatilidade realizada esparsamente (VR) anualizada para cada dia de negociação (eixo das ordenadas) para cada janela de tempo utilizada (eixo das abscissas) da PETR4, ressaltando que foram calculadas 33 volatilidades para cada dia, com respectivas janelas de 1 a 9 minutos, com um passo de 15 segundos para cada volatilidade. Na Figura 4.11 mostramos a superfície de VRDE de PETR4 (eixo Z), para as janelas de 1 a 9 minutos (eixo X), para os dez dias de negociação (eixo Y).

Estatística Descritiva PETR4	27/09	28/10	29/09	30/09	01/10
Número de transações	37.640	22.997	34.751	25.324	24.367
Tempo médio entre as transações(seg)	0,6695	1,0959	0,7252	0,9951	1,0342
Maior log-retorno das transações(%)	0,61	0,37	0,22	0,25	0,29
Menor log-retorno das transações(%)	-0,64	-0,15	-0,22	-0,22	-0,18

Tabela 4.1: Estatística descritiva de PETR4 de 27/09/2010 a 01/10/2010.

Estatística Descritiva PETR4	04/10	05/10	06/10	07/10	08/10
Número de transações	30.608	33.156	45.310	40.390	26.623
Tempo médio entre as transações(seg)	0,8233	0,7600	0,5562	0,6239	0,9465
Maior log-retorno das transações(%)	0,18	0,19	0,27	0,20	0,24
Menor log-retorno das transações(%)	-0,15	-0,19	-0,23	-0,24	-0,24

Tabela 4.2: Estatística descritiva de PETR4 de 04/2010 a 08/10/2010.

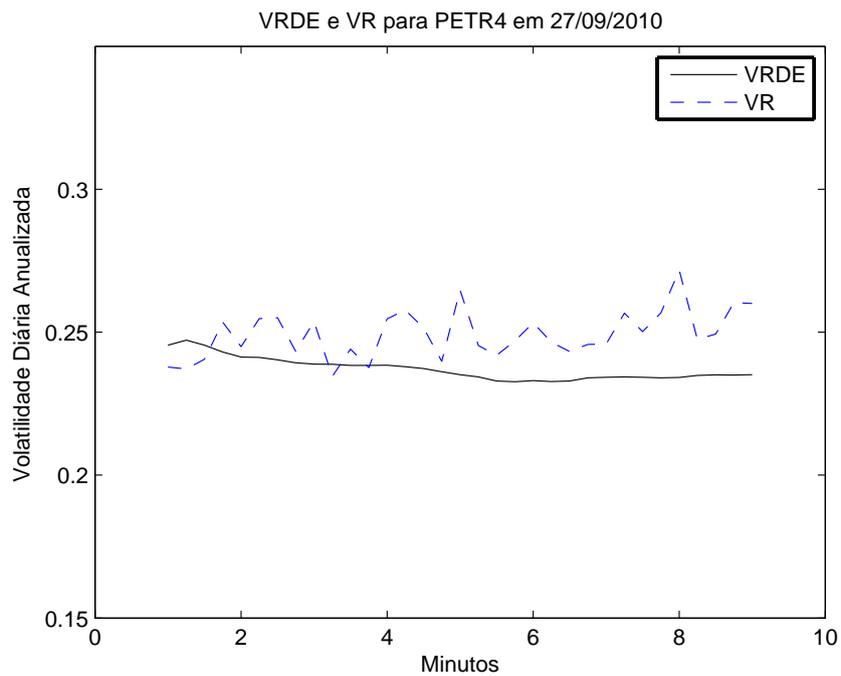


Figura 4.1: Estimação da VRDE e VR para PETR4 em 27/09/2010.

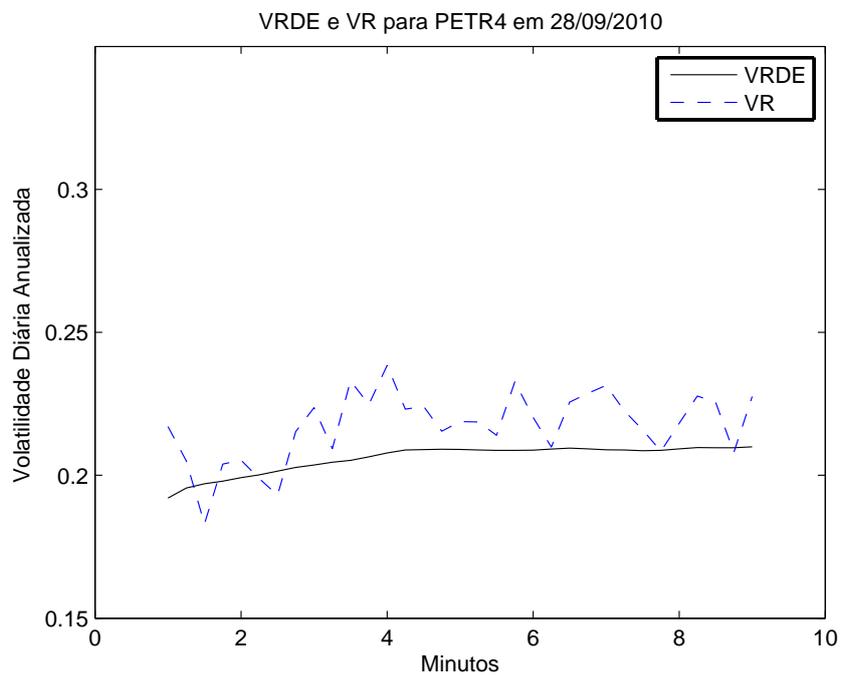


Figura 4.2: Estimação da VRDE e VR para PETR4 em 28/09/2010.

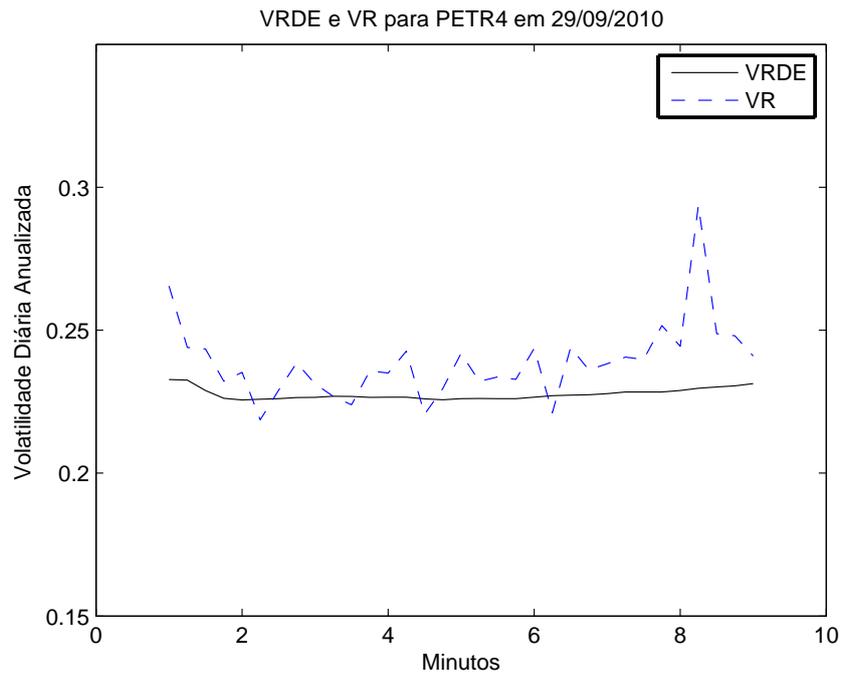


Figura 4.3: Estimação da VRDE e VR para PETR4 em 29/09/2010.

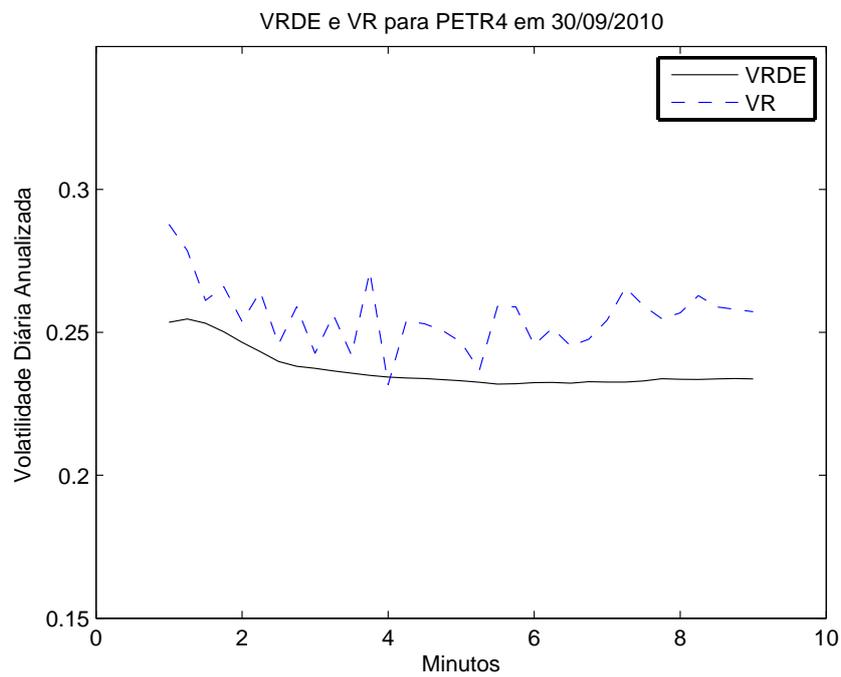


Figura 4.4: Estimação da VRDE e VR para PETR4 em 30/09/2010.

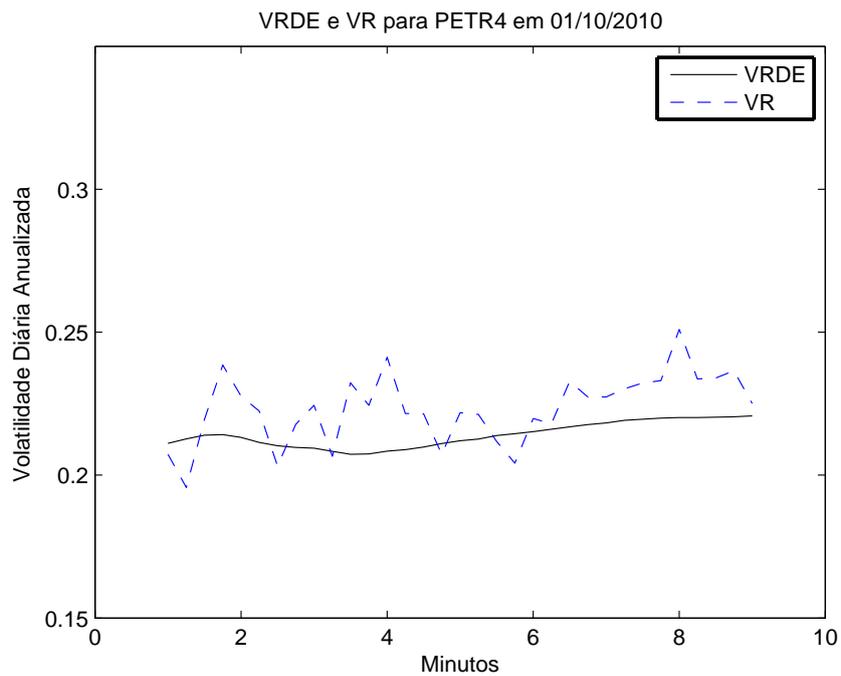


Figura 4.5: Estimação da VRDE e VR para PETR4 em 01/10/2010.

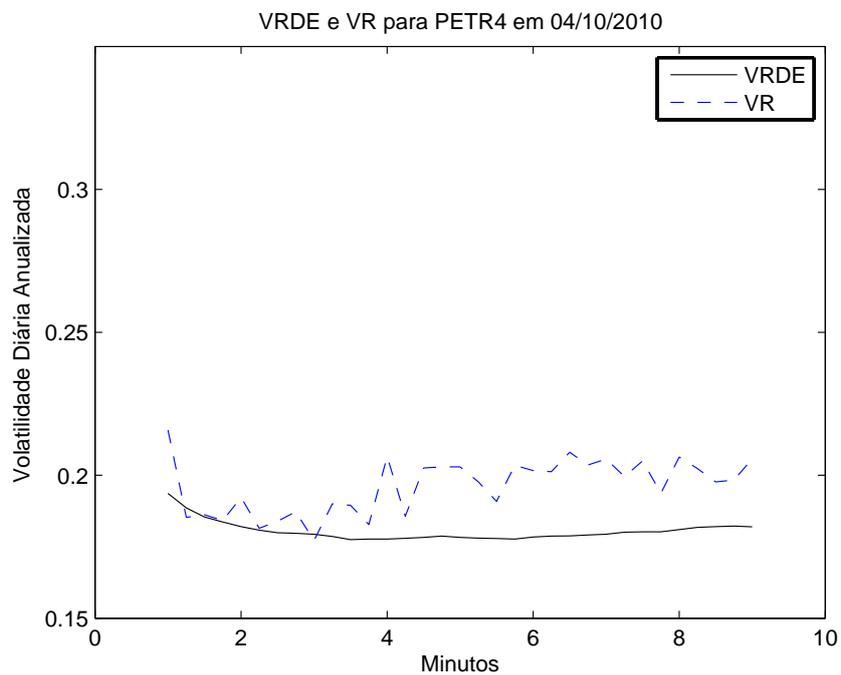


Figura 4.6: Estimação da VRDE e VR para PETR4 em 04/10/2010.

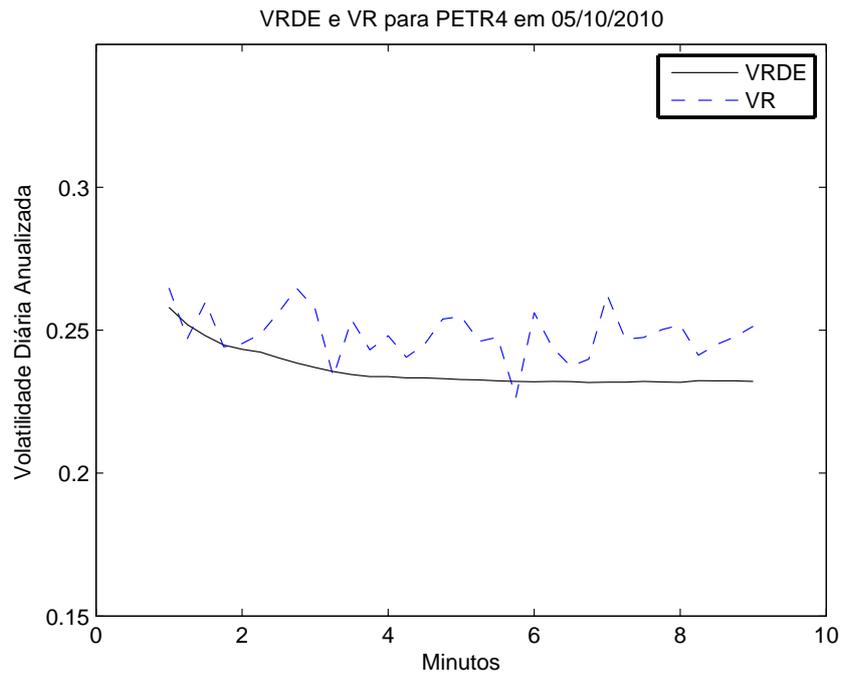


Figura 4.7: Estimação da VRDE e VR para PETR4 em 05/10/2010.

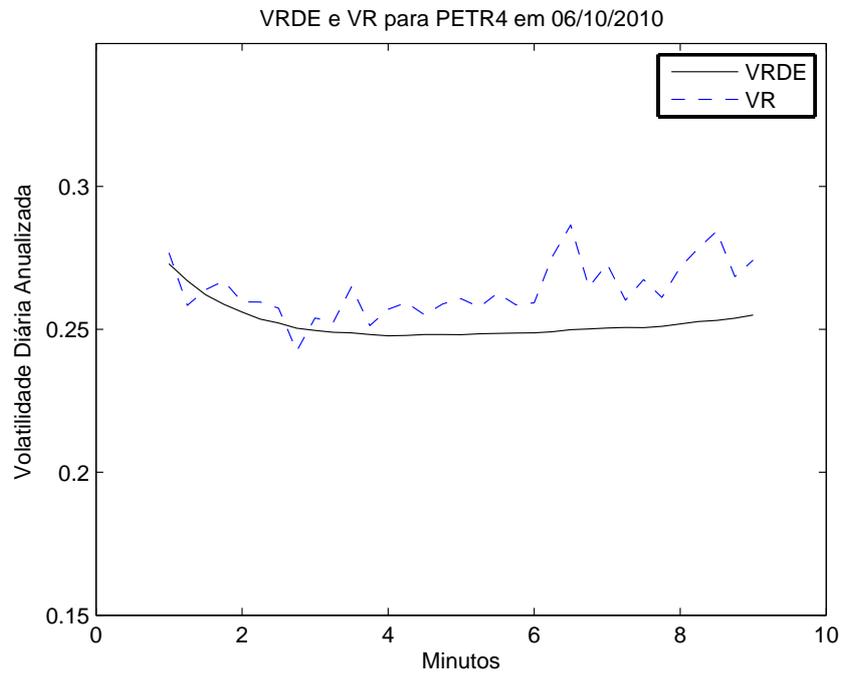


Figura 4.8: Estimação da VRDE e VR para PETR4 em 06/10/2010.

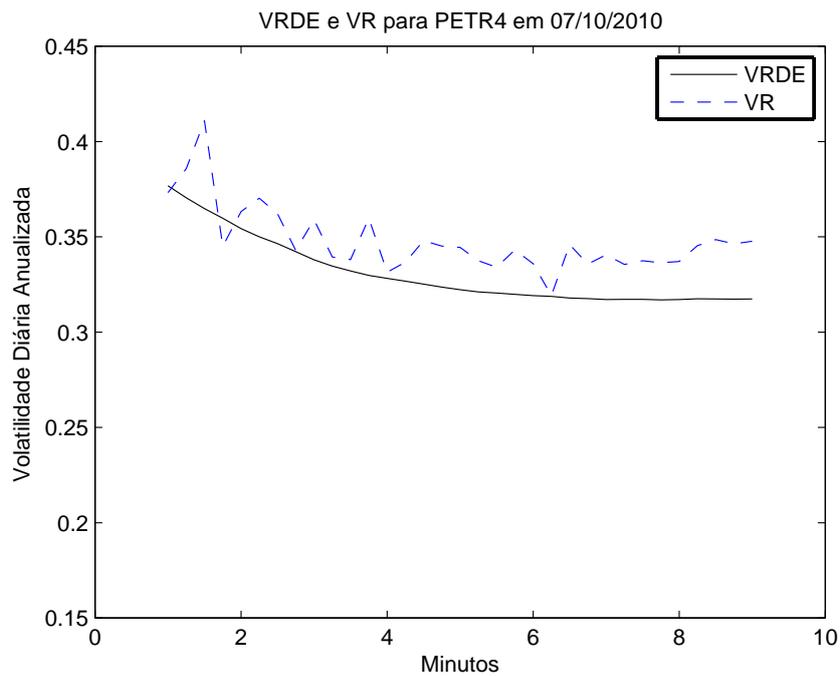


Figura 4.9: Estimação da VRDE e VR para PETR4 em 07/10/2010.

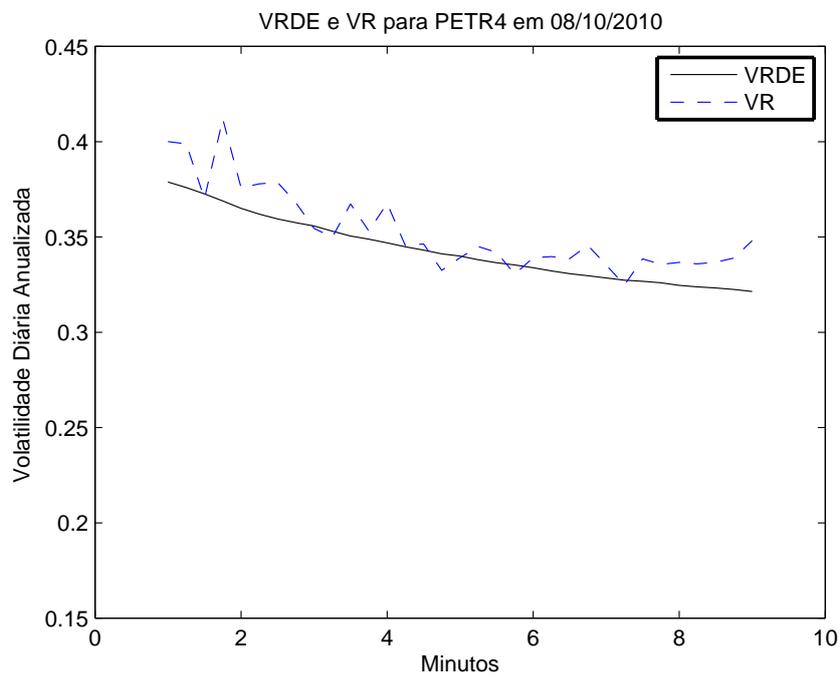


Figura 4.10: Estimação da VRDE e VR para PETR4 em 08/10/2010.

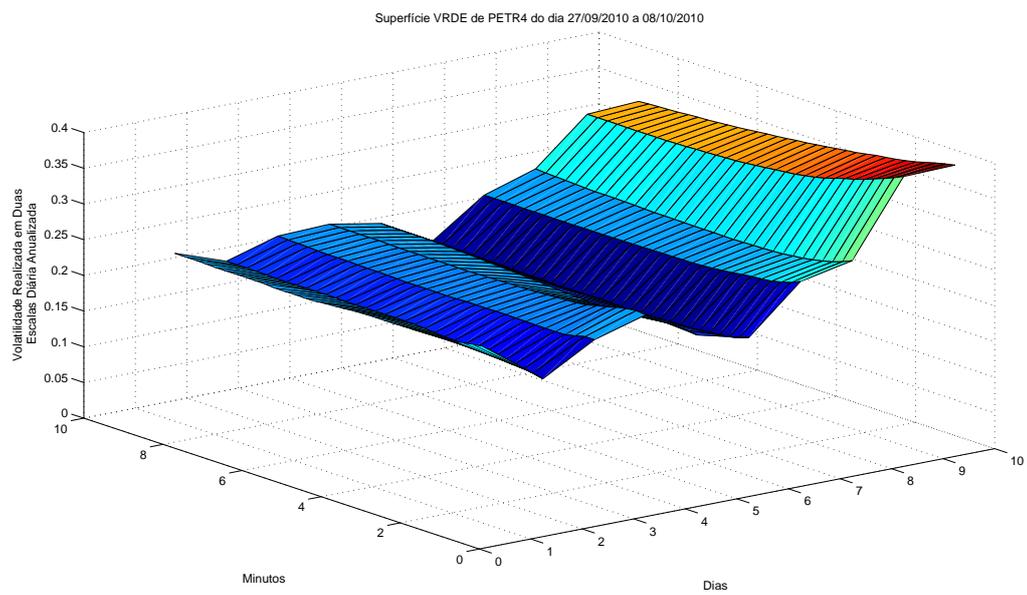


Figura 4.11: Superfície de VRDE dos 10 dias de transação e por intervalo de tempo utilizado na estimação.

4.2.2 VRDE para VALE5

Nas Tabelas 4.3 e 4.4 mostramos a estatística descritiva da VALE5 dos dez dias de negociação. Nas Figuras 4.12 a 4.21 representamos os gráficos da volatilidade realizada em duas escalas (VRDE) e da volatilidade realizada esparsamente (VR) anualizada para cada dia de negociação (eixo das ordenadas) para cada janela de tempo utilizada (eixo das abscissas) da VALE5, ressaltando que foram calculadas 33 volatilidades para cada dia, com respectivas janelas de 1 a 9 minutos, com um passo de 15 segundos para cada volatilidade. Na Figura 4.22 mostramos a superfície de VRDE de VALE5 (eixo Z), para as janelas de 1 a 9 minutos (eixo X), para os dez dias de negociação (eixo Y).

Estatística Descritiva VALE5	27/09	28/10	29/09	30/09	01/10
Número de transações	25.646	16.880	17.701	19.377	18.904
Tempo médio entre as transações(seg)	0,9826	1,4929	1,4236	1,3005	1,3331
Maior log-retorno das transações(%)	0,18	0,31	0,13	0,19	0,11
Menor log-retorno das transações(%)	-0,20	-0,15	-0,15	-0,17	-0,17

Tabela 4.3: Estatística descritiva de VALE5 de 27/09/2010 a 01/10/2010

Estatística Descritiva VALE5	04/10	05/10	06/10	07/10	08/10
Número de transações	14.335	19.221	22.780	21.478	20.114
Tempo médio entre as transações(seg)	1,7579	0,7600	1,3111	1,1062	1,1733
Maior log-retorno das transações(%)	0,15	0,19	0,15	0,15	0,13
Menor log-retorno das transações(%)	-0,15	-0,19	-0,30	-0,19	-0,21

Tabela 4.4: Estatística descritiva de VALE5 de 04/10/2010 a 08/10/2010.

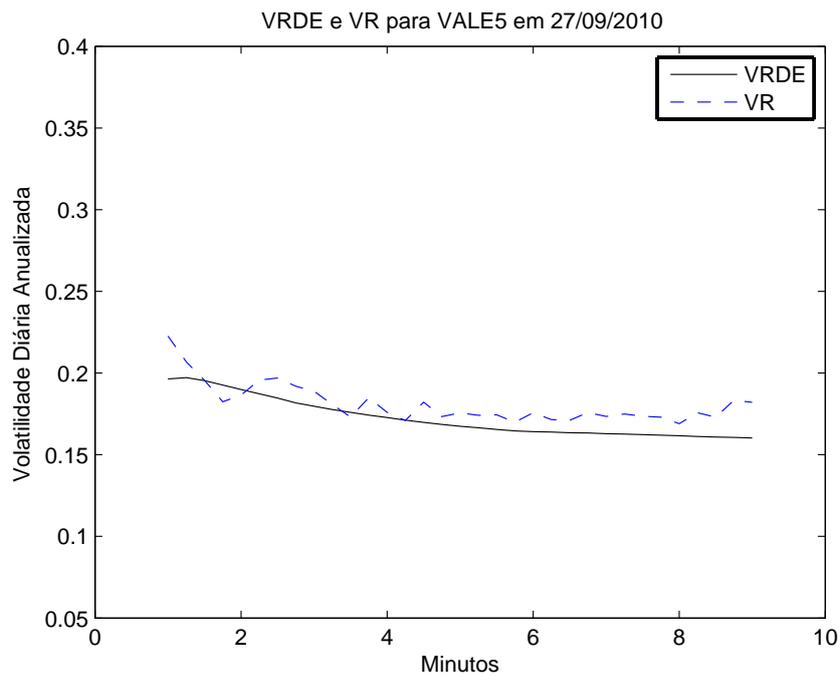


Figura 4.12: Estimação da VRDE e VR para VALE5 em 27/09/2010.

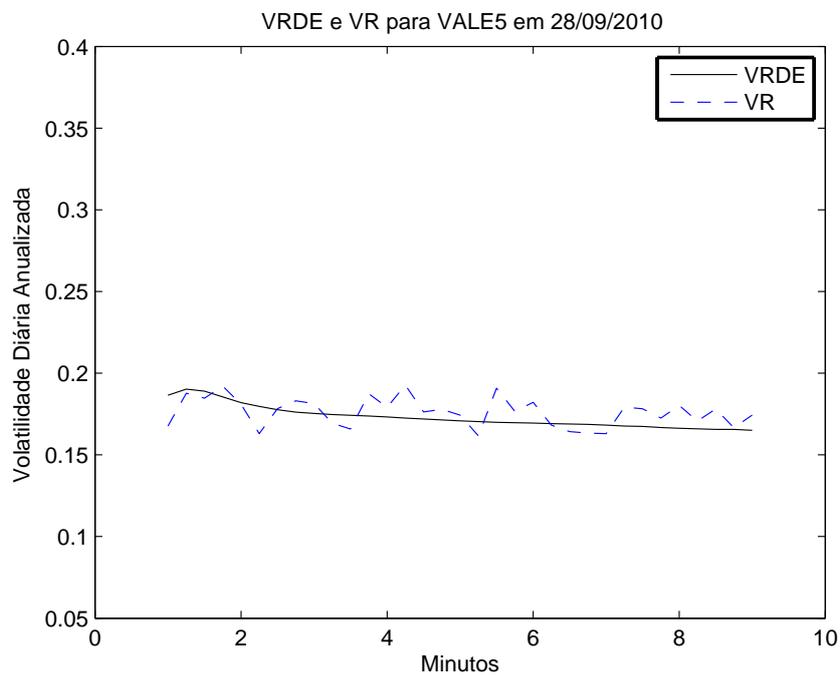


Figura 4.13: Estimação da VRDE e VR para VALE5 em 28/09/2010.

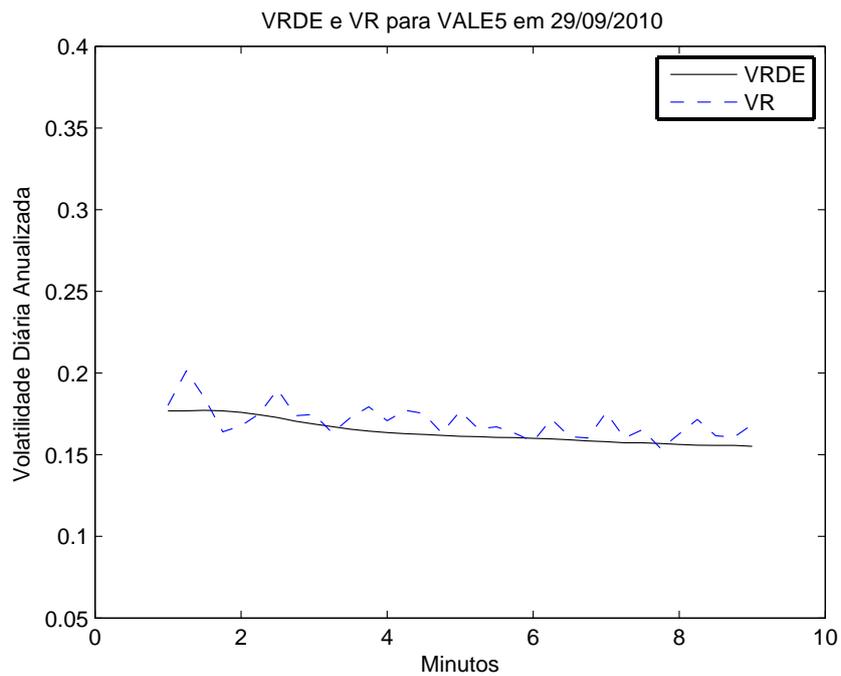


Figura 4.14: Estimação da VRDE e VR para VALE5 em 29/09/2010.

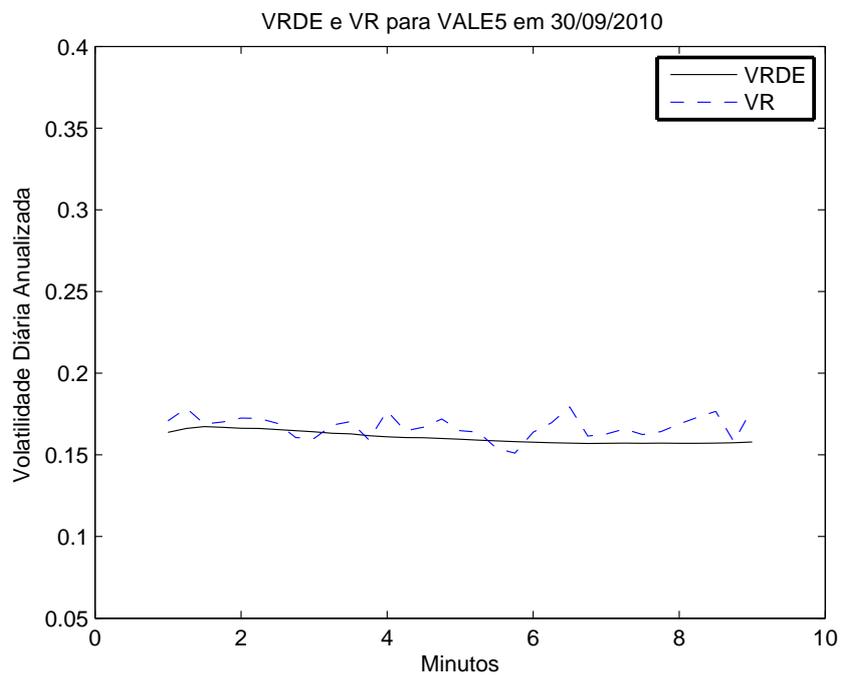


Figura 4.15: Estimação da VRDE e VR para VALE5 em 30/09/2010.

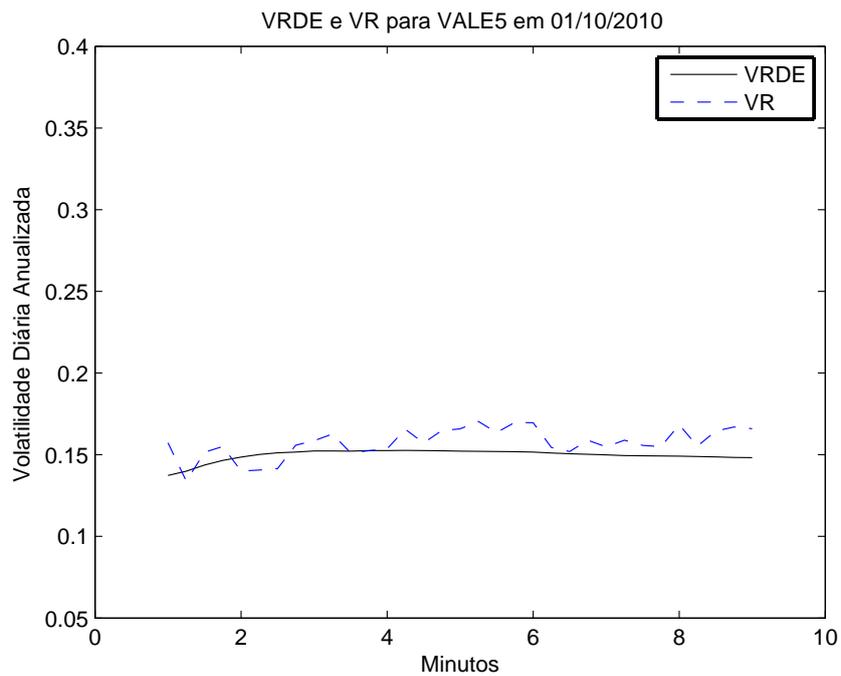


Figura 4.16: Estimação da VRDE e VR para VALE5 em 01/10/2010.

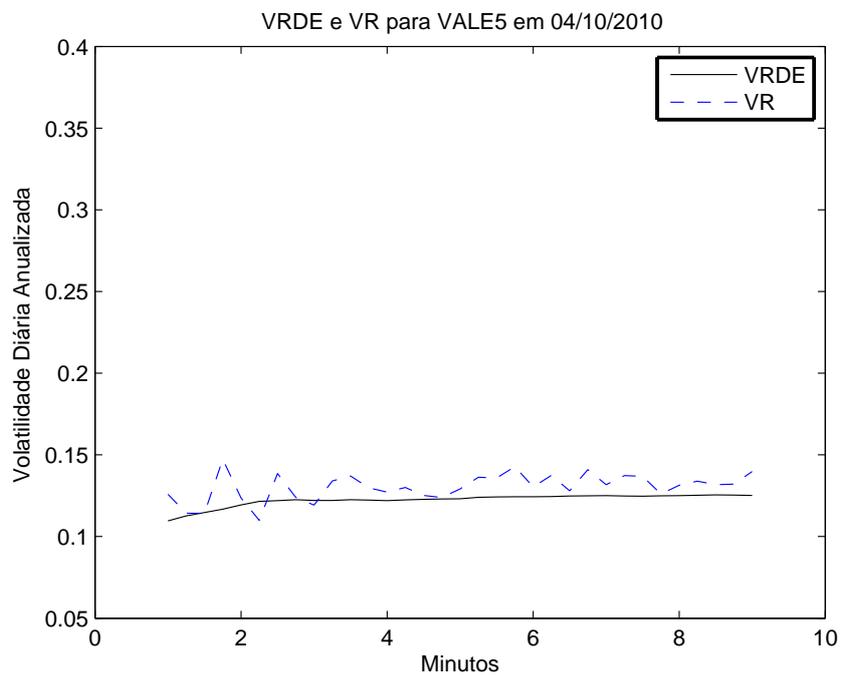


Figura 4.17: Estimação da VRDE e VR para VALE5 em 04/10/2010.

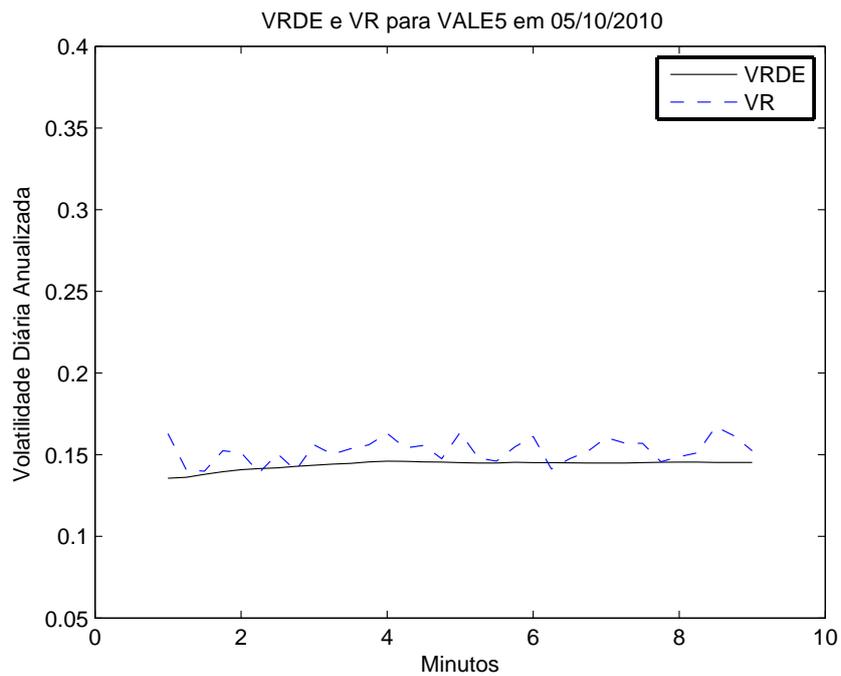


Figura 4.18: Estimação da VRDE e VR para VALE5 em 05/10/2010.

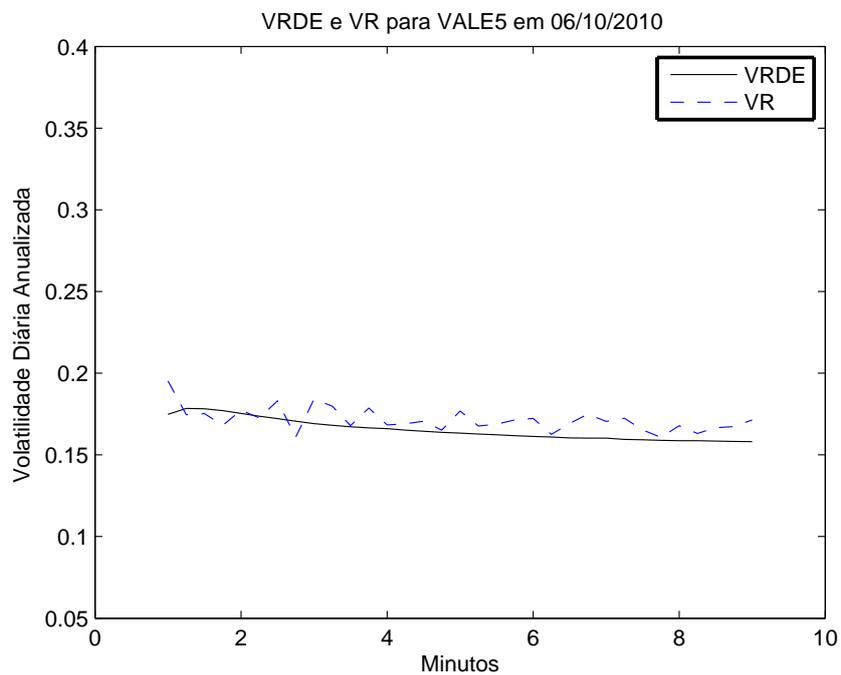


Figura 4.19: Estimação da VRDE e VR para VALE5 em 06/10/2010.

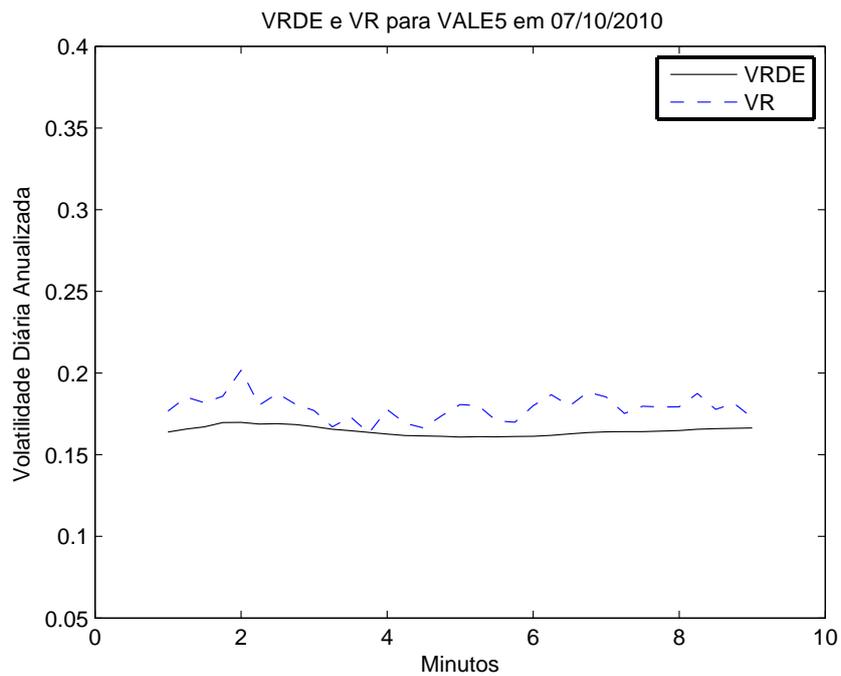


Figura 4.20: Estimação da VRDE e VR para VALE5 em 07/10/2010.

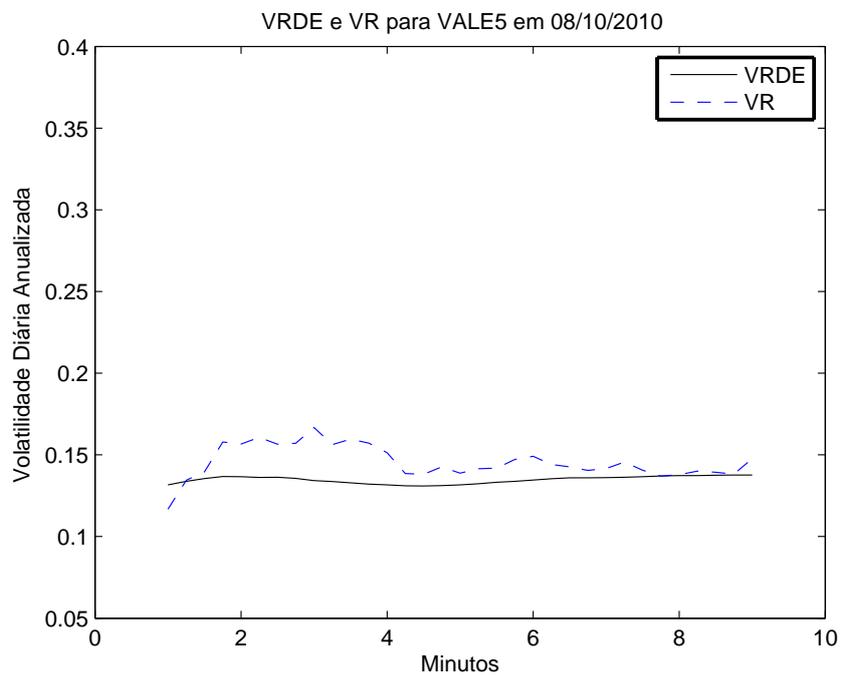


Figura 4.21: Estimação da VRDE e VR para VALE5 em 08/10/2010.

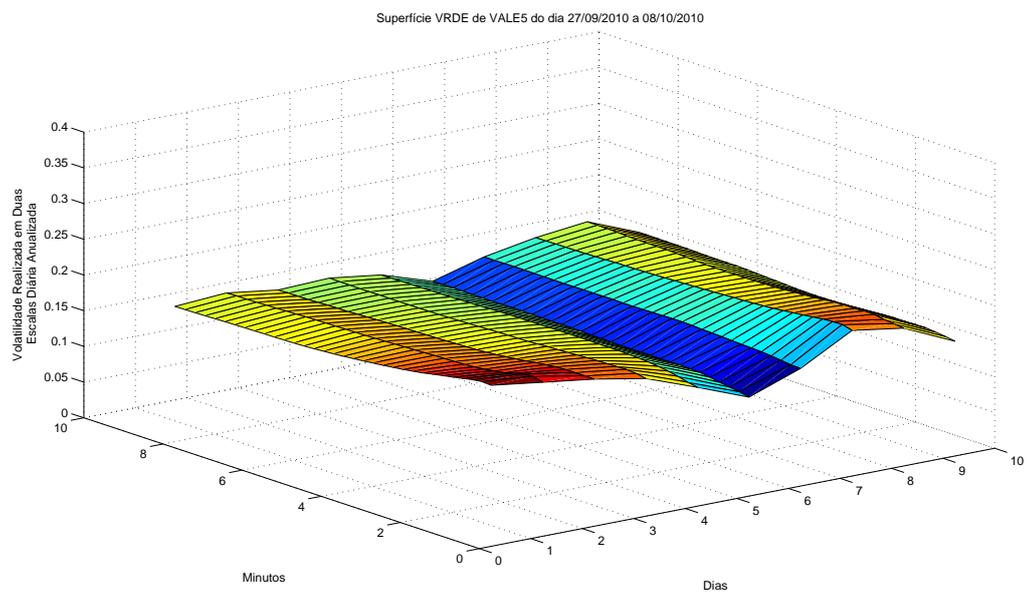


Figura 4.22: Superfície de VRDE de VALE 5 em 10 dias de negociação e por intervalo de tempo utilizado na estimação.

4.2.3 VRDE para OGXP3

Nas Tabelas 4.5 e 4.6 mostramos a estatística descritiva da OGXP3 dos dez dias de negociação. Nas Figuras 4.23 a 4.32 representamos os gráficos da volatilidade realizada em duas escalas (VRDE) e da volatilidade realizada esparsamente (VR) analisada para cada dia de negociação (eixo das ordenadas) para cada janela de tempo utilizada (eixo das abscissas) da OGXP3, ressaltando que foram calculadas 33 volatilidades para cada dia, com respectivas janelas de 1 a 9 minutos, com um passo de 15 segundos para cada volatilidade. Na Figura 4.33 mostramos a superfície de VRDE de VALE5 (eixo Z), para as janelas de 1 a 9 minutos (eixo X), para os dez dias de negociação (eixo Y).

Estatística Descritiva OGXP3	27/09	28/10	29/09	30/09	01/10
Número de transações	7.297	7.678	15.394	15.663	19.338
Tempo médio entre as transações(seg)	3,4535	3,2821	1,6370	1,6089	1,3031
Maior log-retorno das transações(%)	0,44	0,25	0,38	0,27	0,45
Menor log-retorno das transações(%)	-0,39	-0,25	-0,43	-0,45	-0,45

Tabela 4.5: Estatística descritiva de OGXP3 de 27/09/2010 a 01/10/2010.

Estatística Descritiva OGXP3	04/10	05/10	06/10	07/10	08/10
Número de transações	28.397	14.774	14.114	9.269	8.407
Tempo médio entre as transações(seg)	0,8874	1,7057	1,7855	2,7187	2,9975
Maior log-retorno das transações(%)	0,41	0,26	0,48	0,44	0,39
Menor log-retorno das transações(%)	-0,31	-0,26	-0,22	-0,31	-0,30

Tabela 4.6: Estatística descritiva de OGXP3 de 04/10/2010 a 08/10/2010.

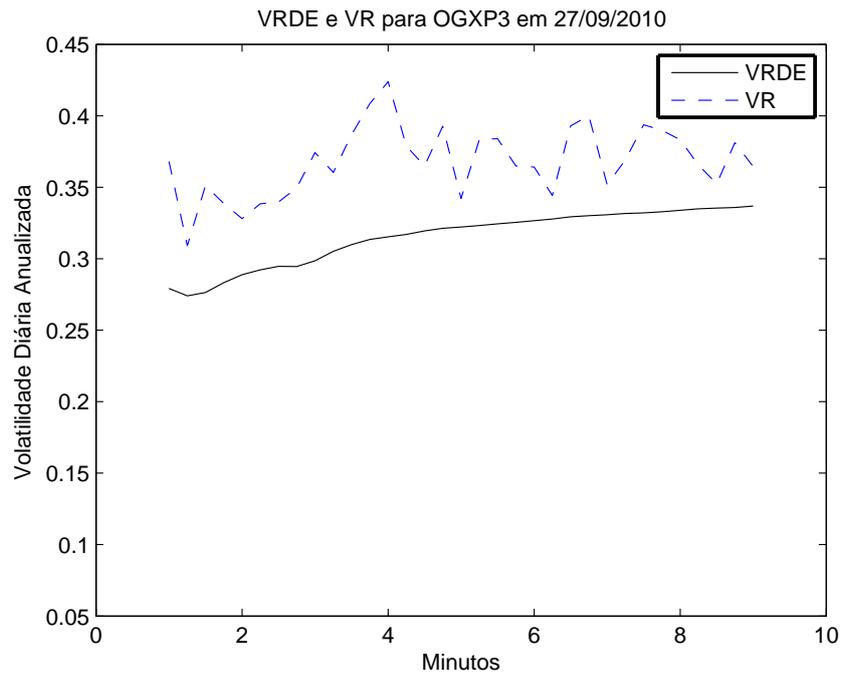


Figura 4.23: Estimação da VRDE e VR para OGXP3 em 27/09/2010.

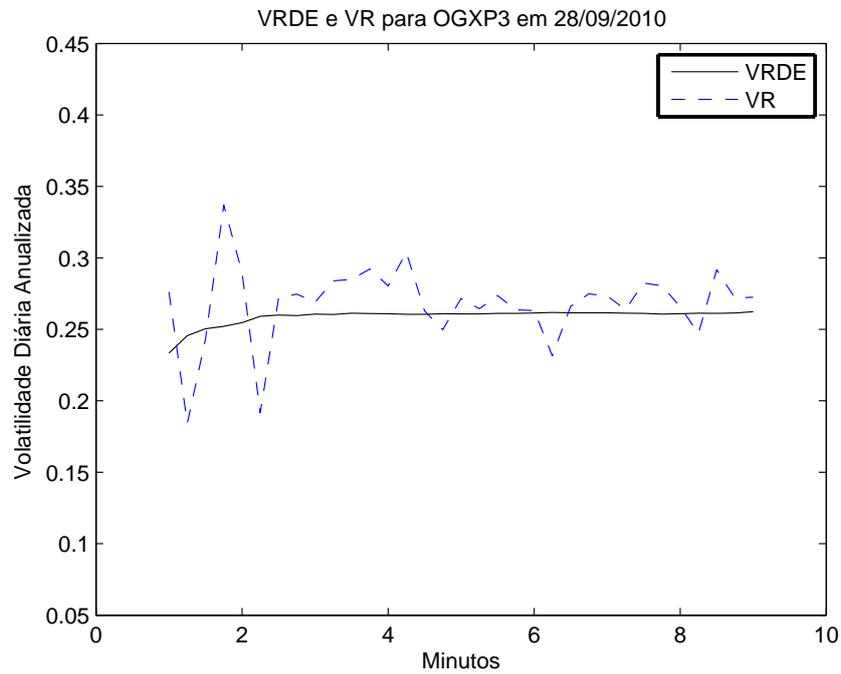


Figura 4.24: Estimação da VRDE e VR para OGXP3 em 28/09/2010.

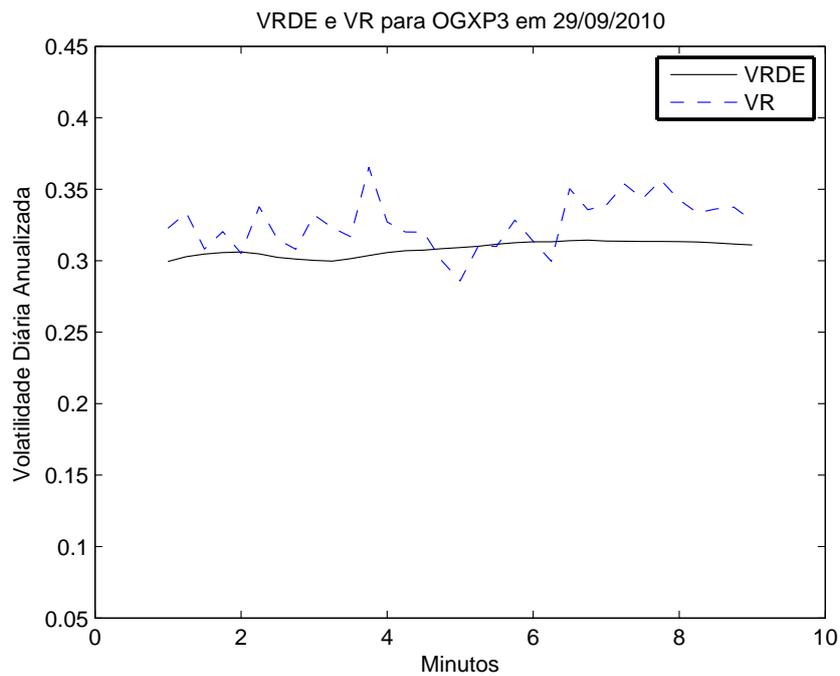


Figura 4.25: Estimação da VRDE e VR para OGXP3 em 29/09/2010.

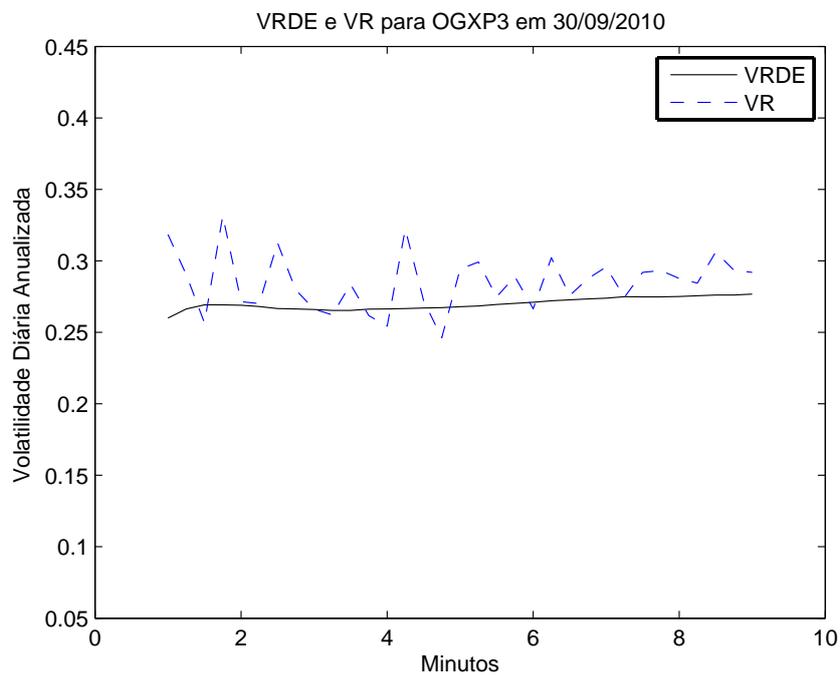


Figura 4.26: Estimação da VRDE e VR para OGXP3 em 30/09/2010.

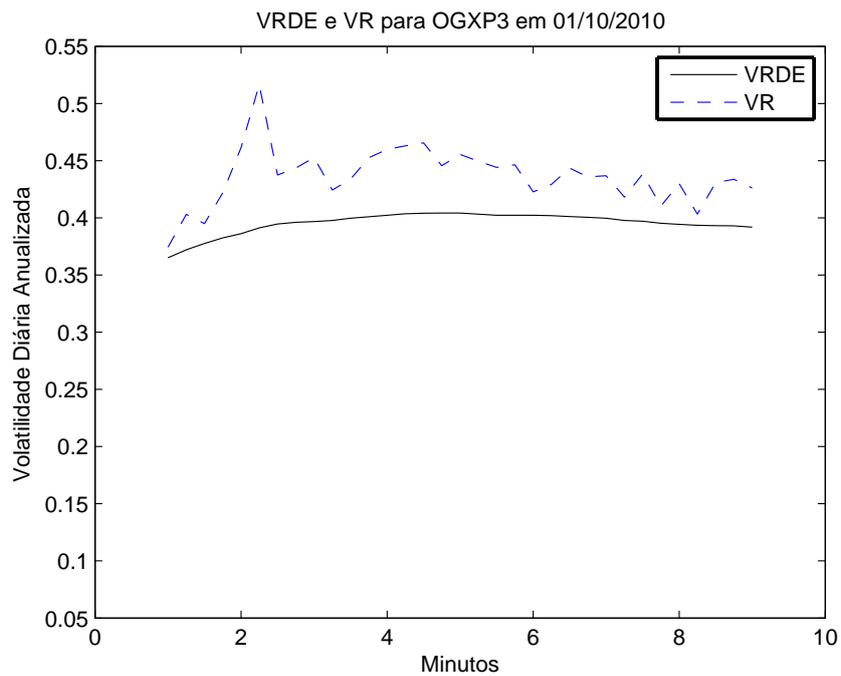


Figura 4.27: Estimação da VRDE e VR para OGXP3 em 01/10/2010.

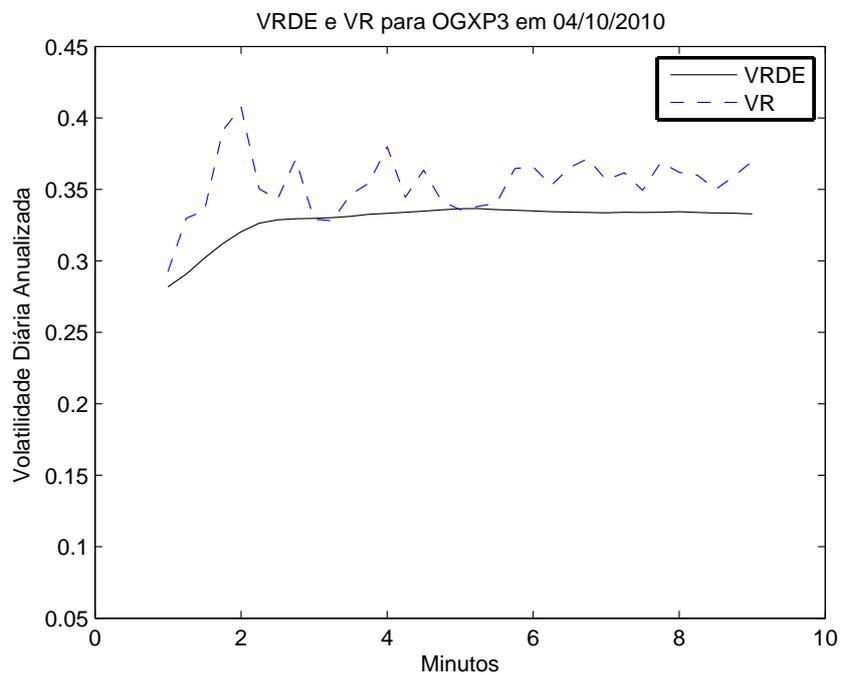


Figura 4.28: Estimação da VRDE e VR para OGXP3 em 04/10/2010.

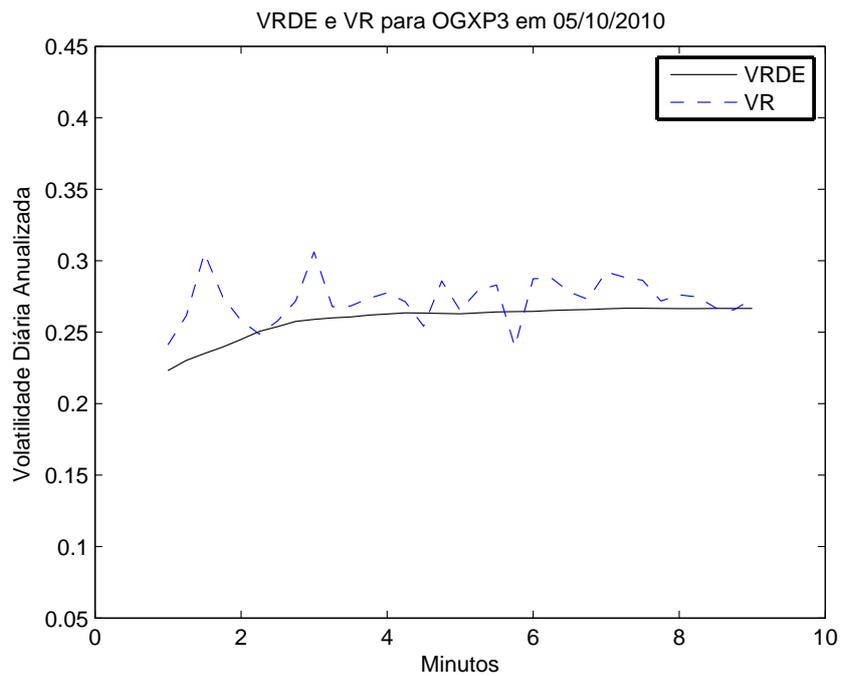


Figura 4.29: Estimação da VRDE e VR para OGXP3 em 05/10/2010.

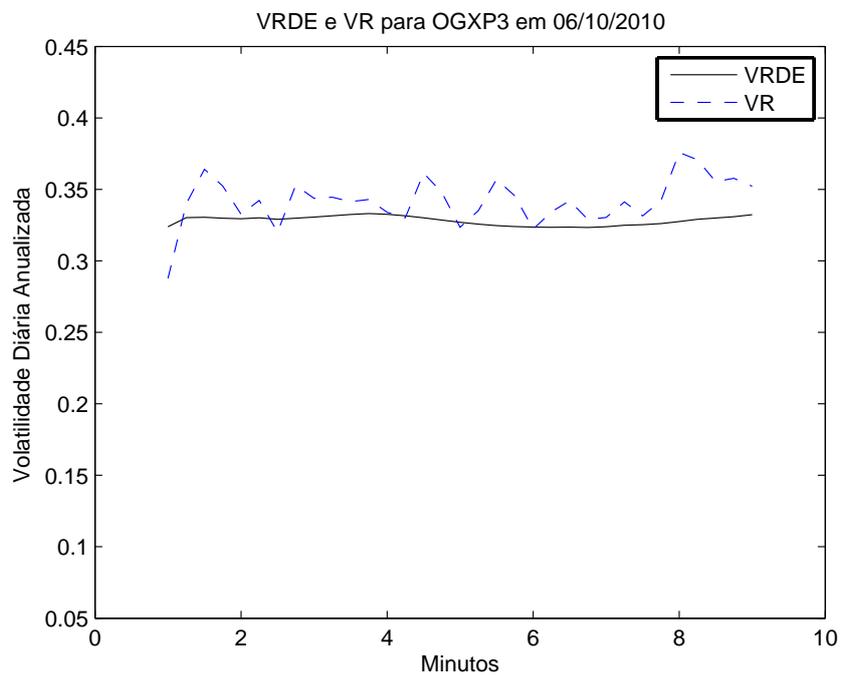


Figura 4.30: Estimação da VRDE e VR para OGXP3 em 06/10/2010.

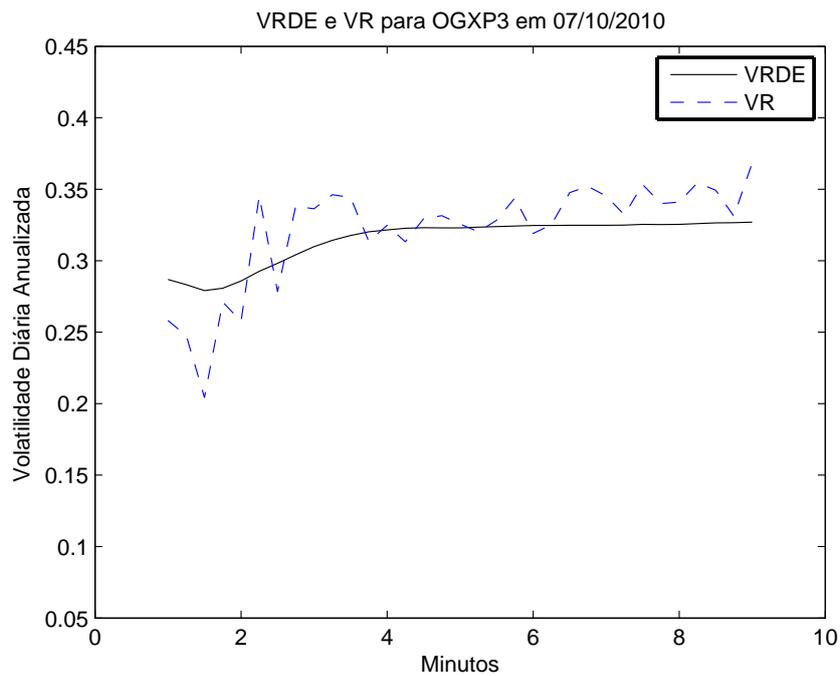


Figura 4.31: Estimação da VRDE e VR para OGXP3 em 07/10/2010.

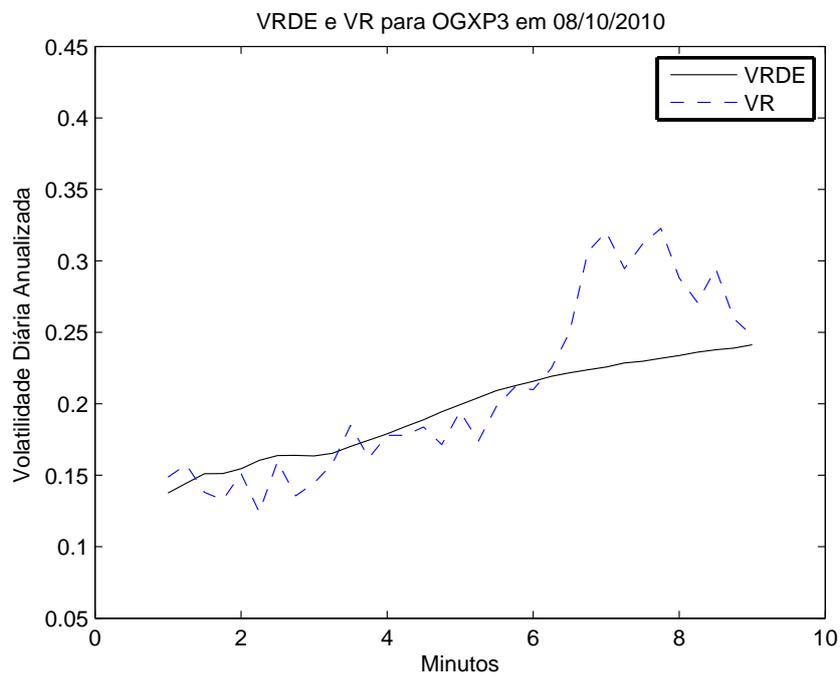


Figura 4.32: Estimação da VRDE e VR para OGXP3 em 08/10/2010.

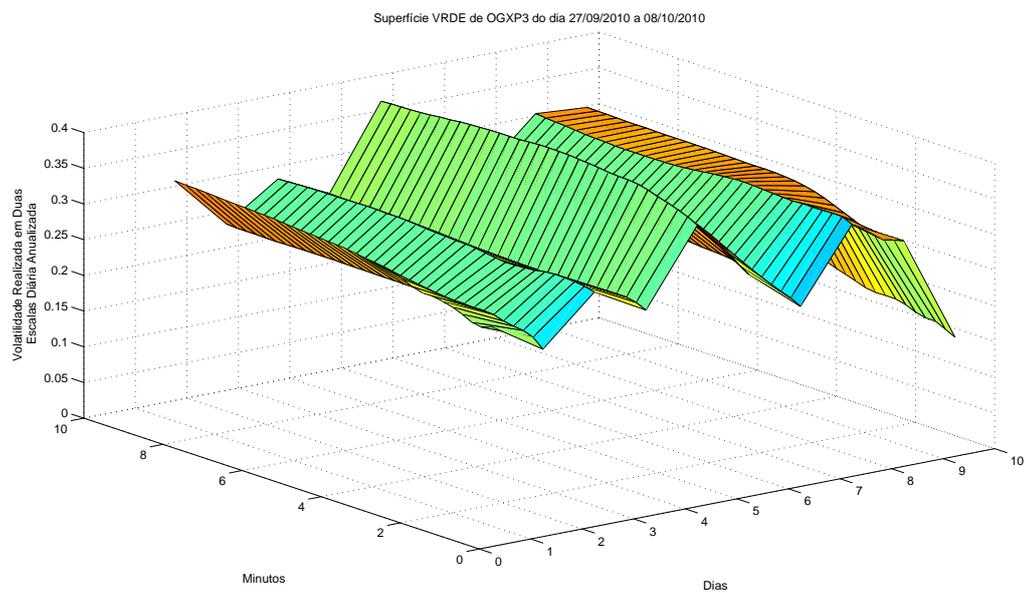


Figura 4.33: Superfície de VRDE dos 10 dias de transação e por intervalo de tempo utilizado na estimação.

4.2.4 VRDE para ITUB4

Nas Tabelas 4.7 e 4.8 mostramos a estatística descritiva da ITUB4 dos dez dias de negociação. Nas Figuras 4.34 a 4.43 representamos os gráficos da volatilidade realizada em duas escalas (VRDE) e da volatilidade realizada esparsamente (VR) anualizada para cada dia de negociação (eixo das ordenadas) para cada janela de tempo utilizada (eixo das abscissas) da ITUB4, ressaltando que foram calculadas 33 volatilidades para cada dia, com respectivas janelas de 1 a 9 minutos, com um passo de 15 segundos para cada volatilidade. Na Figura 4.44 mostramos a superfície de VRDE de VALE5 (eixo Z), para as janelas de 1 a 9 minutos (eixo X), para os dez dias de negociação (eixo Y).

Estatística Descritiva ITUB4	27/09	28/10	29/09	30/09	01/10
Número de transações	6.844	9.167	12.702	19.377	11.869
Tempo médio entre as transações(seg)	3,6821	2,7490	1,9839	1,3005	2,1232
Maior log-retorno das transações(%)	0,34	0,31	0,44	0,19	0,15
Menor log-retorno das transações(%)	-0,18	-0,21	-0,44	-0,17	-0,15

Tabela 4.7: Estatística descritiva de ITUB4 de 27/09/2010 a 01/10/2010.

Estatística Descritiva ITUB4	04/10	05/10	06/10	07/10	08/10
Número de transações	8.277	9.826	13.632	9.033	8.459
Tempo médio entre as transações(seg)	3,0446	2,5646	1,8486	2,7898	2,9791
Maior log-retorno das transações(%)	0,20	0,20	0,36	0,46	0,31
Menor log-retorno das transações(%)	-0,20	-0,27	-0,26	-0,60	-0,31

Tabela 4.8: Estatística descritiva de ITUB4 de 04/10/2010 a 08/10/2010.

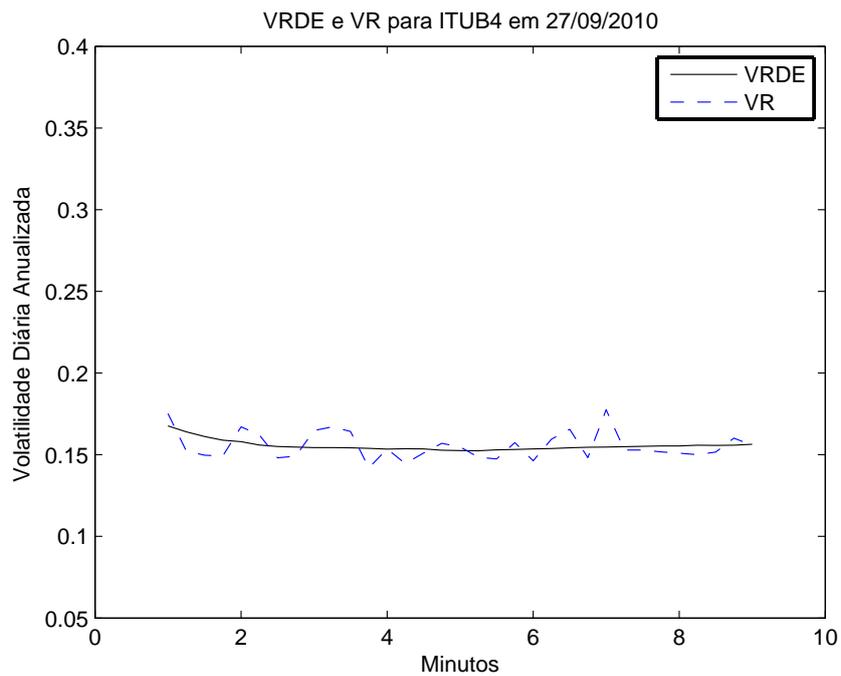


Figura 4.34: Estimação da VRDE e VR para ITUB4 em 27/09/2010.

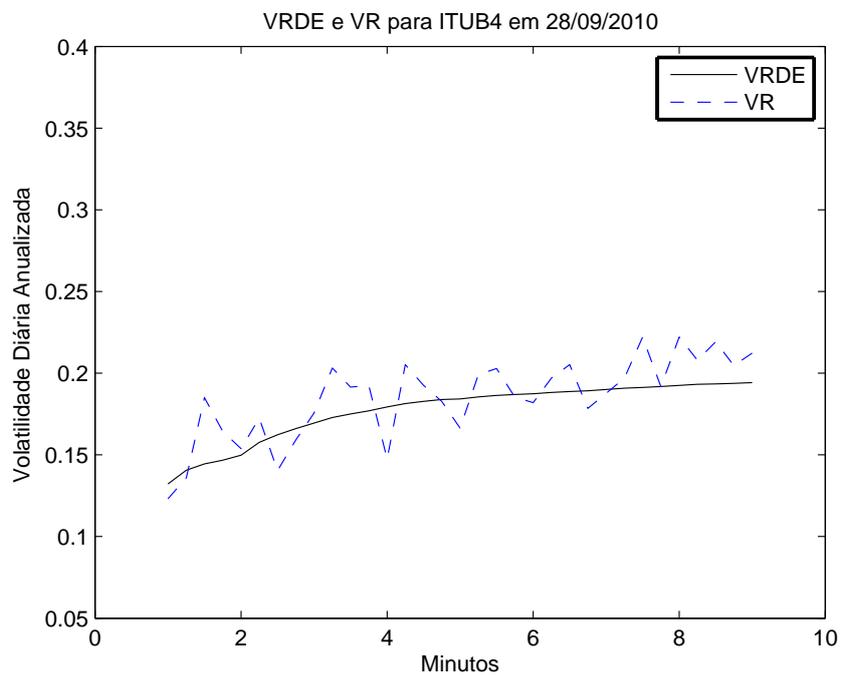


Figura 4.35: Estimação da VRDE e VR para ITUB4 em 28/09/2010.

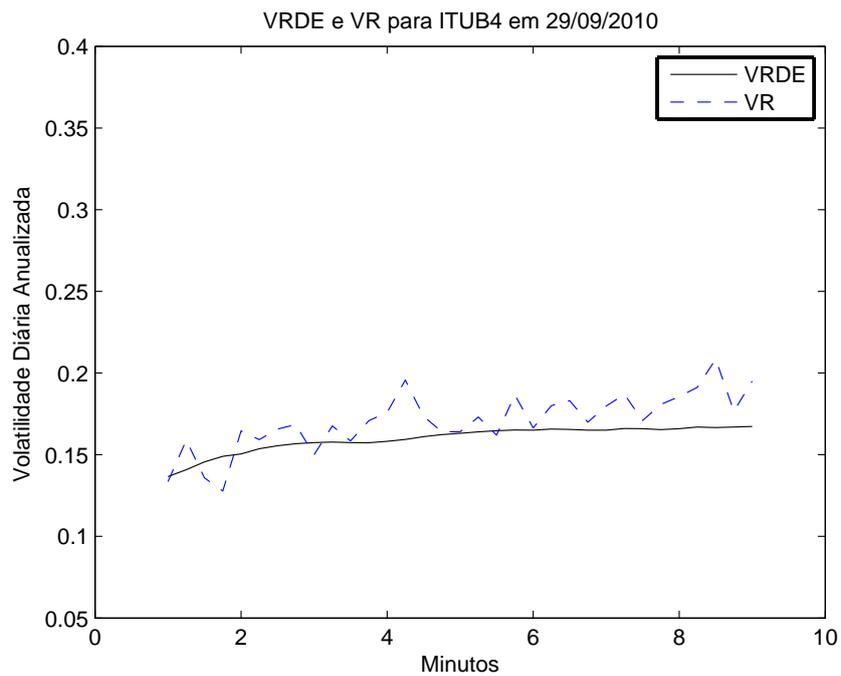


Figura 4.36: Estimação da VRDE e VR para ITUB4 em 29/09/2010.

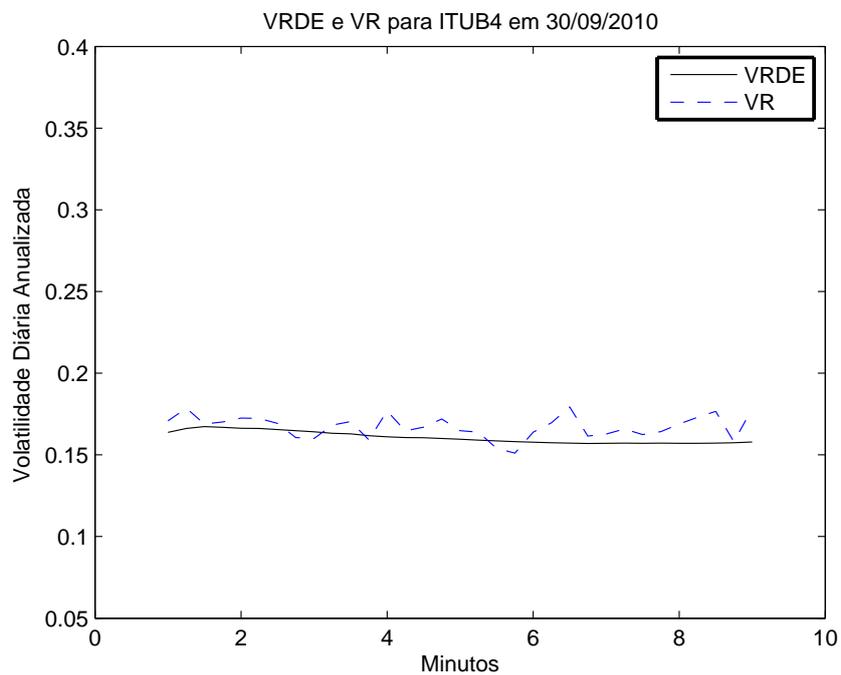


Figura 4.37: Estimação da VRDE e VR para ITUB4 em 30/09/2010.

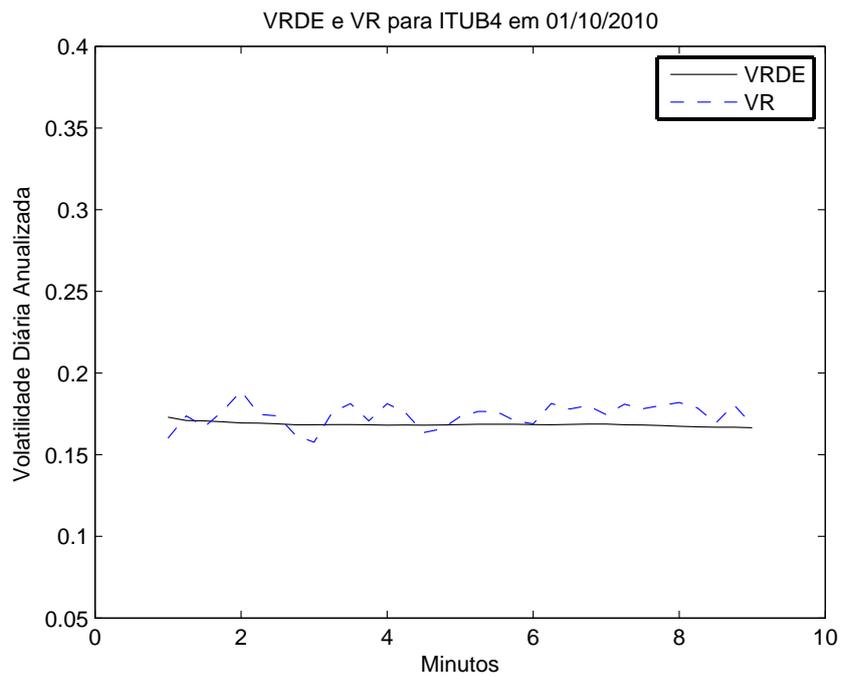


Figura 4.38: Estimação da VRDE e VR para ITUB4 em 01/10/2010.

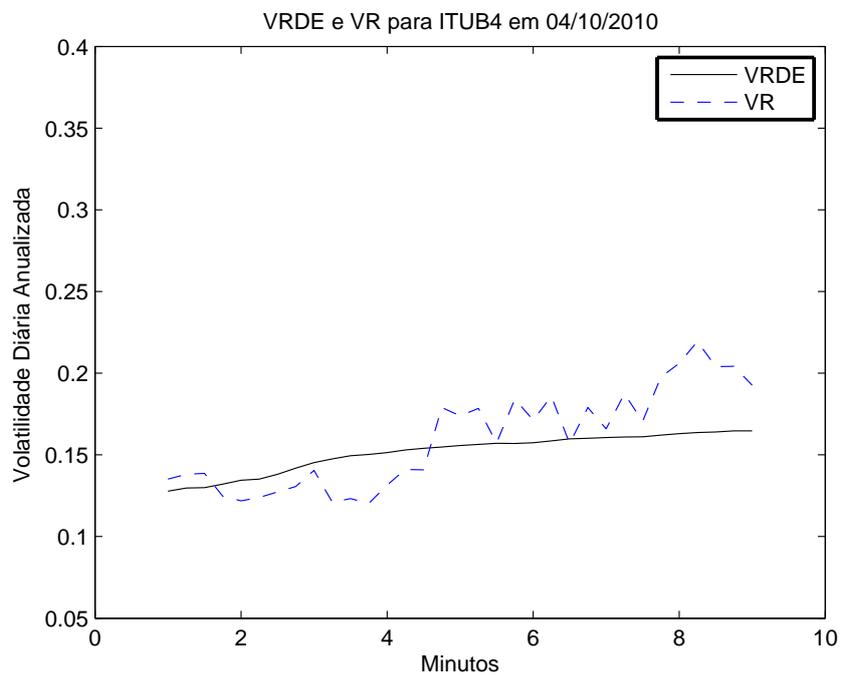


Figura 4.39: Estimação da VRDE e VR para ITUB4 em 04/10/2010.

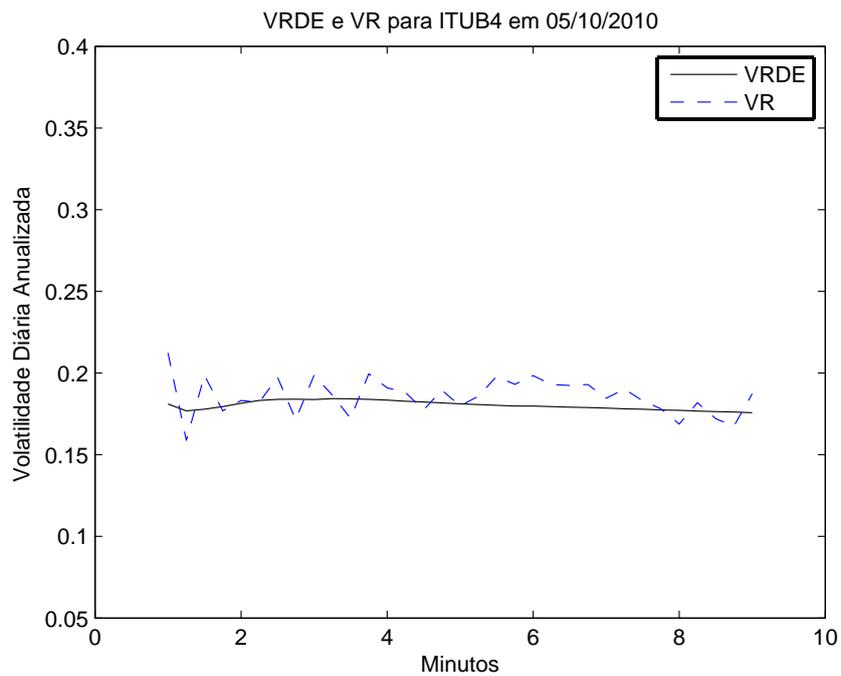


Figura 4.40: Estimação da VRDE e VR para ITUB4 em 05/10/2010.

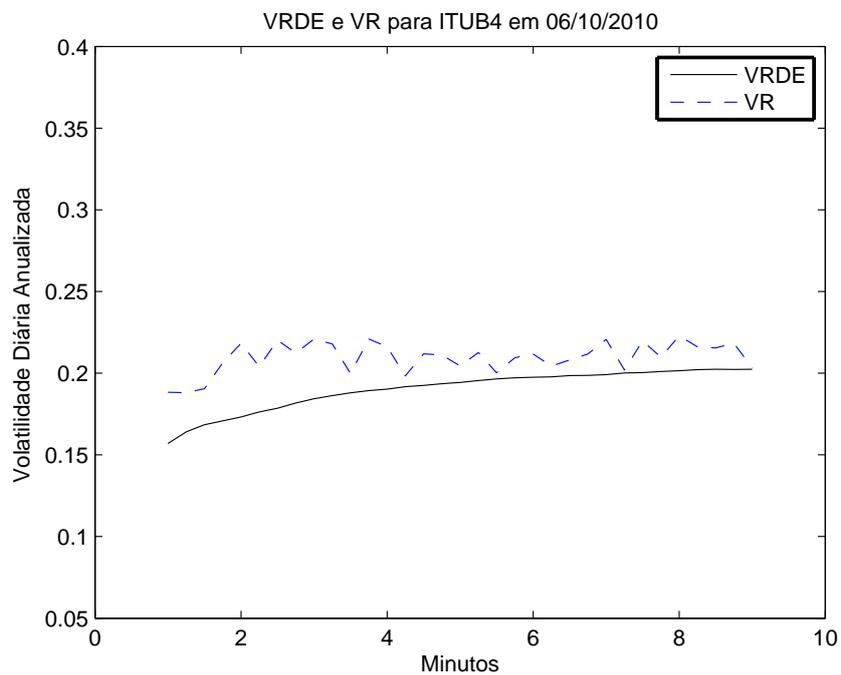


Figura 4.41: Estimação da VRDE e VR para ITUB4 em 06/10/2010.

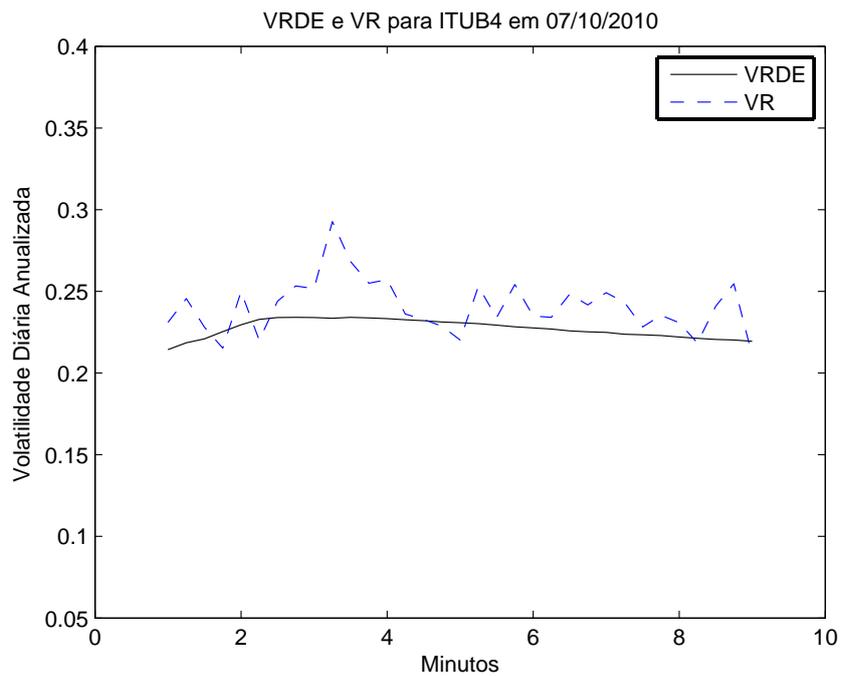


Figura 4.42: Estimação da VRDE e VR para ITUB4 em 07/10/2010.

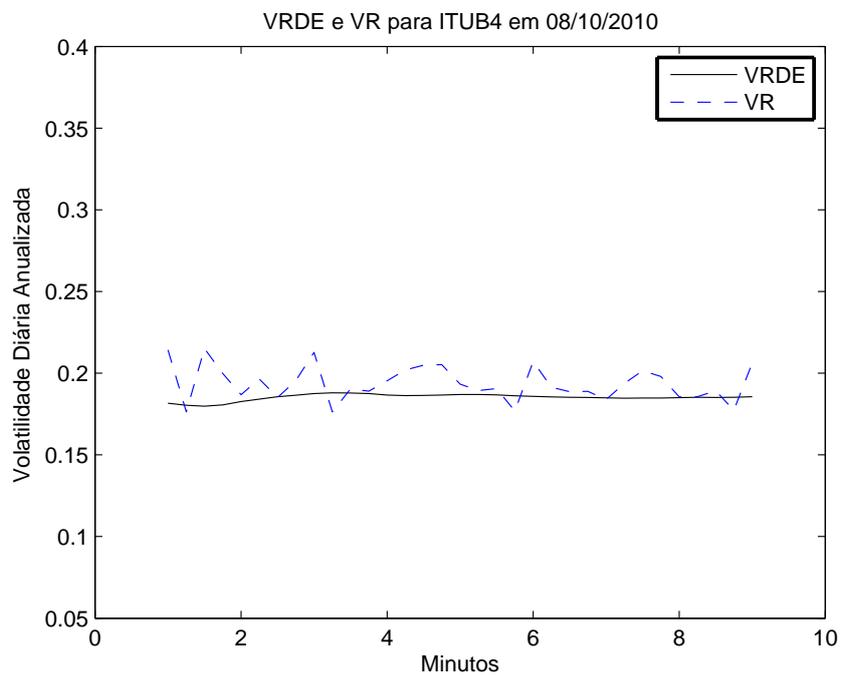


Figura 4.43: Estimação da VRDE e VR para ITUB4 em 08/10/2010.

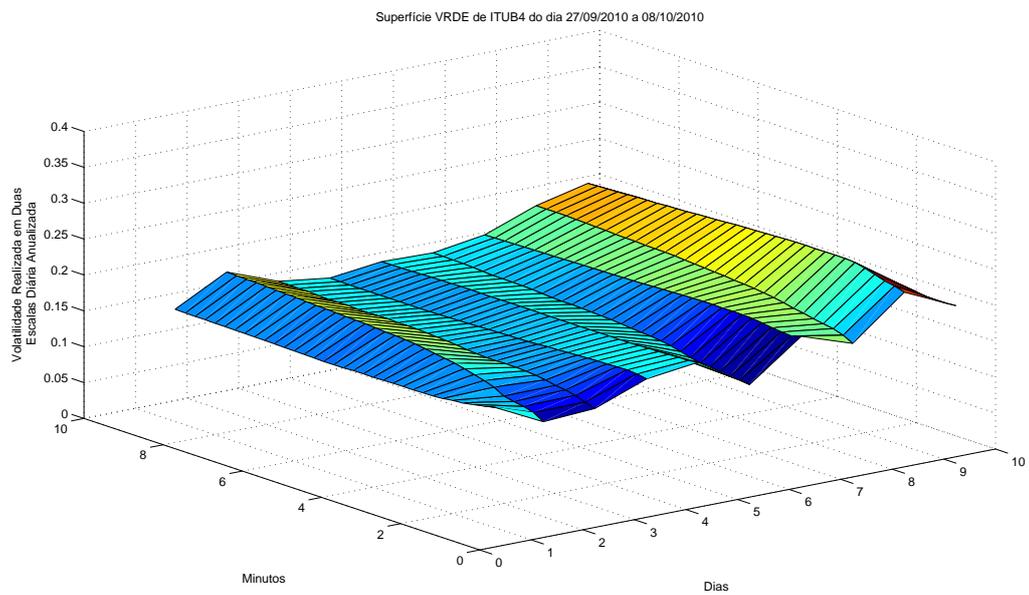


Figura 4.44: Superfície de VRDE dos 10 dias de transação e por intervalo de tempo utilizado na estimação.

Os dias escolhidos na amostra foram propositais por sucederem o último IPO (*initial public offering*) da Petrobras, quando na segunda semana, os preços da ação caíram drasticamente e houve um grande aumento na volatilidade do papel. A OGXP3 por ser uma empresa do mesmo setor da Petrobras, acabou sofrendo influência da mudança do nível de volatilidade da PETR4 nesse período. Dessa forma podemos comparar o desempenho da VRDE e da VR em um período de alta volatilidade para PETR4 e OGXP3 versus o da VALE5 e ITUB4, que não apresentaram nenhum evento relevante que alterasse expressivamente suas respectivas volatilidades no mesmo período.

Podemos notar que para PETR4 e OGXP3, o erro no cálculo da VRDE em função da janela escolhida é significativamente menor, em especial nas Figuras (4.3), (4.24) e (4.31). Para a VALE5 e ITUB4, a VRDE também possui um desempenho melhor que a VR, no entanto, seria aceitável utilizar a estimação da VR para essas ações nesses períodos.

Temos que destacar que em alguns casos, a VRDE apresentou viés significativo como em (4.10), (4.23) e (4.32). Isto pode ter acontecido em dias que os retornos eram significativamente independentes e nosso modelo de ruído não foi suficiente para representar toda microestrutura desses dias.

Capítulo 5

Aplicação do Estimador de VRDE para Delta Hedge em Modelo de Um Fator

Se considerarmos o problema de estimar a variância de um ativo conhecido tanto podemos estimar a variância passada do processo de preço, como podemos estimá-la via um modelo de apreçamento de derivativo. A primeira abordagem é formalmente chamada de volatilidade realizada ou histórica, e a segunda é a volatilidade implícita. No modelo de Black-Scholes-Merton, ambas volatilidades representam o mesmo objeto, as variações dos retornos do ativo objeto. Entretanto, na prática ambas volatilidades raramente coincidem.

5.1 Volatilidade Implícita e Realizada

Consideramos uma ação que não paga dividendo S_t e um título zero cupom Λ_t , e um intervalo de tempo $[0, T]$, onde $\Lambda_T = 1$. O preço descontado da ação é dado por $\tilde{S}_t = S_t/\Lambda_t$, e, em geral, para outros instrumentos de preço V_t , $\tilde{V}_t = V_t/\Lambda_t$. No caso

especial da taxa de juros não aleatória r_t , o preço do título zero cupom no tempo t pode ser expressa como $\Lambda_t = \exp\{-\int_t^T r_u dt\}$. Por simplicidade, devemos desenvolver a teoria para preços descontados.

Vamos assumir que sob a medida física real \mathbb{P} , o preço descontado da ação seja:

$$d\tilde{S}_t = \mu_t \tilde{S}_t dt + \sigma_t \tilde{S}_t dW_t \quad (5.1)$$

onde o termo do *drift* μ_t e o termo de difusão σ_t são estocásticos e variantes com o tempo, e W segue um Browniano padrão. Sob a medida neutra ao risco, (5.1) se torna $d\tilde{S}_t = \sigma_t \tilde{S}_t dW_t^*$, onde W^* é o Browniano padrão sob a medida neutra ao risco. Vamos assumir também que todas as quantidades em (5.1) são processos adaptados a uma filtração simples (\mathcal{F}_t) , a qual não é necessariamente gerada por S_t ou W_t . No entanto, assumimos que W_t é um processo de Wiener. Também supomos que S_t é um processo de Itô, que em adição ao último, requer que $|\mu_u|$ e σ_u^2 sejam integráveis (q.c.) no intervalo de $[0, T]$. Finalmente, supomos que $\sigma_t^2 > 0$ para todo t .

Considere um opção européia sobre o ativo objeto com *payoff* $f(S_T)$ no vencimento T , seu preço no modelo de (Black & Scholes 1973) no tempo t pode ser escrito como $C(S_t, -\log(\Lambda_t), \sigma^2(T-t))$, onde

$$C(S, R, \Xi) = \exp(-R) \mathbb{E} f(S \exp(R - \Xi/2 + \sqrt{\Xi} Z)). \quad (5.2)$$

A maturidade T e o formato do *payoff* f são dados pelo contrato da opção, Z é a distribuição normal padrão e \mathbb{E} denota o valor esperado.

Hipótese 5.1. *Assuma que a função f satisfaz: (i) $f : (0, +\infty) \rightarrow \mathbb{R}$, (ii) f é convexa e não é uma função afim em $(0, +\infty)$; e (iii) $\mathbb{E}|f(\exp(U))| < \infty$ para qualquer variável aleatória normal U .*

Da *Hipótese 5.1*, segue que C em (5.2) está bem definida e é diferenciável infinitas

vezes, com $C_{\Xi} > 0$ ¹, para $\Xi > 0$, $S > 0$, e R qualquer. Seja, também, $C(S, R, \infty) = \lim_{\Xi \rightarrow \infty} C(S, R, \Xi)$.

Suponhamos que o preço de mercado atual desta opção é dado por V_t , e suponha também que $\tilde{V} = V/\Lambda$ é um processo de Itô (\mathcal{F}_t) -adaptado. Devemos destacar que o Browniano que dirige V não é necessariamente o mesmo que dirige S , entretanto se espera que os dois estejam conectados de uma forma geral. A hipótese de um processo de Itô é natural a S , assim como para V , desde que seus valores descontados sejam martingais sob a medida neutra ao risco, e desde que a medida original P e a medida neutra ao risco sejam mutuamente absolutamente contínuas sob hipóteses de não arbitragem (ver (Delbaen & Schachermayer 1995a), (Delbaen & Schachermayer 1995b) e (Delbaen & Schachermayer 1998) para argumentos de não arbitragem).

Das condições normais de não arbitragem temos,

$$\forall t \in [0, T), \quad C(S, -\log(\Lambda_t), \infty) > V_t > C(S, -\log(\Lambda_t), 0) \quad \text{q.c.} \quad (5.3)$$

Então, segue que a definição a seguir, como em (Zhang 2009), está bem definida.

Definição 5.1. *Sob a Hipótese 5.1, a volatilidade implícita acumulada (VIA) no tempo t é definida como a única solução Ξ_t de*

$$V_t = C(S_t, -\log(\Lambda_t), \Xi_t). \quad (5.4)$$

A volatilidade realizada pode estar tanto na sua forma instantânea (σ_t^2) ou na sua forma de variação quadrática integrada ($\langle \log \tilde{S}, \log \tilde{S} \rangle_t = \int_0^t \sigma_u^2 du$).

¹ C_{Ξ} é definido como $\partial C / \partial \Xi$.

Se dividirmos ambos os lados da Equação (5.4) por Λ_t segue,

$$\begin{aligned}\tilde{V}_t &= \frac{1}{\Lambda_t} C(S_t, -\log(\Lambda_t), \Xi_t) \\ &= C(\tilde{S}_t, -\log(\Lambda_t), \Xi_t)\end{aligned}\tag{5.5}$$

Na Equação (5.5), utilizando o título zero cupom como numerário reduzimos as dimensões do modelo do preço C da opção, onde C se torna função apenas do preço futuro \tilde{S}_t e da VIA Ξ_t . Em particular, Ξ_t é também um processo de Itô pelo Teorema da Função Implícita, em vista de nossas hipóteses sobre f (e conseqüentemente C) e porque \tilde{S}_t e \tilde{V}_t são processos de Itô.

5.2 Modelo de Zero Fator

A noção de volatilidade implícita em escala cumulativa, de t até T , segue o conceito descrito em (Zhang 2009), contrasta com uma parte da literatura, que considera a volatilidade implícita instantânea. O autor argumenta que a volatilidade implícita instantânea quase nunca existe, e se existir, essa deve ser idêntica a volatilidade instantânea histórica. Esta idéia é descrita pelo teorema:

Teorema 5.1. (Zhang2009). *Assuma a Hipótese 5.1. Suponha \tilde{S}_t e \tilde{V}_t sejam processos de Itô, com \tilde{S}_t satisfazendo (5.1). Seja a volatilidade implícita acumulada Ξ dada por (5.4). Assuma que existe (pelo menos) uma medida P^* , equivalente a P , sob as quais \tilde{S}_t e \tilde{V}_t são martingais (em particular, não existe arbitragem). Se Ξ é absolutamente contínua como uma função de T no intervalo (t_1, t_2) , $\xi_t = -\frac{d}{dt}\Xi_t$, então para esse*

mesmo intervalo, sob ambas medidas,

$$\sigma_t^2 = \xi_t, \quad e \quad (5.6)$$

$$d\tilde{V}_t = \Delta_t d\tilde{S}_t, \quad \text{onde} \quad (5.7)$$

$$\Delta_t = C_{\tilde{S}}(\tilde{S}_t, 0, \Xi_t). \quad (5.8)$$

Neste modelo, volatilidade implícita leva a um delta hedge exato. Entretanto, a predição do *Teorema 5.1* raramente ocorre na realidade, em outras palavras, “volatilidade instantânea implícita” ($-\frac{d}{dt}\Xi_t$) tipicamente não existirá. Intuitivamente, o motivo é: se Ξ_t é absolutamente contínua no intervalo de $(0, T)$, então sujeito a condições de regularidade, a Equação (5.6) leva a $\int_0^T \sigma_t^2 dt = -\Xi_T + \Xi_0$. Se $\Xi_T = 0$, implica que para $\int_0^T \sigma_t^2 dt$ será conhecida e igual a Ξ_0 no tempo zero. Esta condição não é usual.

O modelo descrito acima é chamado de *modelo zero fator*, uma vez que não existe impacto de algum fator estrutural na volatilidade quando está se hedgeando o derivativo.

Vamos mostrar, então, uma conexão entre volatilidade realizada e implícita. Seja $\Xi^- < \Xi^+$, e considere o conjunto

$$\Xi^- \leq \int_0^T \sigma_t^2 dt \leq \Xi^+. \quad (5.9)$$

Temos o seguinte resultado:

Teorema 5.2. *Assuma a Hipótese 5.1. Suponha \tilde{S} e \tilde{V} são processos de Itô, com \tilde{S} satisfazendo (5.1). Seja a volatilidade implícita acumulada Ξ seja dada por (5.4). Assuma que exista (pelo menos) uma medida P^* , equivalente a P , sob a qual \tilde{S} e \tilde{V} são martingais (em particular, não existe arbitragem). Finalmente, assumo que a Equação*

(5.9) vale com probabilidade um. Então o conjunto

$$\forall t \in [0, T] : \Xi^- - \int_0^t \sigma_u^2 du \leq \Xi_t \leq \Xi^+ - \int_0^t \sigma_u^2 du \quad (5.10)$$

também vale com probabilidade um.

Demonstração. Seja $\epsilon > 0$. Suponha $\Xi_t + \int_0^t \sigma_u^2 du$ atinja $\Xi^+ + \epsilon$ no tempo $\tau \in (0, T)$. Neste tempo τ , vende-se uma unidade de V_τ , e inicia-se uma estratégia de negociação no tempo τ com valor inicial (descontado) $C(\tilde{S}_\tau, 0, \Xi^+ - \int_0^\tau \sigma_u^2 du)$ e a taxa de hedge (delta) da ação de $C_{\tilde{S}}(\tilde{S}_t, 0, \Xi^+ - \int_0^t \sigma_u^2 du)$ no tempo t em $[\tau, T]$. Como em (Mykland 2000), esta estratégia produz um *payoff* no tempo T igual a $C(\tilde{S}_T, 0, \Xi^+ - \int_0^T \sigma_u^2 du) \leq V_T$. Então, a estratégia produz um lucro positivo no tempo τ , e não se pode perder dinheiro até o tempo T . Isto prova o limite superior. A prova para o limite inferior é análogo. \square

5.2.1 Identidades de Não Arbitragem

Como vimos acima, fica claro que esperamos que a volatilidade implícita seja não diferenciável com respeito a t e possua, então, variação quadrática não nula. Utilizando o *Lema de Itô* em (5.5), temos

$$\begin{aligned} d\tilde{V}_t &= C_{\tilde{S}} d\tilde{S}_t + C_{\Xi} d\Xi_t \\ &\quad + \frac{1}{2} C_{\tilde{S}\tilde{S}} d\langle \tilde{S}, \tilde{S} \rangle_t \\ &\quad + \frac{1}{2} C_{\Xi\Xi} d\langle \Xi, \Xi \rangle_t + C_{\tilde{S}\Xi} d\langle \tilde{S}, \Xi \rangle_t \end{aligned} \quad (5.11)$$

Agora vamos definir o seguinte processo

$$\begin{aligned} d\Xi_t^{DR} &= -\frac{1}{C_{\Xi}} \left\{ \frac{1}{2} C_{\tilde{S}\tilde{S}} d\langle \tilde{S}, \tilde{S} \rangle_t + \frac{1}{2} C_{\Xi\Xi} d\langle \Xi, \Xi \rangle_t + C_{\tilde{S}\Xi} d\langle \tilde{S}, \Xi \rangle_t \right\} \\ &= -\sigma_t^2 dt - \frac{1}{C_{\Xi}} \left\{ \frac{1}{2} C_{\Xi\Xi} d\langle \Xi, \Xi \rangle_t + C_{\tilde{S}\Xi} d\langle \tilde{S}, \Xi \rangle_t \right\} \end{aligned} \quad (5.12)$$

desde que $C_{\Xi} = C_{\tilde{S}}\tilde{S}^2/2$. Seja os termos restantes Ξ_t^{MG2} de Ξ_t dado por

$$d\Xi_t = d\Xi_t^{DR} + d\Xi_t^{MG}. \quad (5.13)$$

Então a Equação (5.11) se torna

$$d\tilde{V}_t = C_{\tilde{S}}d\tilde{S}_t + C_{\Xi}d\Xi_t^{MG}, \quad (5.14)$$

e, em particular, Ξ^{MG} é um martingal sob qualquer medida neutra ao risco equivalente. Ξ^{MG} pode ser visto como o valor descontado de Ξ . Para sumarizar o discorrido acima temos:

Teorema 5.3. *Assuma a Hipótese 5.1. Suponha que \tilde{S} e \tilde{V} são processos de Itô, com \tilde{S} satisfazendo (5.1). Seja a volatilidade implícita acumulada Ξ dada pela Equação (5.4). Assuma que exista (pelo menos) uma medida P^* , equivalente a P , sob a qual \tilde{S} e \tilde{V} são martingais. Então, sob P e qualquer medida equivalente, as Equações (5.12) e (5.14) se mantêm.*

5.3 Hedgeando a Volatilidade Implícita para Modelo de Um Fator

Vimos na seção anterior que se a volatilidade implícita é absolutamente contínua, então se produz o hedge exato. Em um caso mais geral, suponha que

$$d\Xi_t = \rho_t d\tilde{S}_t + dZ_t, \quad (5.15)$$

onde $\rho_t = d\langle \Xi, \tilde{S} \rangle_t / d\langle \tilde{S}, \tilde{S} \rangle_t$, para $t \in (0, T)$. A Equação (5.15), pode ser vista como a regressão de Ξ em \tilde{S} , com ρ sendo o coeficiente da regressão e Z sendo o resíduo.

²MG faz referência à martingal e DR à *drift*.

Neste caso, Z pode ou não ser absolutamente contínuo. O hedge mínimo martingal (ver (Föllmer & Sondermann 1985), (Föllmer, Schweizer & Bonn 1990), (Schweizer 1990) e (Schweizer 1991)) fornece a evolução do preço do derivativo da seguinte forma:

$$d\tilde{V}_t = \Delta_t d\tilde{S}_t + C_\Xi dZ_t^{MG}, \text{ onde} \quad (5.16)$$

$$\Delta_t = C_{\tilde{S}} + \rho_t C_\Xi \quad (5.17)$$

onde $dZ_t^{MG} = d\Xi_t^{MG} - \rho_t d\tilde{S}_t$. Então, a Equação (5.17) produz uma correção no delta hedge “implícito”, melhorando a operação baseada apenas na volatilidade implícita. Podemos pensar nisso como uma minimização da exposição ao risco de “Gamma” e ao “Vega” (ver (Hull 2009)).

Da Equação (5.15) temos,

$$d\langle \Xi, \Xi \rangle_t = \rho_t^2 d\langle \tilde{S}, \tilde{S} \rangle_t + 2\rho_t d\langle \tilde{S}, Z \rangle_t + d\langle Z, Z \rangle_t \quad (5.18)$$

$$d\langle \tilde{S}, \Xi \rangle_t = \rho_t d\langle \tilde{S}, \tilde{S} \rangle_t + d\langle \tilde{S}, Z \rangle_t. \quad (5.19)$$

Assim, substituindo as Equações (5.18) e (5.19) em (5.11) e lembrando que $d\langle \tilde{S}, Z \rangle_t = 0$, temos

$$\begin{aligned} d\tilde{V}_t &= C_{\tilde{S}} d\tilde{S}_t + \rho_t C_\Xi d\tilde{S}_t + C_\Xi dZ_t \\ &+ \frac{1}{2} C_{\tilde{S}\tilde{S}} \tilde{S}^2 \left(1 + \rho_t^2 \frac{C_{\Xi\Xi}}{C_{\tilde{S}\tilde{S}}} \right) \\ &+ \frac{1}{2} C_{\Xi\Xi} d\langle Z, Z \rangle_t + \frac{1}{2} C_{\tilde{S}\tilde{S}} \tilde{S}^2 \left(2\rho_t \frac{C_{\tilde{S}\Xi}}{C_{\tilde{S}\tilde{S}}} \right). \end{aligned} \quad (5.20)$$

Analogamente à Equação (5.12), desde que $C_\Xi = C_{\tilde{S}\tilde{S}} \tilde{S}^2 / 2$, definimos,

$$d\Xi_t^{DR} = -\sigma_t^2 \left(1 + \rho_t^2 \frac{C_{\Xi\Xi}}{C_{\tilde{S}\tilde{S}}} + 2\rho_t \frac{C_{\tilde{S}\Xi}}{C_{\tilde{S}\tilde{S}}} \right) - \frac{1}{2} \frac{C_{\Xi\Xi}}{C_\Xi} d\langle Z, Z \rangle_t. \quad (5.21)$$

Dessa forma, utilizando a Equação (5.13), obtemos a Equação (5.16).

O coeficiente ρ_t é uma forma de *leverage effect* da volatilidade implícita, e é empiricamente esperado que tenha sinal negativo. Então, se C_{Ξ} é positivo para *payoffs* convexos, obtemos um delta hedge na Equação (5.17) menor que o delta implícito de Black-Scholes. Em outras palavras, ao utilizarmos apenas o delta de Black-Scholes, tenderíamos a realizar um superhedge. Um resultado similar é encontrado em (Renault & Touzi 1996).

É importante salientar que o hedge na Equação (5.17) pode ser obtido utilizando diretamente os preços de mercado das opções V_t , uma vez que $C_{\tilde{S}} + \rho_t C_{\Xi} = d\langle \tilde{V}, \tilde{S} \rangle_t / d\langle \tilde{S}, \tilde{S} \rangle_t$. A vantagem de se utilizar a volatilidade implícita Ξ é que, em geral, a volatilidade se move na mesma ordem pelos *strikes*, conservando assim as mesmas propriedades estatísticas, enquanto os preços das opções fora e dentro dinheiro se comportam de forma bastante distintas.

Ao realizarmos um hedge de acordo com a Equação (5.16), teremos que nos preocupar com o tamanho do resíduo $C_{\Xi} Z_t^{MG}$, que advém da existência de outros fatores além de ρ_t . No entanto, se consideramos que o modelo possui apenas um fator, obtemos um hedge perfeito.

Corolário 5.1. *Assuma as condições do Teorema 5.3. Suponha também que a volatilidade implícita acumulada Ξ satisfaça a Equação (5.15) de forma que*

$$dZ_t = -\xi_t. \quad (5.22)$$

Então

$$\xi_t = \sigma_t^2 + \rho_t^2 \sigma_t^2 \frac{C_{\Xi\Xi}}{C_{\tilde{S}\tilde{S}}} + 2\rho_t \sigma_t^2 \frac{C_{\tilde{S}\Xi}}{C_{\tilde{S}\tilde{S}}}, \quad e \quad (5.23)$$

$$d\tilde{V}_t = \Delta_t d\tilde{S}_t, \quad (5.24)$$

onde Δ_t é dado pela Equação (5.17), sob a medida física \mathbb{P} e qualquer medida equivalente.

Este resultado segue diretamente de (5.16), (5.17) e (5.21) desde que $dZ^{MG} \equiv 0$.

5.4 Estimação do Delta no Modelo de um Fator

No trabalho (Mykland & Zhang 2008) os autores propõem o seguinte método para estimar o parâmetro ρ_t do modelo de um fator:

$$\hat{\rho}_t = \frac{\langle \widehat{S}, \widehat{\Xi} \rangle_t}{\langle \widehat{S}, \widehat{S} \rangle_t} = \frac{\sum_{t-h < t_i < t_{i+1} \leq t} (\Delta \log(S_{t_i})) (\Delta \log(\Xi_{t_i}))}{\sum_{t-h < t_i < t_{i+1} \leq t} (\Delta \log(S_{t_i}))^2} \quad (5.25)$$

onde $[0, T]$ é uma partição não aleatória com $0 \leq t_1 \leq t_2 \leq \dots \leq t_n = T$. A estimação do erro de $\langle \widehat{S}, \widehat{\Xi} \rangle_t$ pode ser decomposta em duas partes, que possuem ordem de $O_p(\sqrt{h})$ e $O_p(\sqrt{\frac{\Delta t^{(n)}}{h}})$, respectivamente, onde

$$\overline{\Delta t}^{(n)} = \frac{1}{n} \sum_{i=1}^n \Delta t_i^{(n)} = \frac{T}{n}. \quad (5.26)$$

pela hipótese alocação regular. Pela expansão estocástica de Taylor, a estimação do erro de ρ pode ser expressada como

$$\begin{aligned} \hat{\rho}_t - \rho_t &= \frac{1}{\langle S, S \rangle_t} [\langle \widehat{S}, \widehat{\Xi} \rangle_t - \langle S, \Xi \rangle_t] \\ &\quad - \frac{\rho_t}{\langle S, S \rangle_t} [\langle \widehat{S}, S \rangle_t - \langle S, S \rangle_t] + o_p \left(\sqrt{h} + \sqrt{\frac{\Delta t^{(n)}}{h}} \right) \end{aligned} \quad (5.27)$$

As propriedades assintóticas da estimação do erro associado à $\langle \widehat{S}, \widehat{\Xi} \rangle_t$ e $\hat{\rho}_t$ estão em (Mykland & Zhang 2008).

Como podemos notar, no método acima existe um custo de oportunidade associ-

ado ao espaçamento h , como explicitado em (5.27), o que sugere novamente que usar todas observações deveria nos levar a uma estimação de ρ_t com erro maior do que estimar os retornos de forma esparsa.

Nosso objetivo final é estimar o delta relativo para o modelo de um fator $\Delta_t/C_{\widehat{S}}$, onde o Δ_t é dado pela Equação (5.17), para quantificarmos qual a porcentagem do delta de Black-Scholes deveríamos estar usando para realizar o delta hedge correto. No entanto, primeiramente, vamos nos concentrar em estimar a covariância $\langle \widehat{S}, \widehat{\Xi} \rangle_t$, pois o delta e o vega de Black-Scholes é amplamente conhecido e o denominador é a VRDE que estimamos no capítulo anterior. Para estimarmos a covariância $\langle \widehat{S}, \widehat{\Xi} \rangle_t$ escolhi duas opções sobre PETR4 que possuíam o *strike* mais próximo do preço da ação naqueles dias. O período se refere a 5 dias de negociação do dia 07/02/2011 a 11/02/2011. As opções são a PETRB26 e a PETRB28, com preços de exercício R\$ 25,71 e R\$ 27,71 respectivamente. A Figura 5.1 mostra como se comportou o preço da PETR4 durante o período analisado. Note que o preço nunca ficou abaixo de R\$ 26,00 e nunca acima de R\$ 28,00.

Ambas opções tinham vencimento no dia 21/02/2011. A taxa de juros considerada foi meta da taxa básica Selic que naquele momento tinha seu valor anualizado em 11,25%. Para calcular as volatilidades implícitas para cada negócio, utilizamos a função `blsimpl`, do Matlab. Nas tabelas que seguem mostramos a estatística descritiva dos retornos da volatilidade implícita ($\Delta \Xi_t$):

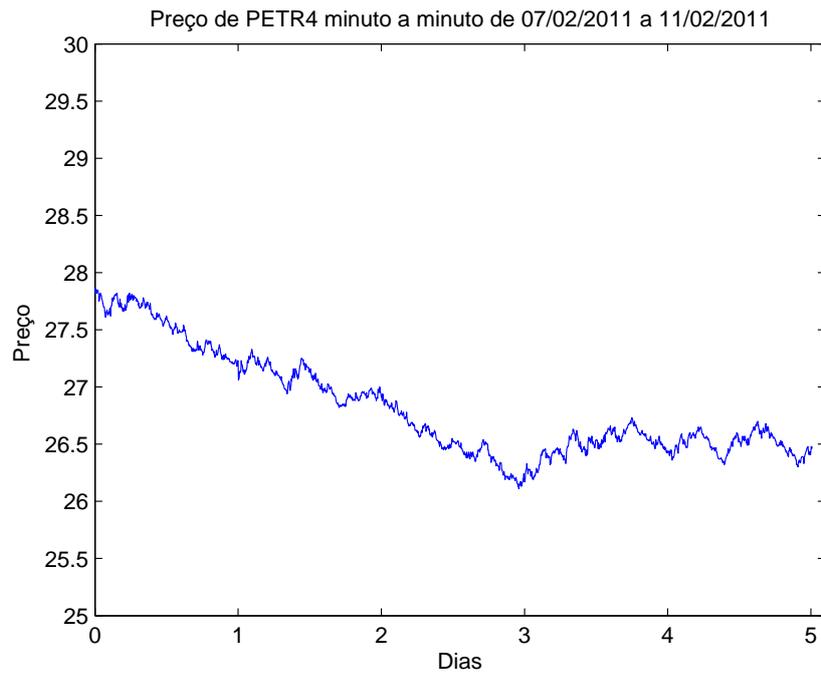


Figura 5.1: Preço de PETR4, minuto a minuto de 07/02/2011 a 11/02/2011

Estatística Descritiva PETRB26	07/02	08/02	09/02	10/02	11/02
Número de transações	6.211	6.927	12.200	12.232	11.740
Tempo médio das transações(seg)	4,06	3,64	2,07	2,06	2,15
Maior log-retorno das volatilidades(%)	54,76	23,59	12,88	13,58	14,77
Menor log-retorno das volatilidades(%)	-67,04	-23,13	-11,60	-12,18	-15,69

Tabela 5.1: Estatística descritiva de PETRB26 de 07/02/2011 a 11/02/2011.

Estatística Descritiva PETRB28	07/02	08/02	09/02	10/02	11/02
Número de transações	13.790	9.039	6.965	3.771	2.123
Tempo médio entre as transações(seg)	1,83	2,79	3,62	6,68	11,87
Maior log-retorno das volatilidades(%)	5,89	5,06	5,81	6,65	8,02
Menor log-retorno das volatilidades(%)	-6,53	-6,65	-6,38	-6,16	-8,67

Tabela 5.2: Estatística descritiva de PETRB28 de 07/02/2011 a 11/02/2011.

Analogamente ao capítulo anterior, vamos comparar o estimador em duas escalas, que denominaremos de covariância realizada em duas escalas (**CRDE**) e dado por

$$\langle \widehat{S}, \widehat{\Xi} \rangle_T^{aju} = \left(1 - \frac{\bar{n}}{n}\right) \langle \widehat{S}, \widehat{\Xi} \rangle_T, \text{ onde} \quad (5.28)$$

$$\langle \widehat{S}, \widehat{\Xi} \rangle_T = [X, Y]_T^{\text{média}} - \frac{\bar{n}}{n} [X, Y]_T^{\text{tudo}}, \text{ onde} \quad (5.29)$$

$$X_t = \log(S_t) + \epsilon_t$$

$$Y_t = \log(\Xi_t) + \delta_t \quad (5.30)$$

com estimador de retornos esparsos, que denominaremos covariância realizada esparsamente (**CRE**), dado por

$$[X, Y]^{esparso} = [X, Y]^{\mathcal{H}} = \sum_{t_i, t_{i,+} \in \mathcal{H}} (X_{t_{i,+}} - X_{t_i})(Y_{t_{i,+}} - Y_{t_i}), \quad (5.31)$$

onde \mathcal{H} é descrito na *Definição 3.2*. Os erros ϵ_t e δ_t são devido a microestrutura nos processos observados X_t e Y_t .

Nos gráficos que seguem estimamos a CRDE e a CRE utilizando janelas dos retornos de 1 a 60 minutos, com um passo de 30 segundos, para cada dia. Com estes gráficos podemos entender como os estimadores se comportam à medida que os vamos estimando mais esparsamente.

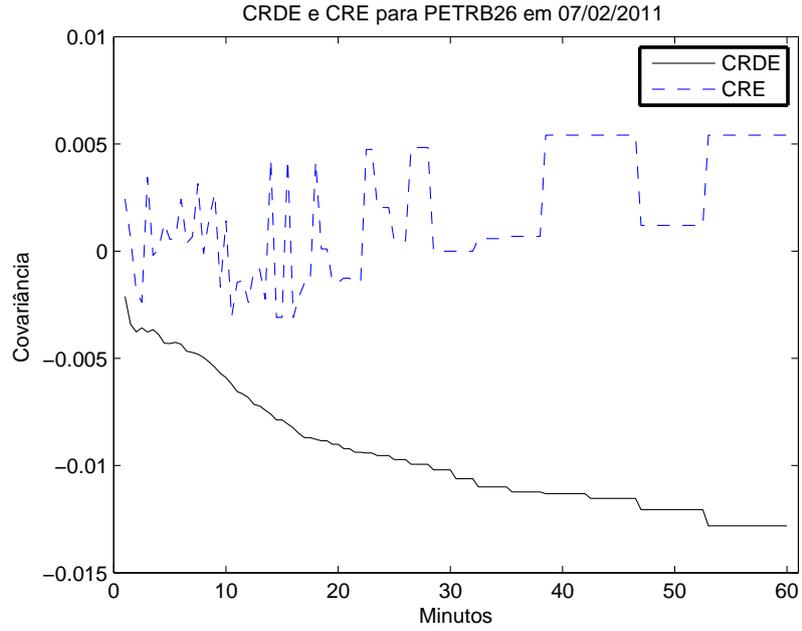


Figura 5.2: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB26 em 07/02/2011.

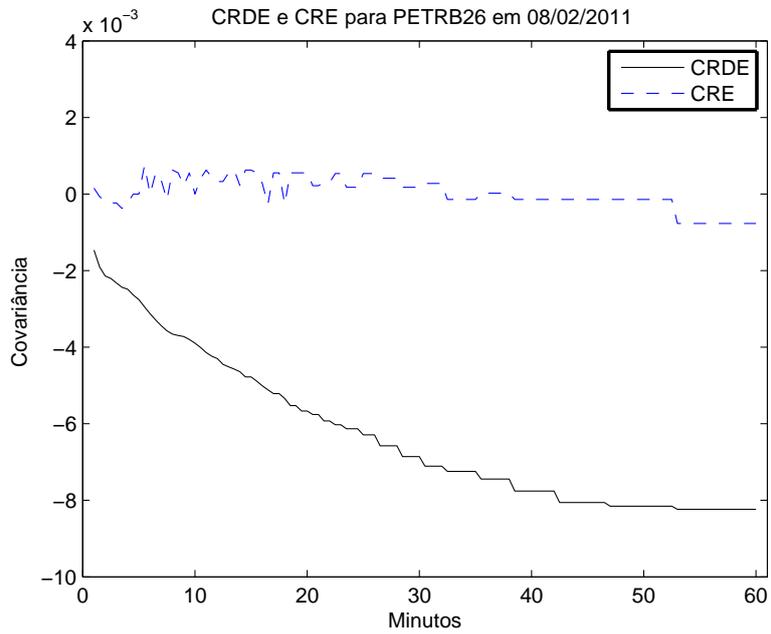


Figura 5.3: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB26 em 08/02/2011.

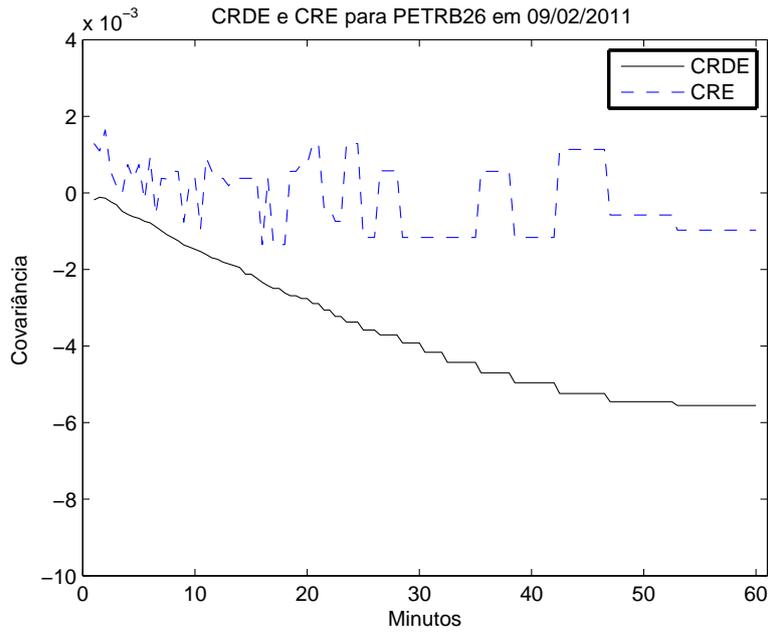


Figura 5.4: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB26 em 09/02/2011.

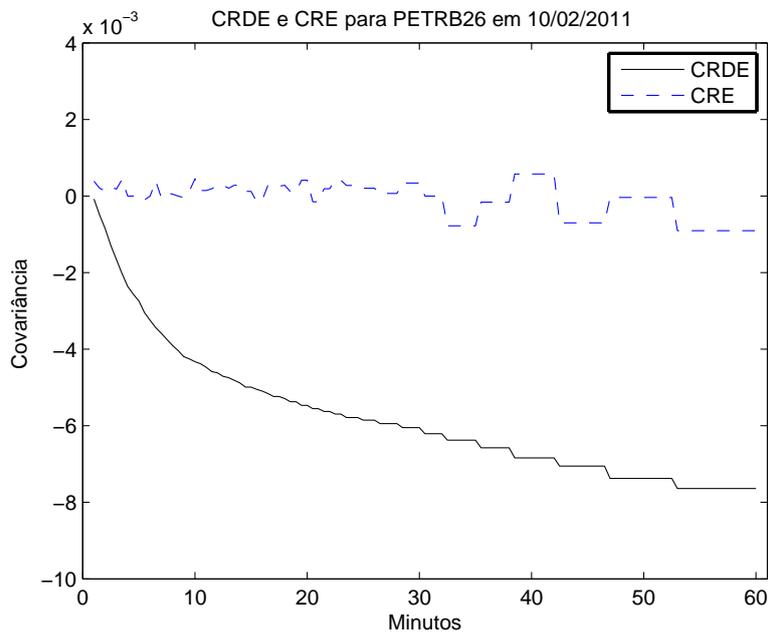


Figura 5.5: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB26 em 10/02/2011.

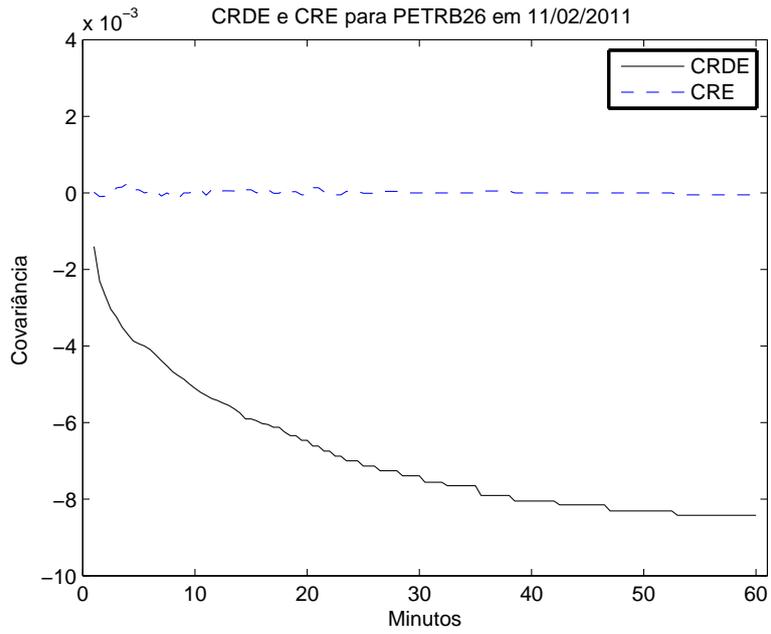


Figura 5.6: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB26 em 11/02/2011.

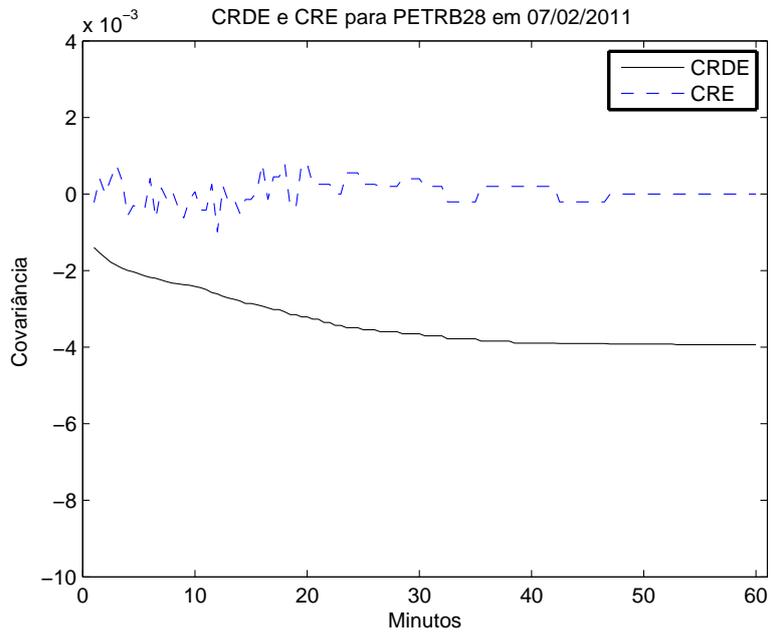


Figura 5.7: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PETRB28 em 07/02/2011.

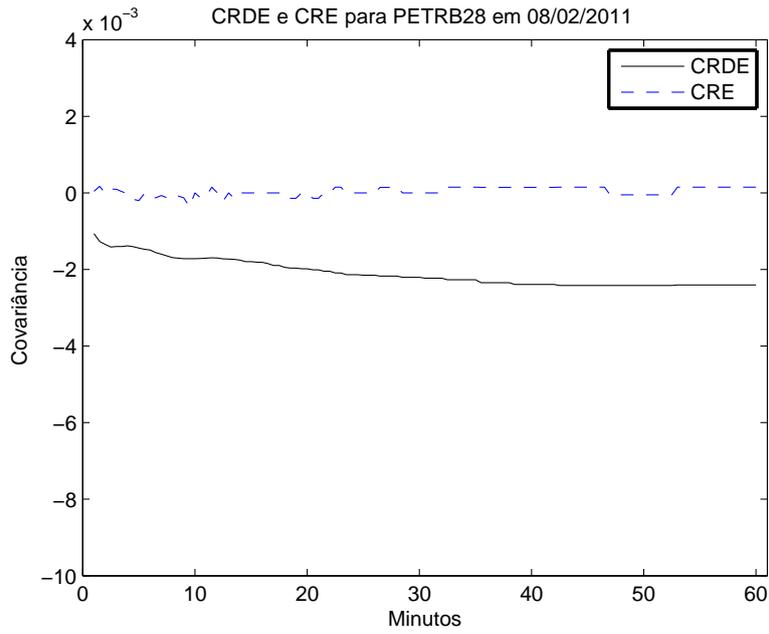


Figura 5.8: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PE-TRB28 em 08/02/2011.

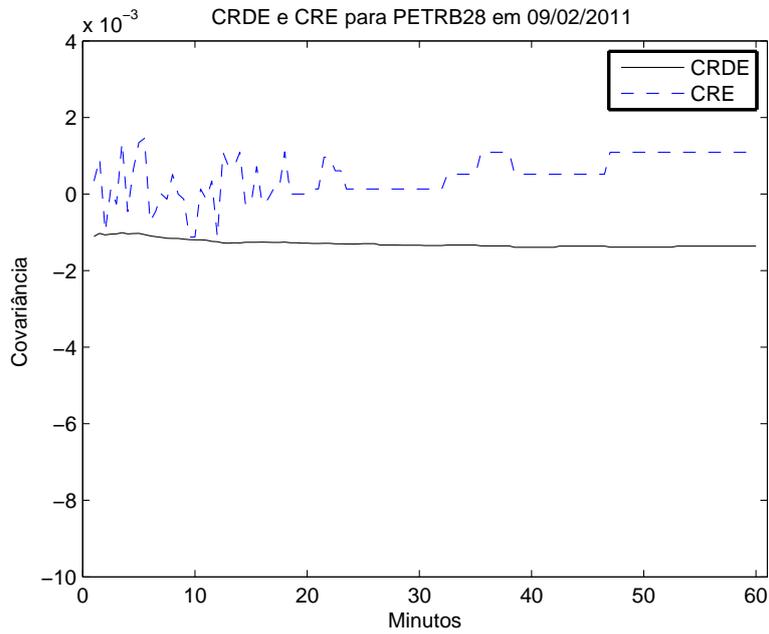


Figura 5.9: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PE-TRB28 em 09/02/2011.

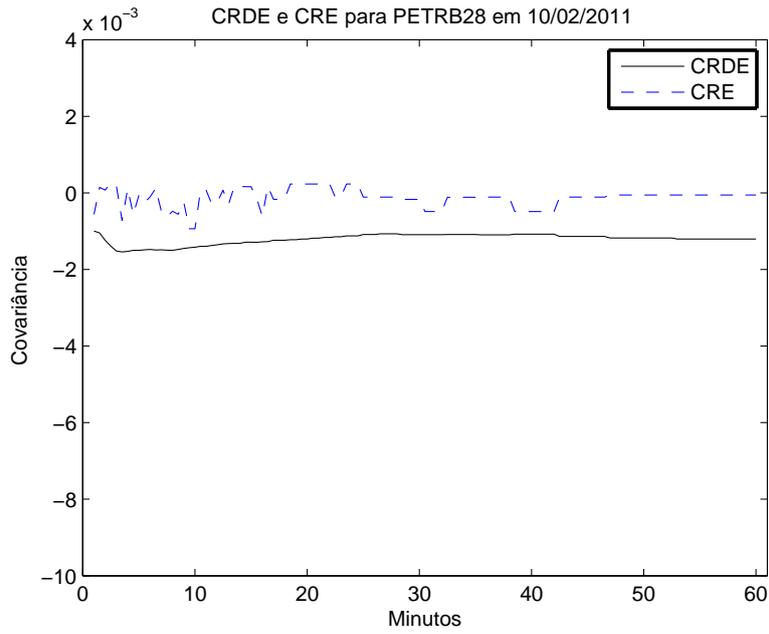


Figura 5.10: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PE-TRB28 em 10/02/2011.

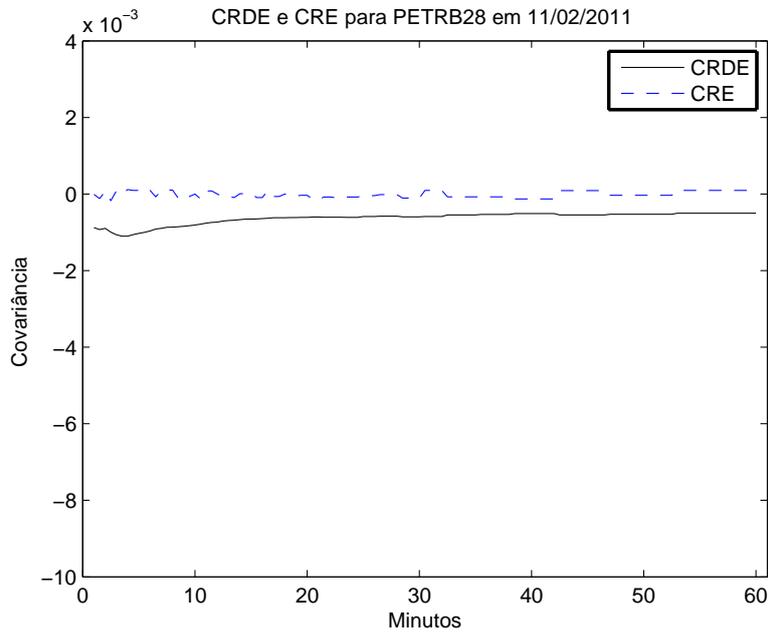


Figura 5.11: Estimação da CRDE (Eq. (5.28)-(5.30)) e CRE (Eq. (5.31)) para PE-TRB28 em 11/02/2011.

Analisando as figuras de comparação da CRDE com a CRE, podemos ressaltar vários pontos interessantes. Primeiramente, quanto à CRE, notamos que o estimador oscila entre uma covariância $\langle \widehat{S}, \widehat{\Xi} \rangle$ negativa e positiva com o *range* de estimação até 20 minutos, mas ainda muito próxima de zero, o que contraria os resultados empíricos, que mostram que o $\hat{\rho}$ e consequentemente a covariância $\langle \widehat{S}, \widehat{\Xi} \rangle$ deveriam ter um sinal negativo bem definido para o mesmo conjunto de dados, independente da janela de tempo utilizada. A partir dos 20 minutos, a CRE é praticamente nula, com exceção para as Figuras 5.2 e 5.4.

Por outro lado, a CRDE é negativa independente da janela utilizada para calculá-la, captando o efeito que foi mostrado empiricamente em (Mykland & Zhang 2008). No entanto, parece existir um viés positivo com o aumento da janela de estimação, especialmente para PETRB26, apesar de se ter uma curva bem mais suave que a CRE. Uma grande diferença que notamos entre as duas opções, nas Tabelas 5.1 e 5.2 acima, é que a magnitude dos retornos da PETRB26 são bem maiores que a da PETRB28. Um outro fator interessante é a ordem dos retornos das volatilidades implícitas que é muito maior que a dos retornos da volatilidade realizada, atingindo níveis em alguns casos, mais de 100 vezes maior. Mesmo na presença do viés a covariância $\langle \widehat{S}, \widehat{\Xi} \rangle$ tende a se estabilizar a medida que a janela aumenta.

Partimos então para o cálculo do fator $\hat{\rho}$ utilizando a CRDE para o numerador e a VRDE para o denominador. Observando os gráficos da estimação da CRDE, utilizamos a janela de 30 minutos, pois este é um ponto onde a maioria dos gráficos de CRDE começam a ter um decaimento mais suave e ainda temos um número significativo de subgrids. Note que dado o gráfico da CRDE, caso utilizássemos outras janelas diferentes das 30 minutos, estaríamos apenas causando uma pequena translação no gráfico de delta relativo. Para VRDE, utilizei a mesma janela. A Tabela 5.3 mostra os valores de $\hat{\rho}$ obtidos para cada dia para PETRB26 e PETRB28 calculados com a janela de 30 minutos.

PETRB26	07/02	08/02	09/02	10/02	11/02
$\hat{\rho}$	-3.86	-2.19	-1.40	-1.70	-2.55
PETRB28	07/02	08/02	09/02	10/02	11/02
$\hat{\rho}$	-1.23	-0.68	-0.51	-0.93	-1.05

Tabela 5.3: Estimação do fator ρ utilizando a CRDE e a VRDE com uma janela de 30 minutos.

Em posse do $\hat{\rho}$ para cada dia e para cada opção calculamos o comportamento delta relativo ($\Delta_t/C_{\hat{\rho}}$) para cada dia de negociação. Para calcular o delta e o vega de Black-Scholes utilizamos as funções `blsdelta` e `blsvega` do Matlab, respectivamente, com os mesmos parâmetros utilizados nos cálculos das volatilidades implícitas. Mostramos os resultados nos gráficos a seguir.

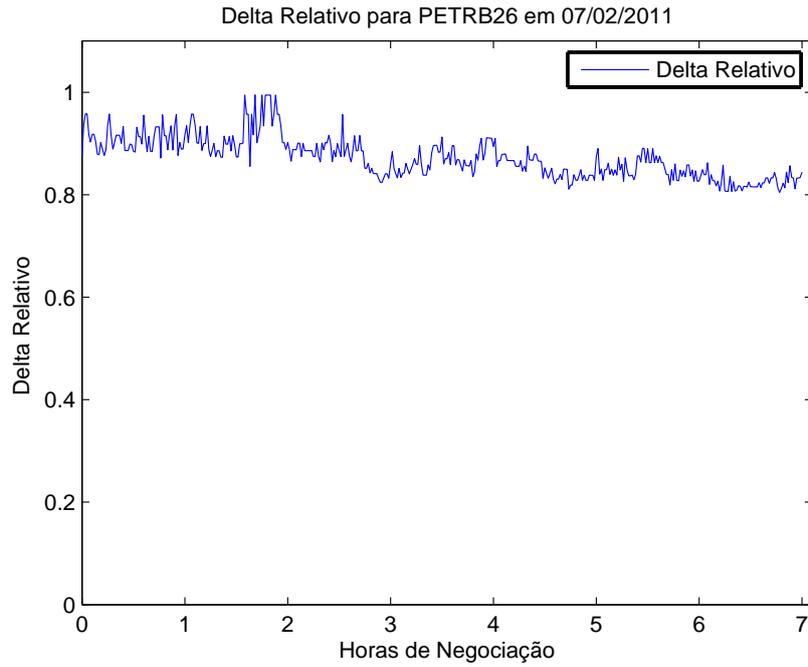


Figura 5.12: Estimação do delta relativo para PETRB26 em 07/02/2011.

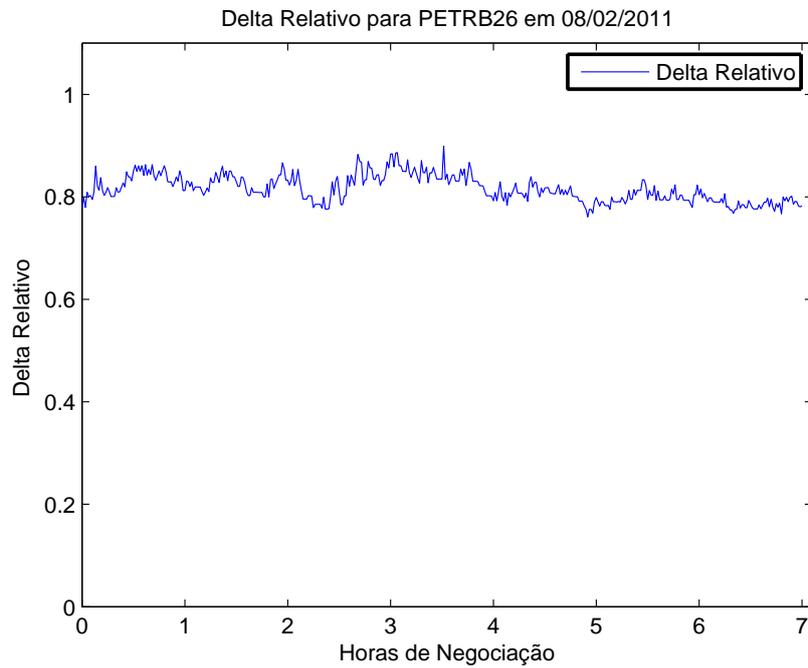


Figura 5.13: Estimação do delta relativo para PETRB26 em 08/02/2011.

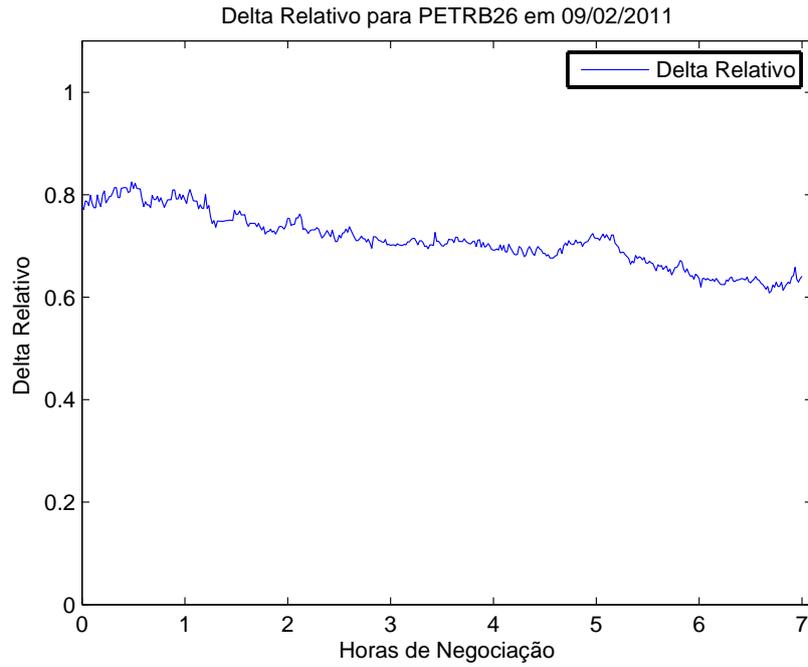


Figura 5.14: Estimação do delta relativo para PETRB26 em 09/02/2011.

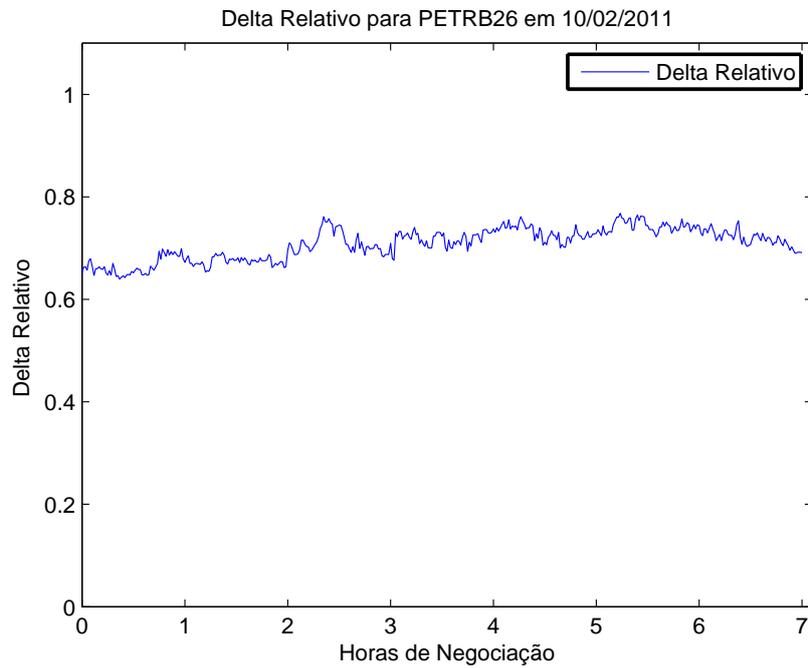


Figura 5.15: Estimação do delta relativo para PETRB26 em 10/02/2011.

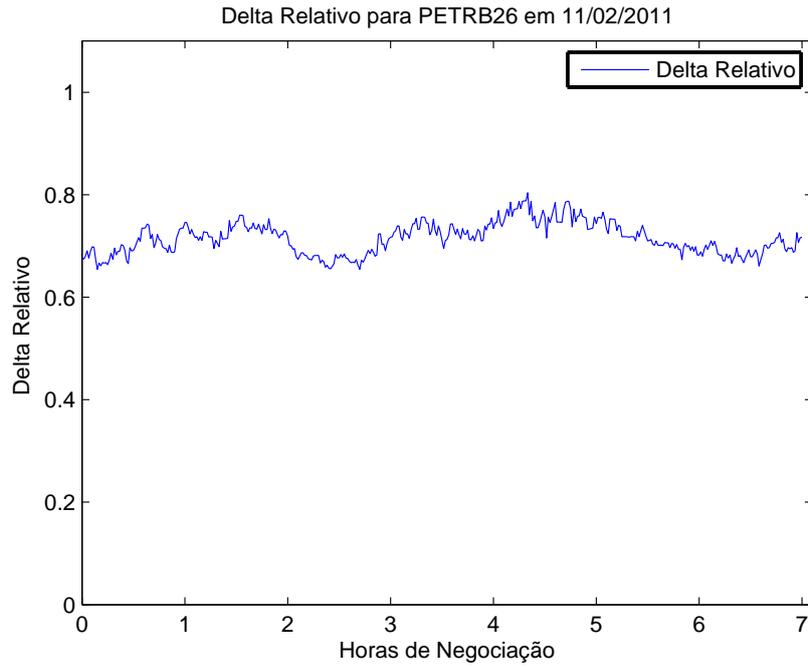


Figura 5.16: Estimação do delta relativo para PETRB26 em 11/02/2011.

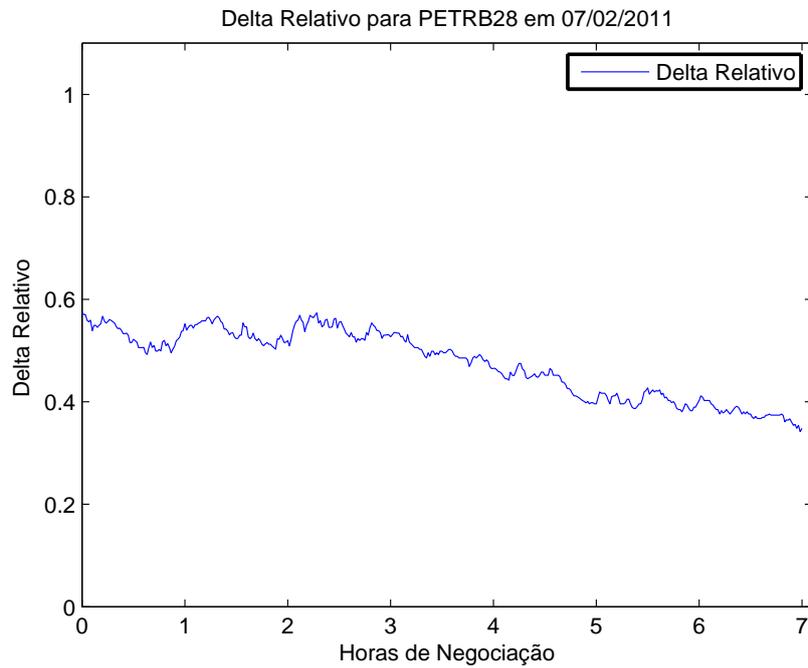


Figura 5.17: Estimação do delta relativo para PETRB28 em 07/02/2011.

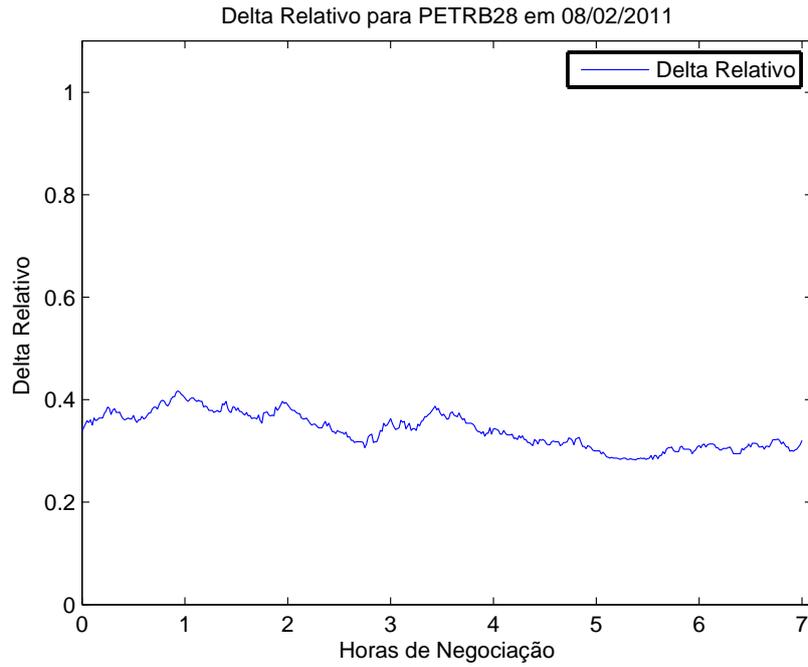


Figura 5.18: Estimação do delta relativo para PETRB28 em 08/02/2011.

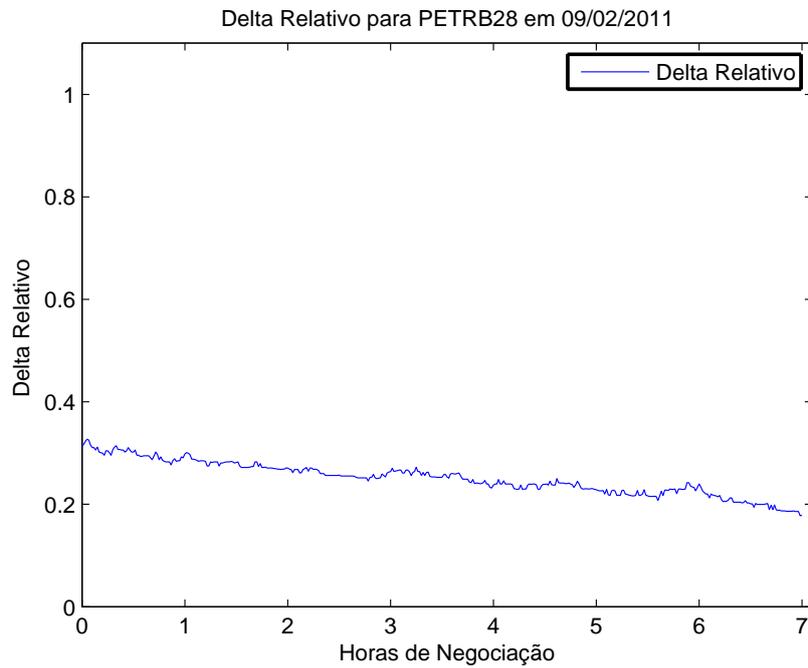


Figura 5.19: Estimação do delta relativo para PETRB28 em 09/02/2011.

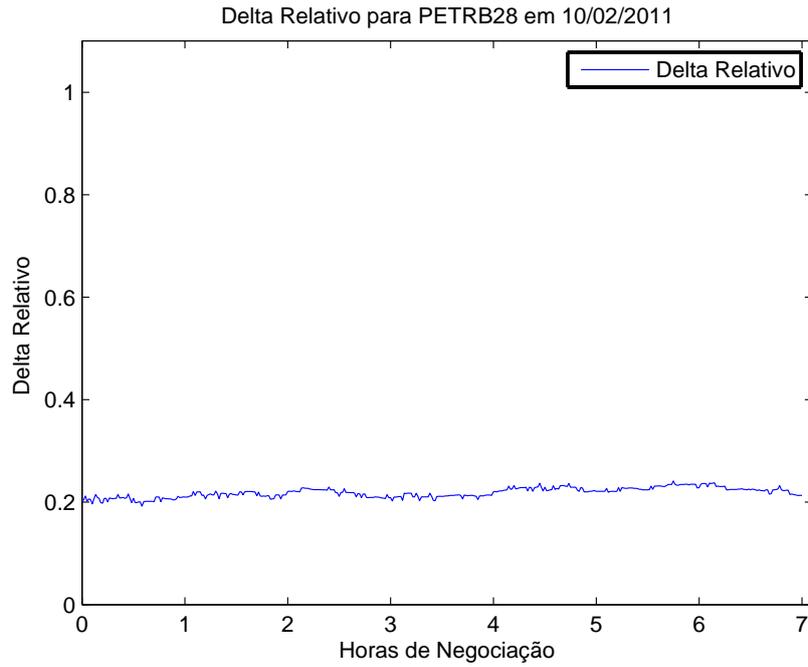


Figura 5.20: Estimação do delta relativo para PETRB28 em 10/02/2011.

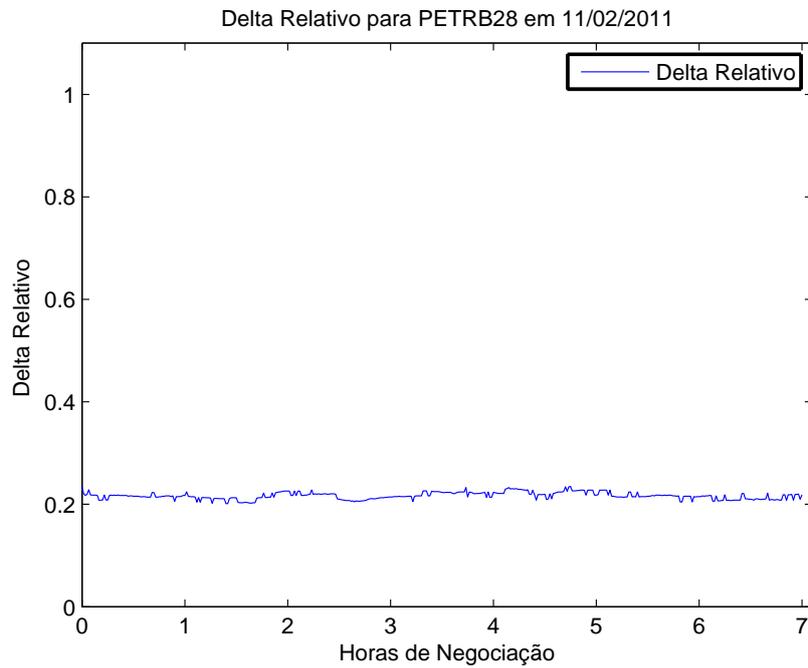


Figura 5.21: Estimação do delta relativo para PETRB28 em 11/02/2011.

Notamos que o delta relativo da PETRB28 é sempre menor que o da PETRB26, mesmo o $\hat{\rho}$ sendo sempre menor. Isto mostra como o vega é relevante na determinação do delta relativo, que é o fator que queremos saber para realizar um delta hedge mais preciso. Como esperado, o delta relativo é sempre menor que um, sugerindo que o delta de Black-Scholes realiza um superhedge. No entanto, é interessante notar que temos delta relativo próximo de um, Figura 5.12, e outros próximos de 20%, como as Figuras 5.20 e 5.21. Outro fator que pode ter influenciado o delta relativo das opções, em especial a PETRB28 (pois estava fora do dinheiro, ver gráfico 5.1), foi a proximidade do vencimento, principalmente nos dias 10 e 11/02/2011.

Conclusão

Neste trabalho estudamos modelos de volatilidade realizada em alta frequência na presença de microestruturas. No entanto, antes de abordarmos estes modelos, estudamos resultados fundamentais para teoria de finanças em alta frequência, em especial os teoremas que permitem a utilização de estimadores não-centrados.

Dos modelos com microestrutura estudados, desenvolvemos o de volatilidade realizada em duas escalas (VRDE), proposto por (Aït-Sahalia, Mykland & Zhang 2010), com uma escala mais rápida, que utiliza todas as transações e outra mais lenta que calcula a volatilidade em vários subgrids e calcula a média dessas volatilidades. Vimos que este método reduz drasticamente o viés e a volatilidade do estimador em relação à janela utilizada para se amostrar. Sendo assim, a VRDE é mais robusta que a volatilidade realizada esparsamente (VR), método mais usado para estimar volatilidade em alta frequência. Aplicamos o estimador nas quatro ações mais líquidas do Ibovespa, que são a Petrobrás PN, Vale PN, OGX ON e Itaú PN, no período de 27/09/2010 a 08/10/2010, correspondendo a dez dias de negociação, para janelas de 1 a 9 minutos, de 15 em 15 segundos, para mostrar a robustez da VRDE em relação à volatilidade realizada com amostragens esparsas.

Em (Aït-Sahalia, Mykland & Zhang 2010), os autores apresentam uma melhor performance da VRDE frente à VR, no entanto eles focam em comparar o desempenho da VRDE e da VR com dados que apresentam correlação serial entre os ruídos contra dados que não apresentam. Apesar da VRDE ter sido desenvolvido com a hipótese da

independência entre os ruídos, esse estimador performa absolutamente bem tanto na presença de correlação serial quanto na ausência. O mesmo não ocorre com a VR, que acaba tendo um desempenho significativamente pior nos dados com correlação serial em comparação ao dados com ruídos independentes.

Nossa abordagem foi testar o desempenho da VRDE frente VR em um período de alta volatilidade. Para tal, a amostra escolhida foram as duas semanas que sucederam o IPO (*initial public offering*) da Petrobras, quando a ação PETR4 sofreu bruscas mudanças de preço e em sua volatilidade como mostra a Figura (4.11). A OGXP3, por também ser uma empresa petrolífera, teve efeito semelhante à PETR4, como mostra a Figura (4.33). No entanto, a VALE5, que é uma mineradora, e a ITUB4, que é uma instituição financeira, não sofreram alterações da mesma magnitude na volatilidade realizada como mostra as Figuras (4.22) e (4.44) respectivamente.

Na Tabela 5.4, mostramos a diferença entre o valor máximo e o mínimo da VR e a VRDE para cada dia, respectivamente. Essa análise é importante, pois geralmente apenas uma janela é escolhida para estimar a volatilidade realizada, então ela retrata a máxima diferença de valores que se poderia estimar. Podemos notar que em geral, para as ações PETR4 e OGXP3, a utilização da VRDE miniza o erro de estimação significativamente, enquanto para VALE5 o desempenho da VR é relativamente satisfatório, apesar da VRDE ser melhor. Para ITUB4, apesar de não ter sofrido grandes alterações no nível de volatilidade no período, a VRDE foi sensivelmente melhor que a VR. Na Tabela 5.5 calculamos o desvio padrão dos 33 valores de VR e VRDE para cada dia. Nessa análise podemos ver que a VR performa consideravelmente pior que a VRDE para a PETR4 e OGXP3, que sofreram com o choque no período. Para a ITUB4 em alguns dias a VRDE estima sensivelmente melhor VR, como dia 04/10 e dia 08/10. Por fim, para a VALE5, a diferença entre os dois estimadores é relativamente pequena.

Ação	PETR4		VALE5		OGXP3		ITUB4	
Dia	VR	VRDE	VR	VRDE	VR	VRDE	VR	VRDE
27/09	3,72%	1,46%	5,37%	3,69%	11,47%	6,29%	3,55%	1,52%
28/09	5,55%	1,79%	3,11%	2,52%	15,27%	2,89%	9,90%	6,21%
29/09	7,50%	0,71%	4,77%	2,20%	7,95%	1,49%	8,08%	3,07%
30/09	5,61%	2,28%	2,83%	1,03%	8,54%	1,67%	2,83%	1,03%
01/10	5,53%	1,34%	3,66%	1,53%	14,21%	3,92%	3,13%	0,65%
04/10	3,86%	1,62%	3,75%	1,60%	11,49%	5,48%	9,93%	3,69%
05/10	3,82%	2,63%	2,82%	1,04%	6,63%	4,36%	5,36%	0,86%
06/10	4,41%	2,52%	3,46%	2,04%	8,81%	0,97%	3,48%	4,55%
07/10	9,18%	5,98%	3,85%	0,89%	16,39%	4,80%	8,20%	1,98%
08/10	8,72%	5,75%	4,99%	0,67%	19,84%	10,37%	3,89%	0,82%

Tabela 5.4: Diferença entre o valor máximo e mínimo calculados para a VR e VRDE para cada dia de negociação.

Ação	PETR4		VALE5		OGXP3		ITUB4	
Dia	VR	VRDE	VR	VRDE	VR	VRDE	VR	VRDE
27/09	0,86%	0,40%	1,19%	1,19%	2,48%	1,95%	0,86%	0,33%
28/09	1,21%	0,49%	0,89%	0,70%	2,77%	0,58%	2,51%	1,78%
29/09	1,38%	0,20%	0,99%	0,73%	1,79%	0,50%	1,75%	0,79%
30/09	1,12%	0,68%	0,69%	0,36%	2,00%	0,42%	0,69%	0,36%
01/10	1,22%	0,44%	0,91%	0,35%	2,51%	0,94%	0,71%	0,12%
04/10	0,96%	0,34%	0,84%	0,38%	2,08%	1,30%	3,03%	1,16%
05/10	0,85%	0,65%	0,75%	0,27%	1,54%	1,15%	1,14%	0,29%
06/10	0,96%	0,57%	0,72%	0,65%	1,66%	0,32%	0,95%	1,24%
07/10	1,76%	1,75%	0,77%	0,27%	3,66%	1,58%	1,66%	0,56%
08/10	2,18%	1,70%	1,00%	0,22%	6,29%	3,29%	1,07%	0,21%

Tabela 5.5: Desvio padrão da VR e da VRDE estimadas para cada dia de negociação.

No último capítulo utilizamos o estimador VRDE para calcular uma covariância realizada em alta frequência (CRDE) para que pudéssemos estimar o fator estrutural do modelo de um fator proposto por (Zhang 2009). É importante salientar que existe uma grande diferença entre o cálculo da VRDE e da CRDE, pois na CRDE os processos são diferentes, logo, não necessariamente, eles ocorrem no mesmo instante, o que faz que a falta de liquidez em um dos dois processos possa distorcer consideravelmente o valor estimado pela CRDE. Em alguns casos, o estimador pareceu apresentar um viés positivo com o aumento da janela de amostragem. No entanto, o método teve um desempenho melhor que o método proposto por (Mykland & Zhang 2008) e captou o fator estrutural, como mostra Tabela 5.3. Depois de estimado o fator, pudemos mostrar que o delta de Black-Scholes realiza um superhedge quando existe covariância negativa entre os retornos e a volatilidade implícita, em concordância com os resultados obtidos

por (Zhang 2009).

Concluimos que o método de volatilidade realizada e covariância realizada em duas escalas foi mais preciso nas duas situações. Ainda podemos destacar que pelo fato da VRDE utilizar todas as observações, sua estimação não fica sujeita a distorções como a de estimadores de baixa frequência, muito utilizados no mercado financeiro, que utilizam apenas os preços de máximo, mínimo, abertura e fechamento.

Outro ponto relevante em relação à VRDE é sua fácil implementação e rápido cálculo computacional. O estimador pode ser melhorado, como o proposto por (Zhang 2006), que seria um estimador de volatilidade em multiescalas. No entanto, o estimador com várias escalas necessita de uma solução numérica e torna o processo computacionalmente mais custoso. Outra possível abordagem para melhorar a estimação é o relaxamento de algumas hipóteses como a alocação regular do tempo e a independência dos retornos. Em relação ao modelo de um fator seria interessante estudar períodos com maior volatilidade e opções mais distantes do vencimento para entender como o delta relativo se comportaria.

Para trabalhos futuros seria interessante conectar a abordagem de se incluir modelo de ruído para a microestrutura no ativo arriscado, que representa o prêmio de risco para o mercado incompleto, para estimação de volatilidade realizada com a teoria de apreçamento de derivativos com volatilidade estocástica como proposto por (Fouque, Papanicolaou & Sircar 2000) que acrescenta termos de correção aos modelos de apreçamento dado a incompletude do mercado.

Bibliografia

- Aït-Sahalia, Y., Mykland, P., & Zhang, L. (2010). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*.
- Barndorff-Nielsen, O. E. & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 64(2), 253–280.
- Billingsley, P. (2008). *Probability and measure*. Wiley-India.
- Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *The journal of political economy*, 81(3), 637–654.
- Delbaen, F. & Schachermayer, W. (1995a). The existence of absolutely continuous local martingale measures. *The Annals of Applied Probability*, 5(4), 926–945.
- Delbaen, F. & Schachermayer, W. (1995b). The no-arbitrage property under a change of numéraire. *Stochastics An International Journal of Probability and Stochastic Processes*, 53(3), 213–226.
- Delbaen, F. & Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(2), 215–250.
- Duffie, D. (2001). *Dynamic asset pricing theory*. Princeton Univ Pr.
- Föllmer, H., Schweizer, M., & Bonn, R. F.-W.-U. (1990). *Hedging of contingent claims under incomplete information*. Rheinische Friedrich-Wilhelms-Universität Bonn.

- Föllmer, H. & Sondermann, D. (1985). Hedging of non-redundant contingent claims.
- Fouque, J., Papanicolaou, G., & Sircar, K. (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge University.
- Hull, J. (2009). *Options, futures and other derivatives*. Pearson Prentice Hall.
- Jacod, J. & Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*, 26(1), 267–307.
- Jacod, J. & Shiryaev, A. (1987). *Limit theorems for stochastic processes*, volume 2003. Springer Berlin.
- Karatzas, I. & Shreve, S. (1991). *Brownian motion and stochastic calculus*. Springer.
- Korn, R. & Korn, E. (2001). *Option pricing and portfolio optimization*. American Mathematical Society.
- Merton, R. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141–183.
- Mykland, P. (2000). Conservative delta hedging. *Annals of Applied Probability*, 664–683.
- Mykland, P. & Zhang, L. (2008). Inference for volatility-type objects and implications for hedging. *Statistics and its Interface*, 1, 255–278.
- Mykland, P. A. & Zhang, L. (2006). ANOVA for diffusions and Itô processes. *Ann. Statist.*, 34(4), 1931–1963.
- Renault, E. & Touzi, N. (1996). Options hedging and implicit volatilities in a stochastic volatility model. *Mathematical Finance*, 6(3), 279–302.
- Rootzen, H. (1980). Limit distributions for the error in approximations of stochastic integrals. *The Annals of Probability*, 8(2), 241–251.
- Schweizer, M. (1990). Risk-minimality and orthogonality of martingales. *Stochastics An International Journal of Probability and Stochastic Processes*, 30(2), 123–131.

- Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2), 339–363.
- Zhang, L. (2001). *From martingales to ANOVA: Implied and realized volatility*. University of Chicago, Dept. of Statistics.
- Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6), 1019–1043.
- Zhang, L. (2009). Implied and realized volatility: empirical model selection. *Annals of Finance*, 1–17.
- Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 100(472), 1394–1411.