

Universidade de Brasília
Instituto de Ciências Exatas

**O Modelo de Volatilidade Estocástica de
Heston: Análise Quantitativa para o
Mercado de Opções Brasileiro**

Autor: Pedro Araújo Elias Dib

Orientador: Alberto Masayoshi Faria Ohashi

Brasília, 9 de janeiro de 2026

O Modelo de Volatilidade Estocástica de Heston: Análise Quantitativa para o Mercado de Opções Brasileiro

Dissertação apresentada ao Instituto de Ciências Exatas da Universidade de Brasília, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Orientador: Prof. Dr. Alberto Masayoshi Faria Ohashi

Brasília, 9 de janeiro de 2026

Pedro Araújo Elias Dib

**O Modelo de Volatilidade Estocástica de Heston: Análise
Quantitativa para o Mercado de Opções Brasileiro**

Dissertação aprovada como requisito parcial
para obtenção do grau de Mestre em Mate-
mática pela Universidade de Brasília, em 11
de dezembro de 2025, pela seguinte banca
examinadora:

Prof. Dr. Alberto Masayoshi Faria Ohashi
(Orientador – MAT/UnB)

Prof. Dr. Paulo Henrique Pereira da Costa
(Membro Interno – MAT/UnB)

Prof.^a Dr.^a Cátia Regina Gonçalves
(Membro Suplente – MAT/UnB)

Prof. Dr. Dorival Leão
(Membro Externo – EstatCamp)

*Cada letra, vírgula e ponto neste texto,
eu dedico à minha mãe.*

Agradecimentos

Em primeiro lugar, agradeço à minha mãe por todos os sacrifícios feitos desde a minha infância para que eu tivesse uma boa educação, por nunca ter medido esforços para que eu pudesse priorizar os estudos e por ter me oferecido, além da vida, a paixão pelo conhecimento. Agradeço também à minha tia Layanne, por ser minha maior companheira desde que eu nasci, minha melhor amiga. Por fim, agradeço a meu padrasto Murilo por sempre cuidar de mim com tanto carinho.

Agradeço ao meu orientador pela paciência ao longo do desenvolvimento deste trabalho, pela liberdade intelectual concedida e pela orientação. Serei também sempre grato à professora Cátia Regina, por todas as broncas, conselhos e conversas, e por ter acreditado em mim. Sem ela, este trabalho não existiria.

Sou grato a Matheus, Pedro, Rafael e Sofia, amizades que fiz nesses dois anos, que sofreram comigo nas disciplinas, me ajudaram quando tive dúvidas, compreenderam meus problemas e compartilharam os seus. Eles tornaram meus dias mais felizes em Brasília, e guardarei com carinho a memória de sua presença em minha vida.

Aos meus amigos em Goiânia, Ana Luisa, Henrique, João Paulo, Leonardo, Samuel, Tatyana e Vinicius, agradeço por sua amizade fiel que mantemos ao longo de mais de uma década. Em especial, agradeço a Gabriel Mamede, por retomarmos nossa antiga amizade, que sempre me foi tão cara, por sua companhia nas estradas entre Goiânia e Brasília e pelas longas horas de conversa que conduzimos. Sou grato também à minha namorada Luna, por trazer sua doçura à minha realidade, por me ensinar a sempre enxergar o copo meio cheio e por me ajudar a atravessar este difícil ano de 2025.

Resta agradecer a Carlos e Laryssa, a família que me escolheu para fazer parte dela e que agora leva um pedacinho de mim para outro continente. Sou extremamente grato pela amizade verdadeira que construímos ao longo desses anos.

Por fim, agradeço à agência de fomento CAPES pelo apoio financeiro concedido durante o período de realização deste mestrado.

“...数学者も、美しい作品を創り出す作家や作曲家と同じ意味で「詩人」であり得るのではないかと考えています。”

— Kiyoshi Itô

(Acredito que os matemáticos também são verdadeiros poetas, assim como os romancistas e compositores musicais...)

Resumo

A hipótese de volatilidade constante, central no modelo de Black-Scholes, diverge das observações empíricas do mercado, falhando em capturar o fenômeno do "sorriso de volatilidade". Este trabalho analisa o modelo de volatilidade estocástica de Heston como uma alternativa mais realista para a precificação de opções europeias no mercado brasileiro. O desenvolvimento teórico fundamenta-se no cálculo estocástico para derivar a dinâmica do preço do ativo e da variância, utilizando a transformada de Fourier para a obtenção de uma solução semi-analítica. A implementação numérica e a estimação dos parâmetros, visando minimizar o erro entre as volatilidades implícitas teóricas e as observadas nos dados, constituem o objetivo principal do texto. Os resultados indicam que o modelo de Heston é capaz de se ajustar de maneira satisfatória às superfícies de volatilidade implícita observadas no mercado brasileiro, dentro das limitações impostas pela qualidade dos dados disponíveis e pela metodologia adotada.

Palavras-chave: Cálculo Estocástico. Modelo de Heston. Volatilidade Estocástica. Estimação de Parâmetros. Métodos Numéricos.

Abstract

The constant volatility hypothesis, central to the Black-Scholes model, diverges from empirical market observations, failing to capture the "volatility smile" phenomenon. This work analyzes the Heston stochastic volatility model as a more realistic alternative for pricing European options in the Brazilian market. The theoretical framework relies on stochastic calculus to derive the asset price and variance dynamics, employing the Fourier transform to obtain semi-analytical solutions. Numerical implementation and parameter estimation, aimed at minimizing the error between theoretical and observed implied volatilities, constitute the main objective of this text. The results indicate that the Heston model yields a satisfactory fit to the implied volatility surfaces observed in the Brazilian market, subject to the limitations arising from data quality and the adopted methodology.

Keywords: Stochastic Calculus. Heston Model. Stochastic Volatility. Parameter Estimation. Numerical Methods.

Sumário

Agradecimentos	4
Resumo	6
Abstract	7
Introdução	10
1 Preliminares	11
1.1 Opções	11
1.2 Noções gerais de Cálculo Estocástico	15
1.2.1 A Integral de Itô	15
1.2.2 Processos de Itô	21
1.2.3 O Teorema da Representação Martingale	24
1.2.4 O Teorema de Girsanov	25
1.3 Finanças Matemáticas	29
1.3.1 Mercados, carteiras e arbitragem	29
1.3.2 Replicabilidade e Completude	31
1.3.3 Precificação	32
2 Volatilidade Implícita e o Modelo de Heston	39
2.1 Volatilidade implícita	40
2.1.1 Implementação computacional	41
2.2 Preliminares	42
2.2.1 A decomposição de Cholesky	42
2.2.2 Orstein-Uhlenbeck e Cox-Ingersoll-Ross	46
2.3 O modelo de Heston	49
2.3.1 Descrição matemática do modelo	49
2.3.2 Implementação numérica	56
2.3.3 A SVI no modelo de Heston	58
3 Estimação de Parâmetros	61
3.1 Fundamentos	61
3.1.1 O Método dos Mínimos Quadrados	61
3.1.2 Implicações da escolha da função custo	64
3.1.3 Desafios da estimação	64
3.1.4 Otimização não convexa e o algoritmo AC-PG	66

3.2	Metodologia	67
3.2.1	Descrição dos dados	67
3.2.2	Limpeza dos dados	68
3.2.3	Estratégia de estimação	69
3.3	Resultados	71
3.3.1	Verificação com <i>Multi-Start</i>	76
3.4	Melhorando a performance	77
3.4.1	Preços	77
3.4.2	Otimização	78
3.5	Limitações do modelo de Heston	79
3.6	Alternativas ao modelo de Heston	80
3.6.1	O paradigma da Rough Volatility	81
A	Disponibilidade dos códigos	85
B	Resultados do <i>multi-start</i>	86
	Referências	90

Introdução

A compreensão da volatilidade do ativo subjacente é fundamental na precificação e gestão de riscos de derivativos financeiros. O modelo de Black–Scholes, embora forneça uma solução analítica elegante para o preço de opções europeias, fundamenta-se na hipótese restritiva de volatilidade constante. A incompatibilidade dessa suposição com a realidade observada no mercado manifesta-se principalmente através do "sorriso de volatilidade", fenômeno que evidencia sua natureza estocástica.

Como resposta a essas limitações, surgiram os modelos de volatilidade estocástica, nos quais a própria volatilidade é modelada como um processo aleatório. Dentre esses, o modelo de Heston destacou-se por capturar dinâmicas mais realistas de volatilidade mantendo a tratabilidade analítica por meio de funções características, obtendo expressões quase fechadas para o preço de opções *vanilla*.

Partimos da modelagem financeira para o problema de estimação de parâmetros, uma aplicação prática importante desses modelos. Mais especificamente, estamos interessados em determinar o conjunto de parâmetros que minimiza a diferença entre valores observados no mercado, como preço e volatilidade implícita, e valores teóricos. Abordamos a tarefa de estimação como um problema de otimização não linear e não convexo, potencialmente mal-posto e sensível às condições iniciais. Para sua resolução, adotamos o algoritmo de gradiente projetado *Auto-Conditioned Projected Gradient* (AC-PG), que ajusta automaticamente o tamanho do passo a partir de informações locais da função custo, dispensando a escolha manual de parâmetros de passo.

A metodologia é aplicada a dados reais de opções de ações negociadas na B3. A análise empírica permite avaliar o desempenho do procedimento de estimação em diferentes ativos, bem como interpretar financeiramente os parâmetros obtidos e identificar situações nas quais o ajuste do modelo se mostra limitado, em especial em função da qualidade dos dados disponíveis.

Por fim, o trabalho situa o modelo de Heston em um quadro mais amplo de abordagens para a descrição da volatilidade, discutindo suas limitações e alternativas presentes na literatura. Em particular, é apresentado o paradigma da *rough volatility*, motivado por evidências empíricas que indicam uma estrutura temporal da volatilidade consideravelmente mais irregular do que aquela capturada pelos modelos clássicos. Embora a estimação de parâmetros em modelos rugosos não seja o objetivo deste estudo, a apresentação do modelo *Rough Heston* permite contextualizar o modelo clássico como um caso limite de uma classe mais geral de modelos, além de apontar direções naturais para investigações futuras.

Capítulo 1

Preliminares

Esta seção toma como referência principal Pfützenreuter (2021).

1.1 Opções

No mercado financeiro, são exemplos padrão de títulos derivativos, isto é, títulos cujo valor depende dos preços de outros títulos mais básicos (chamados títulos primários ou ativos subjacentes) como ações, por exemplo. Uma opção pode ser semelhante a uma apólice de seguro: a seguradora emite uma apólice e, em troca, cobra uma taxa - o prêmio - cujo valor garantirá o seu patrimônio durante a vigência dessa apólice.

Existem opções sobre inúmeros tipos de ativos, por exemplo, sobre contratos de futuros, mas trabalharemos apenas com opções sobre ações.

Nosso foco será nas opções denominadas *vanilla*, isto é,

- * Opções de compra (*call*): título que dá o direito de comprar algo a um preço determinado.
- * Opções de venda (*put*): título que dá o direito de vender algo a um preço determinado.

Opções de compra e venda

Voltando à analogia com o seguro: se contratamos um seguro para o carro e ele permanecer intacto, ficamos com ele. Entretanto, se o carro for roubado ou der perda total, recebemos o dinheiro da seguradora. O seguro pode ser visto como uma "opção de venda" do carro à seguradora.

O uso mais natural da opção de venda é a garantia do capital, a proteção do investimento.

Por outro lado, podemos enxergar as opções de compra como o sinal pago ao comprar um imóvel: queremos garantir a preferência na compra de um imóvel, então damos um sinal (uma pequena porcentagem do valor total) e, se concretizarmos o negócio, ele constitui parte do pagamento. Se desistirmos, perdemos o valor do sinal.

O uso mais comum de uma opção de compra é planejar uma compra futura, garantindo o preço de hoje.

As opções de venda podem ser também adquiridas com fins especulativos: em vez de planejar um investimento futuro, pode-se simplesmente especular que o prêmio de uma opção vai subir.

Valor intrínseco e extrínseco

No dia do pagamento da opção, só vale a pena exercer a opção se ela nos garantir um preço melhor que o de mercado. Para uma opção vantajosa, a diferença entre o preço de mercado e o preço garantido é chamada **valor intrínseco**. No caso das opções, ele é sempre não-negativo.

Para o comprador da opção, um valor intrínseco maior é sempre melhor, tanto para opções de venda quanto de compra. Já para o vendedor, é mais interessante um valor intrínseco menor (idealmente, zero), para *calls* e *puts*.

Antes do vencimento, é natural que o prêmio da opção seja bem maior que o valor intrínseco. Chamamos de **valor extrínseco** a diferença entre o preço da opção e seu valor intrínseco.

Moneyness

Como a opção é um derivativo, seu prêmio depende totalmente do preço atual (*spot*) do ativo subjacente. A relação entre preço de exercício (*strike*) e *spot* é de extrema importância e recebe o nome de *moneyness*, apresentando-se numa das três formas a seguir:

- * Valor intrínseco positivo, chamado *in the money* (ITM). Se tudo continuar como está, a opção vencerá valendo alguma coisa.
- * Valor intrínseco zero e *spot* significativamente distante do preço de exercício. Dizemos que a opção está *out of the money* (OTM). Se tudo continuar como está, a opção irá expirar sem valor (informalmente, chamamos de "virar pó").
- * Valor intrínseco próximo de zero e *spot* muito próximo ou igual ao *strike*. Nesse caso, dizemos que a opção está *at the money* (ATM). Nesse caso, o *spot* pode mover-se tanto para um lado quanto para outro, jogando a opção para ITM ou OTM.

Independentemente do ativo subjacente ou do mercado específico, opções com o mesmo *moneyness* exibem comportamento análogo:

- * Uma opção ITM tende a acompanhar mais de perto as variações de preço *spot* do ativo subjacente, e seu valor extrínseco fica cada vez menor.
- * Uma opção ATM possui o maior valor extrínseco da série, ou seja, seu prêmio é o "mais caro". Também tende a ser a opção mais negociada.
- * Uma opção OTM tem valor extrínseco menor que uma opção ATM. Quanto mais fortemente OTM, menos o prêmio acompanha as variações do *spot*.

Opções na B3

O *ticket* (símbolo) de uma opção da B3 possui a seguinte forma:

AAAASNNN

onde:

- * AAAA : prefixo da ação subjacente.
- * S : letra que indica o código de vencimentos, conforme a tabela 1.1 a seguir.
- * NNN : número de dois ou três dígitos que identifica o valor garantido (preço de exercício).

Por exemplo, PETRA50 é uma opção de compra sobre a ação PETR4 (Petrobras preferencial), que provavelmente vence em janeiro e provavelmente tem preço de exercício de \$ 50.

Entretanto, nem sempre o sufixo NNN corresponde ao valor do strike. É preciso consultar os dados da opção para descobrir o valor real do preço de exercício.

Vencimento	Opção de Compra	Opção de Venda
Janeiro	A	M
Fevereiro	B	N
Março	C	O
Abril	D	P
Maior	E	Q
Junho	F	R
Julho	G	S
Agosto	H	T
Setembro	I	U
Outubro	J	V
Novembro	K	W
Dezembro	L	X

Tabela 1.1: Códigos de vencimento de opções na B3

Parâmetros de uma opção

O prêmio (ou pagamento) de uma opção é completamente determinado pelos seguintes parâmetros:

- * Spot *S*: preço atual do ativo subjacente à opção.
- * Strike (ou preço de exercício) *K*: preço do ativo subjacente que a opção garante ao seu detentor.
- * Vencimento *T*: data limite para o exercício da opção, após a qual a opção deixa de existir.

- * Estilo: define em que datas ou faixas de tempo a opção pode ser exercida (europeia, americana, etc).
- * Direito garantido: opção de compra ou opção de venda.
- * Taxa base de juros r : no Brasil, a taxa SELIC, paga pelo governo.
- * Volatilidade σ : oscilação dos rendimentos do ativo subjacente.

Na tabela abaixo, resumimos as informações acerca do *moneyness* de uma opção:

Condição	Opção de Compra (Call)	Opção de Venda (Put)
$S > K$	In the Money (ITM) Valor Intrínseco = $S - K$	Out of the Money (OTM) Sem valor intrínseco
$S = K$	At the Money (ATM) Prêmio = 100% Valor Extrínseco	At the Money (ATM) Prêmio = 100% Valor Extrínseco
$S < K$	Out of the Money (OTM) Sem valor intrínseco	In the Money (ITM) Valor Intrínseco = $K - S$

Tabela 1.2: Moneyness e valor intrínseco de opções

Referências principais adotadas no restante deste capítulo: Øksendal (2003), Musiela e Rutkowski (2005), Wiersema (2008), Oosterlee e Grzelak (2020) e Shreve (2004).

1.2 Noções gerais de Cálculo Estocástico

A conhecida não diferenciabilidade das trajetórias do movimento browniano impede a aplicação direta de teoremas do cálculo tradicional no contexto de funções desse processo estocástico.

Em 1944, Kiyoshi Itô, no artigo intitulado "*Stochastic Integral*", estabeleceu uma maneira de resolver esse tipo de problema. No texto, ele introduz um novo tipo de integral e mostra o resultado que ficou conhecido posteriormente como *fórmula de Itô*. Essa fórmula é frequentemente vista como análoga ao Teorema Fundamental do Cálculo, agora expresso em termos de integrais de Itô, e que acrescenta um "termo de correção", resultante da variação quadrática não nula do movimento browniano. Ver mais detalhes em Kuo (2006).

1.2.1 A Integral de Itô

Para simplificar a notação, omitiremos o termo $\omega \in \Omega$ sempre que possível. Além disso, as notações W_t e $W(t)$ serão utilizadas de maneira intercambiável ao longo do texto, ambas representando o valor do movimento browniano W no tempo t (a menos que seja explicitamente definido de outra forma).

Começemos com alguns conceitos fundamentais para o estudo do cálculo estocástico.

Definição 1.1. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Um **movimento browniano** (unidimensional), também conhecido como *processo de Wiener*, é um processo estocástico caracterizado pelas seguintes propriedades:

- (i) $W_0 = 0$ (mais precisamente, $\mathbb{P}(W_0 = 0) = 1$),
- (ii) W_t é contínuo q.c. $[\mathbb{P}]$,
- (iii) W_t possui incrementos independentes e normalmente distribuídos, isto é, $W_t - W_s \sim N(0, t - s)$, para $s < t$.

Um vetor n -dimensional $W = (W_1(t), \dots, W_n(t))$ onde

- (i) Cada W_i é um processo de Wiener unidimensional,
- (ii) W_i e W_j são independentes para $i \neq j$,

é chamado um **movimento browniano n -dimensional**.

Definição 1.2. Seja W_t um movimento browniano n -dimensional. Definimos $\mathcal{F}_t = \mathcal{F}_t^{(n)}$ como sendo a σ -álgebra gerada pelas variáveis aleatórias $\{W_i(s)\}$, com $1 \leq i \leq n$ e $0 \leq s \leq t$. Em outras palavras, \mathcal{F}_t é a menor σ -álgebra contendo todos os conjuntos da forma

$$\{\omega : W_{t_1}(\omega) \in B_1, \dots, W_{t_k}(\omega) \in B_k\},$$

onde $t_j \leq t$ e $B_j \in \mathcal{B}(\mathbb{R}^n)$ são borelianos, $j = 1, 2, \dots$ (Assumimos que \mathcal{F}_t é completa).

Definição 1.3. Seja $\{\mathcal{N}_t\}_{t \geq 0}$ uma sequência crescente de σ -álgebras de Ω . Um processo $g(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}^n$ é dito ser \mathcal{N}_t -**adaptado** se, para cada $t \geq 0$, a função $\omega \rightarrow g(t, \omega)$ é \mathcal{N}_t -mensurável.

Definição 1.4. Seja $0 \leq t < T$ e $\mathcal{V} = \mathcal{V}(0, T)$ a classe de funções $f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ tais que:

- (i) $(t, \omega) \rightarrow f(t, \omega)$ é $\mathcal{B} \times \mathcal{F}$ -mensurável,
- (ii) $f(t, \omega)$ é \mathcal{F}_t -adaptada,
- (iii) $\mathbb{E} \int_0^T f(t, \omega)^2 < \infty$.

Definição 1.5. Dados $X_t, Y_t : \Omega \rightarrow \mathbb{R}$ processos estocásticos, definimos a **variação quadrática** de X_t como sendo, para $\omega \in \Omega$, o processo

$$\langle X, X \rangle_t = \lim_{\Delta t_k \rightarrow 0} \sum_{t_k \leq t} |X_{t_{k+1}} - X_{t_k}|^2 \text{ (limite em probabilidade)}, \quad (1.1)$$

onde $0 = t_1 < t_2 < \dots < t_n = t$ e $\Delta t_k = t_{k+1} - t_k$. Similarmente, a **variação cruzada** de X_t e Y_t é dada pelo processo

$$\langle X, Y \rangle_t = \lim_{\Delta t_k \rightarrow 0} \sum_{t_k \leq t} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}). \quad (1.2)$$

Teorema 1.6. *Seja W um movimento browniano. Então, $\langle W, W \rangle_t = t$, $\forall t \geq 0$, quase certamente.*

Demonstração. Shreve (2004). □

Informalmente, escrevemos esse fato como

$$d\langle W, W \rangle_t = dW(t)dW(t) = (dW_t)^2 = dt.$$

Observação 1.1. Variação cruzada.

Além de calcularmos a variação quadrática de W , da mesma forma podemos também calcular as variações cruzadas e quadrática entre W e t , resultando em: $dt dW(t) = dW(t)dt = (dt)^2 = 0$, como detalhado em Shreve (2004).

Observação 1.2. A caracterização de Levy.

Dentre as várias caracterizações existentes para o MB, uma que se destaca por sua elegância é a caracterização de Levy, que constitui uma espécie de recíproca do teorema 1.6: se $\{X_t\}$ é um processo contínuo, com $X_0 = 0$, tal que X_t é uma martingale com respeito à \mathcal{F}_t (definição 1.11) e $\langle X, X \rangle_t = t$, $\forall t \in [0, T]$, então X_t é um movimento browniano em \mathcal{F}_t .

Introduzimos, a seguir, a definição de integral de Itô.

Definição 1.7. Sejam ϕ um processo simples, $f \in \mathcal{V}$ e $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$ uma partição de $[0, T]$. Definimos a **integral de Itô** de ϕ e f como sendo:

$$\int_0^t \phi(u) dW_u = \sum_{j=0}^{k-1} \phi(t_j) [W_{t_{j+1}} - W_{t_j}] + \phi(t_k) [W_t - W_{t_k}] \quad (1.3)$$

e

$$\int_0^t f(u) dW_u = \lim_{n \rightarrow \infty} \int_0^t \phi_n(u) dW_u \quad (\text{limite em } L^2(P)), \quad (1.4)$$

onde $\{\phi_n\}$ é uma sequência de funções simples tal que $\|f - \phi_n\|_2 \rightarrow 0$ quando $n \rightarrow \infty$.

Exemplo 1.8. Calculando $\int_0^T W(t) dW_t$ pela definição.

Tome $n \in \mathbb{Z}$ grande. Vamos aproximar $f(t) = W(t)$ pela seguinte sequência de funções simples:

$$\phi_n(t) = \begin{cases} W(0) = 0 & , 0 \leq t < \frac{T}{n}, \\ W\left(\frac{T}{n}\right) & , \frac{T}{n} \leq t < \frac{2T}{n}, \\ \vdots & \vdots \\ W\left(\frac{(n-1)T}{n}\right) & , \frac{(n-1)T}{n} \leq t < T. \end{cases}$$

Pode-se mostrar que

$$\|W - \phi_n\|_2 \rightarrow 0, \quad n \rightarrow \infty.$$

Agora, por definição,

$$\begin{aligned} \int_0^T W(t) dW_t &= \lim_{n \rightarrow \infty} \int_0^T \phi_n(t) dW_t \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right) \left\{ W\left[\frac{(j+1)T}{n}\right] - W\left[\frac{jT}{n}\right] \right\}. \end{aligned}$$

Para simplificar a notação, façamos $W_j = W\left(\frac{jT}{n}\right)$. Antes de avaliar o limite acima, vamos obter algumas expressões que serão úteis.

$$\begin{aligned} \frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1} - W_j)^2 &= \frac{1}{2} \sum_{j=0}^{n-1} W_{j+1}^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} \sum_{k=1}^n W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \quad (\text{com } k = j + 1). \\ &= \frac{1}{2} W_n^2 + \frac{1}{2} \sum_{k=1}^{n-1} W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \quad (W_0 = 0) \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j^2 - \sum_{j=0}^{n-1} W_j W_{j+1} \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j (W_j - W_{j+1}). \end{aligned}$$

Da expressão acima, concluímos que

$$\sum_{j=0}^{n-1} W_j(W_j - W_{j+1}) = \frac{1}{2}W_n^2 - \frac{1}{2}\sum_{j=0}^{n-1}(W_{j+1} - W_j)^2.$$

Na notação anterior,

$$\begin{aligned} \sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right) \left\{ W\left[\frac{jT}{n}\right] - W\left[\frac{(j+1)T}{n}\right] \right\} &= \frac{1}{2}W(T)^2 \\ &\quad - \frac{1}{2}\sum_{j=0}^{n-1} \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right) \right]^2. \end{aligned}$$

Tomando $n \rightarrow \infty$, obtemos

$$\int_0^T W(t)dW_t = \frac{1}{2}W^2(T) - \frac{1}{2}\langle W, W \rangle(T) = \frac{1}{2}W^2(T) - \frac{1}{2}T.$$

□

O resultado a seguir relaciona integrais de Itô e integrais de Lebesgue.

Lema 1.9 (Isometria de Itô). *Para toda $f \in \mathcal{V}$, vale a seguinte relação:*

$$\mathbb{E}\left(\int_s^t f(u)dW_u\right)^2 = \mathbb{E}\int_s^t f^2(u)dt. \quad (1.5)$$

Demonstração. Oosterlee e Grzelak (2020). □

Algumas propriedades que seguem da definição:

Proposição 1.10. *Sejam $f, g \in \mathcal{V}(0, T)$ e considere $0 < t < T$. Temos:*

- (a) $\int_0^T f_u dW_u = \int_0^t f_u dW_u + \int_t^T f_u dW_u$, para quase todo ω ;
- (b) $\int_0^T (cf + g)_u dW_u = c \int_0^T f_u dW_u + \int_0^T g_u dW_u$, c constante, para quase todo ω ;
- (c) $\int_0^T f_u dW_u$ é \mathcal{F}_T -mensurável.

Definição 1.11. Uma **filtração** em (Ω, \mathcal{F}) é uma família crescente $\mathcal{M} = \{\mathcal{M}_t\}_{t \geq 0}$ de σ -álgebras em \mathcal{F} . Um processo n -dimensional $\{X_t\}_{t \geq 0}$ em $(\Omega, \mathcal{F}, \mathbb{P})$ é dito ser uma **martingale** com respeito à \mathcal{M} (sob a medida \mathbb{P}) se, para todo t ,

- (i) X_t é adaptado, isto é, X_t é \mathcal{M}_t -mensurável
- (ii) X_t é integrável, isto é, $\mathbb{E}|X_t| < \infty$

- (iii) e X_t não apresenta tendência sistemática de crescimento ou queda, ou seja, $\mathbb{E}[X_s | \mathcal{M}_t] = X_t, \forall s \geq t$.

Observação 1.3. Martingales e preços de ativos.

Os processos de preços de ativos considerados neste trabalho, denotados por S_t , satisfazem $\mathbb{E}|S_t| < \infty, \forall t \geq 0$, já que exigimos integrabilidade na definição 1.11 acima.

Essa condição é essencial, pois a precificação de derivativos financeiros, como as opções, depende fortemente do cálculo do valor esperado de um pagamento futuro com respeito a um processo subjacente. Assumimos que esse processo seja uma martingale, uma vez que essa propriedade está intimamente relacionada à ausência de arbitragem.

Exemplo 1.12. O movimento browniano W_t em \mathbb{R}^n é uma martingale com respeito à $\mathcal{F}_t = \sigma(W_s, s \leq t)$.

De fato, $W_t = (W_t^1, \dots, W_t^n)$ satisfaz as condições da definição anterior:

- (i) imediato pois W_t é \mathcal{F}_t -mensurável por construção;
(ii) analisando o 2º momento, obtemos, da desigualdade de Jensen:

$$\begin{aligned} (\mathbb{E}|W_t|)^2 &\leq \mathbb{E}\|W_t\|^2 = \mathbb{E} \sum_{i=1}^n (W_t^i)^2 = \sum_{i=1}^n \mathbb{E}(W_t^i)^2 \\ &= \sum_{i=1}^n [\text{var } W_t^i + (\mathbb{E}W_t^i)^2] = nt \end{aligned}$$

$\therefore \mathbb{E}|W_t| < \sqrt{nt} < \infty, \forall t$, e W_t é integrável;

- (iii) Por fim, tome $s \geq t$. Basta notar que $W_s = W_t + (W_s - W_t)$. Assim, das propriedades de esperança condicional e da definição do movimento browniano, segue que:

$$\begin{aligned} \mathbb{E}[W_s | \mathcal{F}_t] &= \mathbb{E}[W_t + (W_s - W_t) | \mathcal{F}_t] \\ &= \mathbb{E}[W_t | \mathcal{F}_t] + \mathbb{E}[(W_s - W_t) | \mathcal{F}_t] \\ &= W_t + 0 \\ &= W_t. \end{aligned}$$

□

Outra propriedade importantíssima das integrais de Itô é que elas são martingales:

Proposição 1.13. Para todo $f \in \mathcal{V}(0, T)$, existe uma \mathcal{F}_t martingale em $[0, T]$, $\{I(t)\}$, tal que, para todo $t \in [0, T]$,

$$I(t) = \int_0^t f(w, s) dW_s. \quad (1.6)$$

Em particular,

$$\mathbb{E} \int_0^t f(w, s) dW_s = 0. \quad (1.7)$$

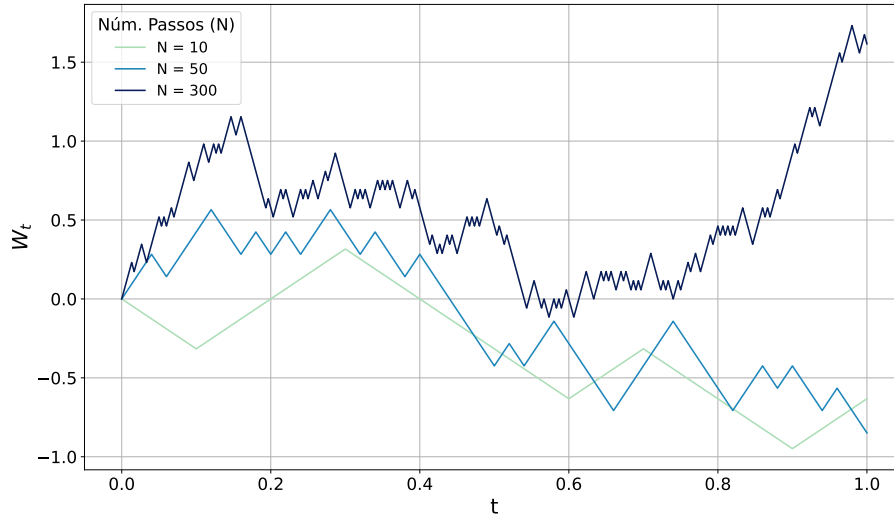


Figura 1.1: O movimento browniano padrão como limite de um passeio aleatório.

Demonstração. Kuo (2006). □

Podemos também estender a noção de Integral de Itô.

Definição 1.14. Seja $W = (W_1, \dots, W_n)$ um movimento browniano n -dimensional. Denote por $\mathcal{V}_{\mathcal{H}}^{m \times n}(S, T)$ a classe das matrizes $v = [v_{ij}(t, w)]_{m \times n}$ tais que cada v_{ij} satisfaz as seguintes propriedades:

- (i) $(t, w) \rightarrow v_{ij}(t, w)$ é $\mathcal{B} \times \mathcal{F}$ -mensurável,
- (ii) existe uma filtração $\mathcal{H} = (\mathcal{H}_t)$, $t \geq 0$, tal que W_t é uma martingale com respeito à \mathcal{H}_t e f_t é \mathcal{H}_t -adaptado,
- (iii) $\mathbb{E} \int_0^T f(t, w)^2 dt < \infty$.

Para $v \in \mathcal{V}_{\mathcal{H}}^{m \times n}(S, T)$, definimos, utilizando notação matricial,

$$\int_0^T v dW = \int_0^T \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mn} \end{pmatrix} \begin{pmatrix} dW_1 \\ \vdots \\ dW_n \end{pmatrix} \quad (1.8)$$

como sendo o vetor coluna cuja i -ésima componente é a seguinte soma de integrais de Itô unidimensionais:

$$\sum_{j=1}^n \int_S^T v_{ij}(s) dW_j. \quad (1.9)$$

Por fim, se $\mathcal{H} = \mathcal{F}^{(n)} = \{\mathcal{F}_t^{(n)}\}_{t \geq 0}$, escrevemos $\mathcal{V}^{m \times n}(S, T)$ e, se $m = 1$, escrevemos $\mathcal{V}_{\mathcal{H}}^n(S, T)$ (respectivamente, $\mathcal{V}^n(S, T)$). Também, fazemos

$$\mathcal{V}^{m \times n} = \mathcal{V}^{m \times n}(0, \infty) = \bigcap_{T > 0} \mathcal{V}^{m \times n}(0, T). \quad (1.10)$$

Exemplo 1.15. Caso $n = 2$.

Para trazer mais clareza à definição anterior, vamos considerar $m = 1$ e $m = n = 2$. Suponha também que estejamos trabalhando com a filtração natural \mathcal{F}_t .

Considere $W = (W_1(t), W_2(t))$. Pela definição anterior, a integral de Itô de uma função $v \in \mathcal{V}^2(0, T)$ é dada por

$$\int_0^T v dW = \int_0^T (v_{11} \ v_{12}) \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix} = \int_0^T v_{11}(s) dW_1 + \int_0^T v_{12}(s) dW_2 ,$$

um escalar.

Agora, tomando $u \in \mathcal{V}^{2 \times 2}(0, T)$, temos:

$$\int_0^T u dW = \int_0^T \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix} = \begin{pmatrix} \int_0^T u_{11}(s) dW_1 + \int_0^T u_{12}(s) dW_2 \\ \int_0^T u_{21}(s) dW_1 + \int_0^T u_{22}(s) dW_2 \end{pmatrix} ,$$

uma matriz 2×1 . □

Naturalmente, as propriedades da integral de Itô unidimensional também se estendem para o caso mais geral. Em especial, a propriedade de que se $W(t) = (W_1(t), \dots, W_n(t))$ é um movimento browniano e $v \in \mathcal{V}^n$, então a integral de Itô

$$X_t = X_0 + \int_0^t v(s) dW_s ; t \geq 0$$

é uma martingale com respeito à $\mathcal{F}_t^{(n)}$.

1.2.2 Processos de Itô

Definição 1.16. Definimos $\mathcal{W}_{\mathcal{H}}(0, T)$ como sendo a classe de processos $f(t, w) \in \mathbb{R}$ satisfazendo os itens (i) e (ii) da definição 1.14 e a condição abaixo:

$$\mathbb{P} \left[\int_0^T f(s, w)^2 ds < \infty \right] = 1.$$

Similarmente, tomamos

$$\mathcal{W}_{\mathcal{H}} = \bigcap_{T>0} \mathcal{W}_{\mathcal{H}}(0, T).$$

Definição 1.17. Seja W_t um movimento browniano unidimensional em $(\Omega, \mathcal{F}, \mathbb{P})$. Um **processo de Itô** (ou integral estocástica) unidimensional é um processo estocástico X_t da forma

$$X_t = X_0 + \int_0^t u(s) ds + \int_0^t v(s) dW_s , \quad (1.11)$$

onde $v \in \mathcal{W}_{\mathcal{H}}$, de modo que

$$\mathbb{P} \left(\int_0^t v^2(s, w) ds < \infty \forall t \geq 0 \right) = 1. \quad (1.12)$$

Além disso, u é \mathcal{H}_t -adaptado e satisfaz

$$\mathbb{P}\left(\int_0^t |v(s, w)| ds < \infty \forall t \geq 0\right) = 1. \quad (1.13)$$

É comum representarmos processos de Itô da forma (1.11) na notação diferencial:

$$dX_t = u dt + v dW_t, \quad X(0) = X_0. \quad (1.14)$$

Proposição 1.18. *Sejam X_t, Y_t processos de Itô da forma $dX_t = u_1 dt + v_1 dW_t$ e $dY_t = u_2 dt + v_2 dW_t$. Então,*

$$\langle X, X \rangle_t = \int_0^t u_1^2(s) ds$$

e

$$\langle X, Y \rangle_t = \int_0^t v_1^2(s) v_2^2(s) ds.$$

Teorema 1.19 (A fórmula de Itô). *Sejam X_t um processo de Itô dado por*

$$dX_t = u dt + v dW_t \quad (1.15)$$

e $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$. *Nessas condições, $Y_t = g(t, X_t)$ é também um processo de Itô*

$$dY_t = \frac{\partial g}{\partial t}(t, X_t) dt + \frac{\partial g}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X_t) (dX_t)^2, \quad (1.16)$$

onde $(dX_t)^2 = (dX_t)(dX_t)$ é calculado de acordo com as regras

$$\begin{cases} dW_t dW_t = dt \\ dW_t dt = dt dW_t = 0 \\ dt dt = 0. \end{cases} \quad (1.17)$$

A fórmula de Itô facilita consideravelmente o cálculo das integrais de Itô, como podemos ver no exemplo a seguir:

Exemplo 1.20. Aplicando a fórmula de Itô.

Considere a função $g(t, x) = \frac{1}{2}x^2$. Utilizando o resultado no processo $Y_t = g(t, W_t)$, obtemos:

$$\begin{aligned} dY_t &= d\left(\frac{1}{2}W^2(t)\right) = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dW_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (dW_t)^2 \\ &= 0 + W_t dW_t + \frac{1}{2} (dW_t)^2 \\ &= W_t dW_t + \frac{1}{2} dt. \end{aligned}$$

$$\therefore \int_0^T W(t) dW_t = \frac{1}{2} W^2(T) - \frac{1}{2} T,$$

como havíamos visto. □

Exemplo 1.21. Movimento Browniano Geométrico.

Considere o seguinte problema de crescimento populacional, como apresentado em Øksendal (2003):

$$\frac{dS_t}{dt} = a_t S_t, \quad (1.18)$$

onde $a_t = \mu + \sigma W_t$, com μ, σ constantes e W_t um "ruído aleatório". Vamos procurar, de maneira informal, uma "interpretação de Itô" para esse problema. Escrevemos:

$$dS_t = \mu S_t dt + \sigma S_t dW_t. \quad (1.19)$$

Assim,

$$\int_0^t \frac{dS_s}{S_s} = \mu t + \sigma W_t \quad (W_0 = 0).$$

Para avaliar a integral do lado esquerdo, utilizamos a fórmula de Itô para a função $g(t, x) = \ln x$, $x > 0$:

$$\begin{aligned} d(\ln S_t) &= \frac{1}{S_t} dS_t + \frac{1}{2} \left(-\frac{1}{S_t^2} \right) (dS_t)^2 \\ &= \frac{dS_t}{S_t} - \frac{1}{2S_t^2} \sigma^2 S_t^2 dt \\ &= \frac{dS_t}{S_t} - \frac{1}{2} \sigma^2 dt \end{aligned}$$

$$\therefore \frac{dS_t}{S_t} = d(\ln S_t) + \frac{1}{2} \sigma^2 dt \Rightarrow \ln \frac{S_t}{S_0} = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t,$$

ou então,

$$S_t = S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\}. \quad (1.20)$$

Processos assim são chamados **movimentos brownianos geométricos** e são muito importantes para a teoria que se segue. \square

Em finanças, o movimento browniano geométrico surge como modelo para o preço de uma ação em alguns casos. Entretanto, qualquer carteira de investimentos simples nos leva a considerar a dinâmica de um *sistema* de processos de Itô.

Definição 1.22. Seja $W(t, w) = (W_1(t, w), \dots, W_m(t, w))$ um movimento browniano m -dimensional. Se cada um dos processos u_i e v_{ij} satisfaz as condições da definição de processo de Itô unidimensional, $1 \leq i \leq n$, $1 \leq j \leq m$, então podemos formar n processos de Itô

$$\begin{cases} dX_1 = u_1 dt + v_{11} dW_1 + \dots + v_{1m} dW_m \\ \dots \\ dX_n = u_n dt + v_{n1} dW_1 + \dots + v_{nm} dW_m \end{cases} \quad (1.21)$$

Ou, em notação matricial,

$$dX(t) = u dt + v dW(t), X(0) = X_0, \quad (1.22)$$

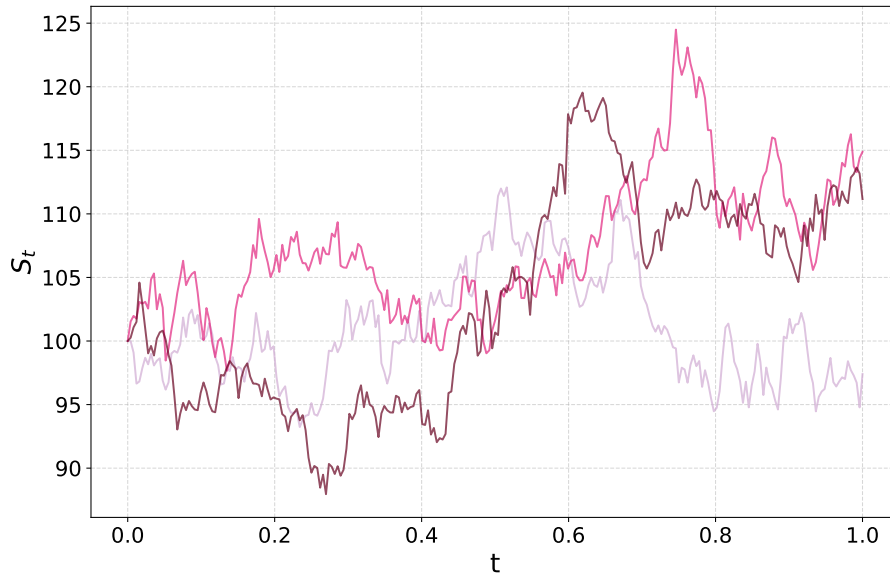


Figura 1.2: Simulação de 3 trajetórias de um Movimento Browniano Geométrico ao longo de um ano, com parâmetros: $S_0 = 100$, $\mu = 0.05$, e $\sigma = 0.2$.

onde

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \quad v = \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}, \quad dW(t) = \begin{pmatrix} dW_1(t) \\ \vdots \\ dW_m(t) \end{pmatrix}$$

Tal processo $X(t)$ é chamado um **processo de Itô n -dimensional**.

Vale ressaltar que resultados válidos para processos de Itô unidimensionais continuam válidos no caso n -dimensional.

1.2.3 O Teorema da Representação Martingale

Nesta seção, apresentaremos dois importantes teoremas de representação, muito utilizados no contexto de finanças.

Seja $W(t) = (W_1(t), \dots, W_n(t))$ um movimento browniano n -dimensional. Vimos anteriormente que a integral de Itô é uma martingale com respeito à $\mathcal{F}_t^{(n)}$. Veremos agora que, surpreendentemente, a recíproca é verdadeira: *qualquer $\mathcal{F}_t^{(n)}$ -martingale (sob a medida \mathbb{P}) pode ser representada como uma integral de Itô.*

Teorema 1.23 (Teorema da Representação de Itô). *Seja $F \in L^2(\mathcal{F}_t^{(n)}, \mathbb{P})$ uma variável aleatória. Então, existe um único processo estocástico $f(t, w) \in \mathcal{V}^n(0, T)$, tal que*

$$F(w) = \mathbb{E}F + \int_0^T f(t, w) dW(t). \quad (1.23)$$

Teorema 1.24 (Representação Martingale, unidimensional). *Sejam $X \in L^2(\mathbb{P})$ uma variável aleatória \mathcal{F}_t -mensurável. Então, existe um processo $\alpha \in \mathcal{V}(0, t)$ tal que:*

$$X = \mathbb{E}X + \int_0^t \alpha(s) dW_s. \quad (1.24)$$

Exemplo 1.25. Considere o processo $M_t = W_t^2 - t$.

Uma conta simples mostra que M_t é uma martingale com respeito à \mathcal{F}_t (note também que $\mathbb{E}|M_t| = 0 < \infty$).

Assim, pelo teorema da Representação Martingale, existe um processo estocástico α tal que

$$M_t = \int_0^t \alpha(s) dW_s.$$

De fato, vimos no exemplo 1.8 que

$$\int_0^t W_s dW_s = \frac{1}{2} (W_t^2 - t) = \frac{1}{2} M_t$$

Portanto, a martingale M_t pode ser representada como a integral de Itô

$$\int_0^t 2W_s dW_s,$$

com $2W_s = \alpha(s) \in \mathcal{V}(0, t)$. □

1.2.4 O Teorema de Girsanov

O Teorema de Girsanov fornece um resultado fundamental acerca da mudança na dinâmica de um processo, através da mudança de medida de probabilidade sob a qual ele está definido. Essencialmente, o teorema formaliza uma "mudança de referencial", permitindo alterar o termo de *drift* de um processo de Itô utilizando uma medida equivalente. Além disso, o resultado fornece explicitamente a derivada de Radon-Nikodym associada a essa mudança de medida.

A capacidade de transitar entre medidas controladamente é fundamental no contexto das finanças, como veremos adiante, pois será do nosso interesse *remover* o coeficiente de *drift* para realizarmos a precificação das opções.

Começemos com um exemplo de motivação:

Exemplo 1.26. Mudança de medida para uma variável aleatória normal.

Seja $X \sim N(0, 1)$ definida sob $(\Omega, \mathcal{F}, \mathbb{P})$, isto é,

$$\mu_X(B) = \mathbb{P}(X \in B) = \int_B n(x) dx = \int_B \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx, \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (1.25)$$

Em particular, tomando $B = (-\infty, x]$, obtemos:

$$\mathbb{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx, \quad \forall x \in \mathbb{R}. \quad (1.26)$$

Note que $\mathbb{E}X = 0$ e $\text{var } X = 1$. Agora, seja $\xi > 0$ uma constante e defina $Y := X + \xi$, de modo que, sob \mathbb{P} , $Y \sim N(\xi, 1)$.

Gostaríamos de obter uma nova medida \mathbb{Q} em Ω , sob a qual $Y \sim N(0, 1)$. Esse interesse surge pois, em finanças, mudar a medida de probabilidade nos permite substituir o retorno esperado do ativo, um parâmetro incerto e que varia com o risco, pela taxa de juros livre de risco, um parâmetro observável. Assim, conseguimos simplificar consideravelmente o procedimento de precificação.

Começamos definindo uma nova variável aleatória:

$$Z(\omega) = \exp \left\{ -\xi X(\omega) - \frac{1}{2}\xi^2 \right\} \forall \omega \in \Omega.$$

Essa variável Z tem duas propriedades muito importantes, que permitirão que ela seja a derivada de Radon-Nikodym necessária para obtermos a medida \mathbb{Q} equivalente a \mathbb{P} :

(i) $Z(\omega) > 0 \forall \omega \in \Omega$;

(ii) $\mathbb{E}Z = 1$.

De fato, (i) é imediato por definição e (ii) segue das propriedades da esperança:

$$\begin{aligned} \mathbb{E}Z &= \int_{-\infty}^{+\infty} \exp \left\{ -\xi x - \frac{1}{2}\xi^2 \right\} n(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{2}(x^2 - 2\xi x + \xi^2) \right\} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{2}(x + \xi)^2 \right\} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{2}y^2 \right\} dy = 1. \end{aligned}$$

Agora, usamos Z para criar uma nova medida de probabilidade $\mathbb{Q} \ll \mathbb{P}$ ao definirmos:

$$\mathbb{Q}(A) = \int_A Z(\omega) d\mathbb{P}(\omega), \forall A \in \mathcal{F}. \quad (1.27)$$

Além disso, Y tem distribuição normal padrão sob \mathbb{Q} . De fato,

$$\begin{aligned} \mathbb{Q}(Y \leq b) &= \int_{\{\omega: Y(\omega) \leq b\}} Z(\omega) d\mathbb{P}(\omega) \\ &= \int_{\Omega} \mathbb{1}_{\{Y(\omega) \leq b\}} Z(\omega) d\mathbb{P}(\omega) \\ &= \int_{\Omega} \mathbb{1}_{\{Y(\omega) \leq b\}} \exp \left\{ -\xi X(\omega) - \frac{1}{2}\xi^2 \right\} d\mathbb{P}(\omega) \\ &= \int_{-\infty}^{+\infty} \mathbb{1}_{\{x \leq b - \xi\}} e^{-\xi x - \frac{1}{2}\xi^2} n(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b - \xi} e^{-\xi x - \frac{1}{2}\xi^2} e^{-\frac{1}{2}x^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b - \xi} e^{-\frac{1}{2}(x + \xi)^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{1}{2}y^2} dy, \end{aligned}$$

como queríamos. □

O exemplo anterior nos mostra como podemos utilizar essa ideia de *mover a média* de uma variável aleatória normal. Essa ideia pode ser aplicada em um contexto mais geral, onde mudamos a média não de uma variável aleatória, mas sim de um processo estocástico inteiro.

Isso nos leva ao resultado a seguir.

Teorema 1.27 (Girsanov, unidimensional). *Sejam W_t um movimento browniano unidimensional definido sob $(\Omega, \mathcal{F}_t, \mathbb{P})$, $t \in [0, T]$, e \mathcal{F}_t sua filtração associada. Considere Ξ_t um processo estocástico adaptado e defina:*

$$Z_t = \exp \left\{ - \int_0^t \Xi_s dW_s - \frac{1}{2} \int_0^t \Xi_s^2 ds \right\}, \quad (1.28)$$

$$W_t^* = W_t + \int_0^t \Xi_s ds, \quad (1.29)$$

e suponha que

$$\mathbb{E} \int_0^T \Xi_s^2 Z_s^2 ds < \infty. \quad (1.30)$$

Por fim, faça $Z = Z(T)$. Nessas condições, $\mathbb{E}Z = 1$ e, sob a medida \mathbb{P}^* definida por

$$\mathbb{P}^*(A) = \int_A Z(w) d\mathbb{P}(w) \quad \forall A \in \mathcal{F}, \quad (1.31)$$

o processo W_t^* é um movimento browniano.

Exemplo 1.28. Transitando entre movimentos brownianos.

Vamos praticar a mudança de medida que vimos no teorema de Girsanov para o caso mais simples, onde $\Xi \equiv \xi$ constante.

Queremos mudar do movimento browniano W_t (sob \mathbb{P}), para outro, $W_t^* = W_t + \xi t$, definido sob outra medida \mathbb{P}^* , induzida por ele (pode-se provar que ele segue um movimento browniano padrão sob ela). Sendo assim, a densidade de W^* avaliada no ponto y é dada por

$$f_{W^*}(y) = \frac{1}{\sqrt{2\pi t}} \exp \left\{ - \frac{1}{2} \left(\frac{y}{\sqrt{t}} \right)^2 \right\},$$

e, da mesma forma, a densidade de W no ponto x é

$$f_W(x) = \frac{1}{\sqrt{2\pi t}} \exp \left\{ - \frac{1}{2} \left(\frac{x}{\sqrt{t}} \right)^2 \right\}.$$

De modo que a razão entre essas densidades no ponto x é (após substituirmos $y = x + \xi t$ e reorganizarmos a expressão):

$$\frac{f_{W^*}(y)}{f_W(x)} = \exp \left\{ - \xi x - \frac{1}{2} \xi^2 t \right\}$$

O que nos motiva a tomar:

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp \left\{ -\xi W_t - \frac{1}{2} \xi^2 t \right\} \text{ q.c } [\mathbb{P}].$$

Agora, quisermos voltar de W^* para W , basta procedermos de maneira similar para obtermos

$$\frac{d\mathbb{P}}{d\mathbb{P}^*} = \exp \left\{ \xi W_t - \frac{1}{2} \xi^2 t \right\} \text{ q.c } [\mathbb{P}^*].$$

Note que as duas mudanças de medida realmente se comportam como se fossem uma razão numérica pois, substituir $W_t = W_t^* - \xi t$ na primeira expressão, resulta:

$$\begin{aligned} \frac{d\mathbb{P}^*}{d\mathbb{P}} &= \exp \left\{ -\xi(W_t^* - \xi t) - \frac{1}{2} \xi^2 t \right\} \\ &= \exp \left\{ -\xi W_t^* + \frac{1}{2} \xi^2 t \right\} \end{aligned}$$

Ou seja,

$$\frac{1}{\frac{d\mathbb{P}}{d\mathbb{P}^*}} = \exp \left\{ \xi W_t - \frac{1}{2} \xi^2 t \right\} = \frac{d\mathbb{P}}{d\mathbb{P}^*} \text{ q.c } [\mathbb{P}^*].$$

Naturalmente, podemos também aplicar diretamente o teorema para obtermos:

$$Z_t = \exp \left\{ -\int_0^t \xi dW_s - \frac{1}{2} \int_0^t \xi^2 ds \right\} = \exp \left\{ -\xi W_t - \frac{1}{2} \xi^2 t \right\}$$

Daí, por Girsanov, concluímos que $W^* = W_t + \xi t$ é um movimento browniano e a derivada de Radon-Nikodym é dada por

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = Z_t = \exp \left\{ -\xi W_t - \frac{1}{2} \xi^2 t \right\}.$$

□

De maneira geral, temos o seguinte resultado, que une os teoremas de Girsanov e da Representação Martingale:

Teorema 1.29 (Girsanov). *Suponha que $\Xi(t) \in \mathcal{V}^m(0, T)$ satisfaça a seguinte condição de Novikov:*

$$\mathbb{E} \exp \left(\frac{1}{2} \int_0^T \Xi^2(s, w) ds \right) < \infty \quad (1.32)$$

e considere \mathbb{Q} uma medida em $\mathcal{F}_T^{(m)}$ tal que

$$d\mathbb{Q}(w) = \exp \left(-\int_0^T \Xi(t, w) dW(t) - \frac{1}{2} \int_0^T \Xi^2(t, w) dt \right) d\mathbb{P}(w). \quad (1.33)$$

Então,

$$W^*(t) := \int_0^t \Xi(s, w) ds + W(t) \quad (1.34)$$

é uma martingale com respeito à $\mathcal{F}_t^{(m)}$ (e, portanto, um movimento browniano em $\mathcal{F}_t^{(m)}$) sob \mathbb{Q} . Além disso, qualquer $F \in L^2(\mathcal{F}_t^{(m)}, \mathbb{Q})$ possui uma única representação

$$F(w) = \mathbb{E}_{\mathbb{Q}}F + \int_0^T \phi(t, w) d\tilde{W}(t), \quad (1.35)$$

onde ϕ é um processo a valores em \mathbb{R}^m que é $\mathcal{F}_t^{(m)}$ -adaptado, (t, w) -mensurável e quadrado integrável em $[0, T]$ com respeito à \mathbb{Q} .

1.3 Finanças Matemáticas

1.3.1 Mercados, carteiras e arbitragem

Definição 1.30. Um **mercado financeiro** (ou, simplesmente, mercado) é um processo de Itô $(n + 1)$ -dimensional e $\mathcal{F}_t^{(m)}$ -adaptado $X(t) = (B(t), X_1(t), \dots, X_n(t))$, $t \in [0, T]$, que possui a forma

$$\begin{cases} dB(t) = r(t)B(t)dt; B(0) = 1 \\ dX_i(t) = \mu_i(t)dt + \sum_{j=1}^m \sigma_{ij}(t)dW_j \end{cases} \quad (1.36)$$

$\{X(t)\}_{t \in [0, T]}$ é chamado **normalizado** se $B(t) \equiv 1$.

Definição 1.31. Uma **carteira** (ou portfólio) no mercado $\{X(t)\}_{t \in [0, T]}$ é um processo estocástico $(n + 1)$ -dimensional $\mathcal{F}_t^{(m)}$ -adaptado e (t, w) -mensurável

$$\theta(t) = (\theta_0(t), \theta_1(t), \dots, \theta_n(t)), \quad t \in [0, T]. \quad (1.37)$$

O **valor** (ou processo de riqueza), no tempo t , de uma carteira $\theta(t)$ é definido por

$$V^\theta(t) = \sum_{i=0}^n \theta_i(t)X_i, \quad (1.38)$$

onde $X_0(t) = B(t)$.

Definição 1.32. O portfólio $\theta(t)$ é chamado **auto-financiado** se satisfaz

$$V(t) = V(0) + \sum_{i=1}^n \int_0^t \theta_i(s) dX_i(s), \quad (1.39)$$

ou, em notação diferencial,

$$dV(t) = \sum_{i=1}^n \theta_i(t) dX_i(t), \quad V(0) = V_0. \quad (1.40)$$

Observação 1.4. Alguns comentários.

- (a) Pensamos em $X_i(t) = X_i(t, w)$ como o preço do i -ésimo ativo da carteira no tempo t . Os ativos $1, \dots, n$ são chamados **arriscados** devido a presença dos seus termos de difusão. Eles podem representar, por exemplo, ações. Já o ativo $B(t)$ é chamado **livre de risco** por não possuir o termo de difusão. Entre outras coisas, representa títulos da dívida pública.
- (b) Podemos sempre tornar o mercado normalizado, basta definirmos

$$\bar{X}_i(t) = B(t)^{-1}X_i(t); 1 \leq i \leq n \quad (1.41)$$

O mercado $\bar{X}(t) = (1, \bar{X}_1(t), \dots, \bar{X}_n(t))$ é chamado a **normalização** de $X(t)$. Essa normalização corresponde a usar o $X_0(t)$ como *numeraire*, isto é, como o ativo de referência usado para expressar os preços dos demais ativos.

Como

$$B(t) = \exp \left\{ \int_0^t r(s) ds \right\}, \quad (1.42)$$

temos:

$$B^{-1}(t) = \exp \left\{ - \int_0^t r(s) ds \right\} > 0, \forall t \in [0, T] \quad (1.43)$$

e também

$$d\bar{X}_i(t) = d(B^{-1}(t)X(t)) = B^{-1}(t)[(\mu_i - rX_i)dt + \sigma_i dW(t)]; 1 \leq i \leq n. \quad (1.44)$$

- (c) Os componentes $\theta_i(t)$ representam o número de cotas dos títulos que um investidor detém no tempo t . Ou seja, representam o valor total de todos os investimentos em t .

Observação 1.5. Se $\theta_1, \dots, \theta_n$ são dados, podemos **sempre** fazer de $\theta = (\theta_0, \theta_1, \dots, \theta_n)$ uma carteira auto-financiada ao escolhermos θ_0 como sendo

$$\theta_0(t) = V^\theta(0) + B^{-1}(t)A(t) + \int_0^t \rho(s)A(s)B^{-1}(s)ds, \quad (1.45)$$

onde

$$A(t) = \sum_{i=1}^n \left(\int_0^t \theta_i(s) dX_i(s) - \theta_i(t)X_i(t) \right) \quad (1.46)$$

Além disso, podemos escolher o valor inicial $V^\theta(0)$ da carteira.

Definição 1.33. Um portfólio auto-financiado $\theta(t)$ é chamado **admissível** se seu processo de valor correspondente, $V^\theta(t)$, é limitado inferiormente, isto é, se existe $K = K(\theta) < \infty$ tal que

$$V^\theta(t, w) \geq -K \text{ para q.t. } (t, w) \in [0, T] \times \Omega. \quad (1.47)$$

Essa condição é natural se pensarmos que deve existir um limite na quantidade de débito que os credores podem tolerar.

Definição 1.34. Um portfólio admissível $\theta(t)$ é chamado uma **arbitragem** no mercado $\{X_t\}_{t \in [0, T]}$ se seu processo de valor $V^\theta(t)$ satisfaz

$$\begin{cases} V^\theta(0) = 0 \\ V^\theta(T) \geq 0 \text{ q.c., com } P(V^\theta(T) > 0) > 0 \end{cases} \quad (1.48)$$

Em outras palavras, $\theta(t)$ é uma arbitragem se, de alguma forma, $\theta(t)$ gera lucro sem qualquer risco de perder dinheiro. Isto é, com probabilidade positiva, há um aumento estritamente positivo no valor da carteira entre 0 e T .

Intuitivamente, a existência de oportunidades de arbitragem reflete uma falta de equilíbrio no mercado. Logo, é importante saber determinar se um dado mercado admite arbitragem ou não. Essa possibilidade está intimamente relacionada com as condições que impomos às carteiras que são admissíveis. Se exigirmos apenas que nossos portfólios sejam auto-financiados, é possível encontrar exemplos de estratégias que permitam arbitragem. Assim, precisamos de mais condições.

Uma maneira simples de identificar quando um mercado $\{X_t\}_{t \in [0, T]}$ possui arbitragem ou não é dada pelo seguinte resultado:

Lema 1.35. *Suponha que exista uma medida $\mathbb{Q} \sim \mathbb{P}$ em $\mathcal{F}_T^{(m)}$ tal que o processo normalizado $\{\bar{X}_t\}_{t \in [0, T]}$ seja uma martingale local com respeito à medida \mathbb{Q} . Então, $\{X_t\}_{t \in [0, T]}$ não possui arbitragem.*

Demonstração. Øksendal (2003). □

Definição 1.36. Uma medida $\mathbb{Q} \sim \mathbb{P}$ tal que o processo normalizado $\{\bar{X}_t\}_{t \in [0, T]}$ é uma martingale (local) com respeito à \mathbb{Q} é chamada uma **medida martingale equivalente** (local).

Assim, o lema anterior afirma que se existe uma medida martingale equivalente, então o mercado não possui arbitragem. Vale ainda a recíproca, de modo que:

um mercado não admite arbitragem se, e somente se, nele existe uma medida martingale equivalente.

Esse resultado é conhecido na literatura como o **Primeiro Teorema Fundamental da Precificação de Ativos**.

1.3.2 Replicabilidade e Completude

Começamos com o simples, mas importante, resultado

Lema 1.37. *Seja $\bar{X}(t) = B^{-1}(t)X(t)$ o processo de preço normalizado. O portfólio $\theta(t)$ é admissível para o mercado $\{X_t\}$, com valor*

$$V^\theta(t) = \langle \theta(t), X(t) \rangle \quad (1.49)$$

se, e somente se, $\theta(t)$ também é admissível para o mercado normalizado, com valor

$$\bar{V}^\theta(t) = \langle \theta(t), \bar{X}(t) \rangle = B^{-1}(t)V^\theta(t). \quad (1.50)$$

Demonstração. Øksendal (2003). □

Definição 1.38. Um **derivativo europeu**, ou simplesmente um derivativo, é uma variável aleatória $F(w) \in L^2(\mathbb{Q})$, limitada inferiormente e $\mathcal{F}_T^{(m)}$ -mensurável. Dizemos que $F(w)$ é **replicável** (no mercado $\{X_t\}_{t \in [0, T]}$) se existe um portfólio admissível $\theta(t)$ e $z \in \mathbb{R}$ tais que:

$$F(w) = V_z^\theta(T) := z + \int_0^T \theta(t) dX(t) \text{ q.c.}[\mathbb{Q}] \quad (1.51)$$

de modo que $\bar{V}_z^\theta(t)$ seja uma \mathbb{Q} -martingale. Se tal $\theta(t)$ existe, o denominamos **portfólio replicante**, ou estratégia replicante, para F .

Definição 1.39. O mercado $\{X_t\}_{t \in [0, T]}$ é chamado **completo** se todo derivativo europeu com vencimento T for replicável.

Em outras palavras, $F(w)$ é replicável se existe um número real z tal que, se z é nossa fortuna inicial, podemos encontrar uma estratégia admissível $\theta(t)$ que gera um valor V_z^θ tal que

$$V_z^\theta(T) = F(w) \text{ para quase todo } w. \quad (1.52)$$

Observação 1.6. Mais ainda, existe a seguinte caracterização para a completude, conhecida como o **Segundo Teorema Fundamental da Precificação de Ativos**:

Um mercado $\{X_t\}$ é completo se, e somente se, existe uma única medida martingale equivalente para o mercado normalizado $\{\bar{X}_t\}$.

1.3.3 Precificação

Definição 1.40. Seja $F(w)$ um título derivativo europeu com vencimento em T . Uma **opção europeia** no título F é uma garantia de pagamento da quantidade $F(w)$ no tempo $t = T > 0$.

Definição 1.41. Sejam F um derivativo com vencimento T , B_t um ativo livre de risco e \mathbb{Q} uma medida martingale equivalente. Definimos o **preço de não arbitragem** de F no instante $t \in [0, T]$ por

$$\pi_t(F) = B_t \mathbb{E}_{\mathbb{Q}}[B_T^{-1} F | \mathcal{F}_t]. \quad (1.53)$$

Observação 1.7. Sobre a precificação neutra ao risco.

Uma das noções mais importantes da teoria de finanças matemáticas é a de precificação neutra ao risco. Ela surge como uma espécie de simplificação do mundo real: supõe-se um mundo onde os investidores são indiferentes aos diversos tipos de risco envolvidos nas transações financeiras.

Como consequência, o retorno esperado de qualquer ativo é exatamente a taxa livre de risco r . Esse é o princípio que motiva a equação (1.53). A maior vantagem nessa abordagem é que ela elimina a taxa de retorno real μ do ativo, que é um parâmetro subjetivo e notoriamente difícil de estimar.

Essa filosofia, contudo, não é apenas um artifício matemático e teórico. Na prática, os preços praticados no mercado são baseados na medida neutra ao risco, e procedimentos usuais como a calibração de modelos, simulação de Monte Carlo para precificação, gestão de risco e modelagem da curva de juros também são fundamentados nessa estrutura.

Observação 1.8. A paridade *put-call*.

Em mercados livres de arbitragem, existe uma famosa e importante expressão que relaciona os preços de opções europeias de compra e de venda, conhecida como **paridade *put-call***. Ela decorre das propriedades do *payoff* e é dada por:

$$C = P + S - Ke^{-r\tau}, \quad (1.54)$$

onde C representa o preço da *call*, P o preço da *put*, S é o *spot*, K o *strike*, r a taxa de juros livre de risco e τ a maturidade comum das opções.

O modelo de Black-Scholes (BS)

Em seu artigo de 1973, Fischer Black e Myron Scholes fizeram uma série de suposições para calcular o valor teórico do valor de uma opção em termos do preço da ação. Apesar de essas condições serem ideais e não refletirem a realidade do mercado como um todo, elas possibilitaram que eles obtivessem uma expressão fechada para esse valor teórico, o que tornou o modelo proposto por eles uma espécie de *benchmarking* para o mercado, ainda hoje utilizado no mundo todo. Além disso,

As condições assumidas foram:

- (a) A taxa de juros r é conhecida e constante ao longo do tempo.
- (b) O preço da ação segue um passeio aleatório a tempo contínuo, com variância proporcional ao quadrado do preço da ação. Logo, a distribuição dos possíveis preços da ação ao final de qualquer intervalo finito é log-normal. A variância do retorno da ação é constante.
- (c) A ação não paga dividendos ou outras distribuições.
- (d) A opção é europeia, isto é, só pode ser exercida no vencimento.
- (e) Não existem custos de transação ao comprar ou vender a ação ou a opção.
- (f) É possível financiar qualquer fração do preço de um ativo, seja para compra ou manutenção, usando a taxa de juros de curto prazo.
- (g) Não há penalidades para a venda a descoberto. Um vendedor que não possui um ativo simplesmente aceitará o preço do ativo de um comprador e concordará em liquidar a operação com o comprador em uma data futura, pagando-lhe um valor igual ao preço do ativo naquela data.

Sob essas hipóteses, o valor da opção dependerá apenas do preço da ação e do tempo, além das variáveis que são assumidas conhecidas e constantes.

Assim, utilizando os conceitos apresentados em 1.3, temos um mercado com uma conta poupança B_t , pagando a uma taxa de juros constante r , e uma ação cuja dinâmica do preço S_t obedece a equação de um movimento browniano geométrico:

$$\begin{cases} dB_t = rB_t dt; B_0 = 1 \\ dS_t = \mu dt + \sigma S_t dW_t \end{cases} \quad (1.55)$$

A etapa crucial do desenvolvimento deste modelo é a de construir uma carteira $\theta = (\theta_0, \theta_1)$ que replique o valor da opção, isto é,

$$V_T^\theta(w) = F(w) \text{ quase certamente.} \quad (1.56)$$

Vejamos, agora, como esse modelo de encaixa na teoria que estudamos até o momento.

Exemplo 1.42. De \mathbb{P} para \mathbb{Q} no modelo de Black-Scholes.

Como vimos, sob a medida física \mathbb{P} , o modelo de BS supõe que o ativo adjacente possui dinâmica

$$dS_t = \mu S_t dt + \sigma S_t dW_t^{\mathbb{P}}, \quad \mu \text{ e } \sigma \text{ constantes.}$$

Vimos também que \mathbb{Q} é a medida que torna o preço descontado $\bar{S}_t = S_t/B_t$ uma martingale. Vamos relacionar esses fatos através da mudança de medida.

Aplicando a fórmula de Itô na função \bar{S} , obtemos:

$$d\bar{S} = d\left(\frac{S}{B}\right) = \frac{1}{B}dS - \frac{S}{B^2}dB + \frac{1}{B^3}(dB)^2 - \frac{1}{B^2}dSdB.$$

De $dB_t = rB_t dt$ e da tabela 1.17, segue que:

$$d\left(\frac{S}{B}\right) = \frac{1}{B}dS - \frac{S}{B^2}dB.$$

Substituindo a dinâmica de S_t e B_t resulta:

$$\begin{aligned} d\left(\frac{S_t}{B_t}\right) &= \frac{1}{B_t}(\mu S_t dt + \sigma S_t dW_t^{\mathbb{P}}) - \frac{S_t}{B_t^2}(rB_t dt) \\ &= (\mu - r)\frac{S_t}{B_t}dt + \sigma\frac{S_t}{B_t}dW_t^{\mathbb{P}}. \end{aligned} \quad (1.57)$$

Lembrando que S/B é uma martingale, que integrais de Itô também o são quando não contém qualquer termo de *drift* e que, como vimos em 1.2.4, a volatilidade não é afetada por uma mudança de medida, concluímos que S/B , na medida \mathbb{Q} , deve possuir a seguinte dinâmica:

$$d\left(\frac{S_t}{B_t}\right) = \sigma\frac{S_t}{B_t}dW_t^{\mathbb{Q}}. \quad (1.58)$$

Igualando as últimas duas equações, encontramos:

$$dW_t^{\mathbb{Q}} = \frac{\mu - r}{\sigma} dt + dW_t^{\mathbb{P}}. \quad (1.59)$$

A equação (1.59) determina a transformação de Girsanov da medida real \mathbb{P} para a medida risco-neutra \mathbb{Q} , sob o modelo de Black-Scholes. Assim sendo, temos

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t^{\mathbb{P}} \\ &= \mu S_t dt + \sigma S_t \left(dW_t^{\mathbb{Q}} - \frac{\mu - r}{\sigma} dt \right) \\ &= r S_t dt + \sigma S_t dW_t^{\mathbb{Q}}, \end{aligned} \quad (1.60)$$

que é a equação da dinâmica de S sob \mathbb{Q} . □

Corolário 1.43. *Sob a medida neutra ao risco \mathbb{Q} , um ativo S_t gerado por (1.60) possui retorno esperado dado por*

$$\mathbb{E}_{\mathbb{Q}}[S_t] = S_0 e^{rt}. \quad (1.61)$$

Demonstração. Com efeito, de (1.20) segue que

$$S_t = S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma W_t^{\mathbb{Q}} \right\}.$$

Tomando a esperança (com respeito à \mathbb{Q}), obtemos:

$$\begin{aligned} \mathbb{E}[S_t] &= \mathbb{E} \left[S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma W_t^{\mathbb{Q}} \right\} \right] \\ &= S_0 \exp \left[\left(r - \frac{\sigma^2}{2} \right) t \right] \mathbb{E} \left[\exp \left(\sigma W_t^{\mathbb{Q}} \right) \right]. \end{aligned}$$

Agora, perceba que esperança na última igualdade acima é, por definição, uma função geradora de momentos e, sendo $W_t^{\mathbb{Q}} \sim N(0, t)$, concluímos que:

$$\mathbb{E} \left[\exp \left(\sigma W_t^{\mathbb{Q}} \right) \right] = \exp \left(\frac{1}{2} \sigma^2 t \right).$$

Substituindo de volta:

$$\mathbb{E}[S_t] = S_0 \exp \left[\left(r - \frac{\sigma^2}{2} \right) t \right] \exp \left(\frac{1}{2} \sigma^2 t \right).$$

$$\therefore \mathbb{E}_{\mathbb{Q}}[S_t] = S_0 e^{rt}.$$

□

Procedemos aplicando nosso arcabouço teórico num problema um pouco mais concreto.

Exemplo 1.44. O preço de uma *call* europeia no modelo Black-Scholes.

Vamos encontrar o preço de não arbitragem para uma opção de compra europeia, isto é, $F(w) = (S_T - K)^+$. Assim,

$$\pi_t(F) = \mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)}(S_T - K)^+ | \mathcal{F}_t].$$

Primeiramente, note que: como o incremento $W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}$ é independente de \mathcal{F}_t e o preço da ação S_t é \mathcal{F}_t -mensurável, segue, das propriedades de esperança condicional, que:

$$\mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)}(S_T - K)^+ | \mathcal{F}_t] = H(S_t, T - t), \quad (1.62)$$

onde a função $H(S_t, T - t)$ é definida por:

$$H(s, T - t) = \mathbb{E}_{\mathbb{Q}} \left\{ (s e^{\sigma(W_T^{\mathbb{Q}} - W_t^{\mathbb{Q}}) + (r - \frac{1}{2}\sigma^2)(T-t)} - K)^+ \right\},$$

para qualquer $(s, t) \in \mathbb{R}_+ \times [0, T]$.

Logo, basta fazermos para o caso $t = 0$ e encontrarmos a esperança incondicional

$$\mathbb{E}_{\mathbb{Q}}(B_T^{-1}(S_T - K)^+) = \mathbb{E}_{\mathbb{Q}}(S_T B_T^{-1} \mathbf{1}_D) - \mathbb{E}_{\mathbb{Q}}(K B_T^{-1} \mathbf{1}_D) = J_1 - J_2,$$

onde $\mathbf{1}_D$ é a indicadora do conjunto $D = \{S_T > K\}$.

Para J_2 , temos:

$$\begin{aligned} J_2 &= e^{-rT} K \mathbb{Q}(S_T > K) \\ &= e^{-rT} K \mathbb{Q}(S_0 \exp(\sigma W_T^{\mathbb{Q}} + (r - \frac{1}{2}\sigma^2)T) > K) \\ &= e^{-rT} K \mathbb{Q}\{-\sigma W_T^{\mathbb{Q}} < \ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T\} \\ &= e^{-rT} K \mathbb{Q}\left\{\xi < \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right\} \\ &= e^{-rT} K N(d_2(S_0, T)), \end{aligned}$$

já que $\xi = -W_T^{\mathbb{Q}}/\sqrt{T} \sim N(0, 1)$ sob a medida martingale \mathbb{Q} .

Para a outra integral, primeiramente note que:

$$J_1 = \mathbb{E}_{\mathbb{Q}}(S_T B_T^{-1} \mathbf{1}_D) = \mathbb{E}_{\mathbb{Q}}(\bar{S}_T \mathbf{1}_D). \quad (1.63)$$

Agora, para o cálculo de J_1 , é conveniente introduzir uma nova medida \mathbb{P}^* em (Ω, \mathcal{F}_T) , a partir de \mathbb{Q} , fazendo

$$\frac{d\mathbb{P}^*}{d\mathbb{Q}} = \exp\left(\sigma W_T^{\mathbb{Q}} - \frac{1}{2}\sigma^2 T\right) \text{ q.c. } [\mathbb{Q}].$$

Assim, pelo teorema de Girsanov, o processo $W_t^{\mathbb{P}^*} = W_t^{\mathbb{Q}} - \sigma t$ segue um movimento browniano padrão no espaço $(\Omega, \mathbb{F}, \mathbb{P}^*)$. Mais ainda, usando a solução de (1.58), obtemos:

$$\bar{S}_T = S_0 \exp\left(\sigma W_T^{\mathbb{P}^*} + \frac{1}{2}\sigma^2 T\right). \quad (1.64)$$

Daí,

$$\begin{aligned}
J_1 &= \mathbb{E}_Q(\bar{S}_T \mathbf{1}_D) \\
&= S_0 \mathbb{P}^*(\bar{S}_T > KB_T^{-1}) \\
&= S_0 \mathbb{P}^* \left[\exp \left(\sigma W_T^{\mathbb{P}^*} + \frac{1}{2} \sigma^2 T \right) > Ke^{-rT} \right] \\
&= S_0 \mathbb{P}^* \left(-\sigma W_T^{\mathbb{P}^*} < \ln(S_0/K) + \frac{1}{2} \sigma^2 T \right).
\end{aligned}$$

De maneira semelhante ao que fizemos para J_2 , concluímos que

$$J_1 = S_0 N(d_1(S_0, T)).$$

Resumindo, mostramos que, para $t = 0$, o preço de uma call europeia é dado por

$$C_0 = c(S_0, T) = S_0 N(d_1(S_0, T)) - Ke^{-rT} N(d_2(S_0, T)), \quad (1.65)$$

onde

$$d_{1,2}(S_0, T) = \frac{\ln(S_0/K) + (r \pm \frac{1}{2} \sigma^2) T}{\sigma \sqrt{T}}. \quad (1.66)$$

O procedimento para demonstrar para $0 < t < T$ é essencialmente o mesmo, basta substituir S_0 por S_t , T por $\tau = T - t$ e assim por diante. Isso nos fornece o preço de uma *call* europeia no modelo de Black-Scholes no tempo t :

$$C_t = S_t N(d_1(S_t, \tau)) - Ke^{-r\tau} N(d_2(S_t, \tau)), \quad (1.67)$$

onde

$$d_{1,2}(S_t, \tau) = \frac{\ln(S_t/K) + (r \pm \frac{1}{2} \sigma^2) \tau}{\sigma \sqrt{\tau}} \quad (1.68)$$

e $N(x)$ é a função de distribuição normal padrão. \square

Exemplo 1.45. As gregas.

Para medir quantitativamente a influência de variáveis do mercado (como *spot*, tempo até o vencimento, preço de exercício, volatilidade, etc.) no prêmio, utilizamos as chamadas **gregas**, que são, basicamente, derivadas parciais do prêmio em relação a alguma dessas variáveis. Em particular, o delta e o gamma correspondem às derivadas de primeira e segunda ordem em relação ao preço do ativo subjacente S , enquanto o theta, o vega e o rho são obtidos derivando-se em relação a τ , σ e r , respectivamente.

Elas recebem esse nome por, em sua maioria, serem representadas por letras gregas no modelo Black-Scholes.

Para um vencimento T fixo e $t \leq T$ arbitrário, considere $\tau = T - t$. Denote por $c(S_t, \tau, K, r, \sigma)$ e $p(S_t, \tau, K, r, \sigma)$ o preço de uma *call* e uma *put* europeias, respectivamente.

As funções c e p são dadas por:

$$c(s, \tau, K, r, \sigma) = sN(d_1) - Ke^{-r\tau} N(d_2) \quad (1.69)$$

e

$$p(s, \tau, K, r, \sigma) = Ke^{-r\tau}N(-d_2) - sN(-d_1), \quad (1.70)$$

onde

$$d_{1,2} = d_{1,2}(s, \tau, K, r, \sigma) = \frac{\ln(s/K) + (r \pm \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}. \quad (1.71)$$

Assim, derivando essas expressões, obtemos cada uma das gregas para os dois tipos de opções:

Grega	Opção	Expressão
Delta (Δ)	<i>Call</i>	$N(d_1)$
	<i>Put</i>	$N(d_1) - 1$
Gamma (Γ)	<i>Call & Put</i>	$\frac{N'(d_1)}{S\sigma\sqrt{\tau}}$
	<i>Call</i>	$-\frac{SN'(d_1)\sigma}{2\sqrt{\tau}} - rKe^{-r\tau}N(d_2)$
Theta (Θ)	<i>Put</i>	$-\frac{SN'(d_1)\sigma}{2\sqrt{\tau}} + rKe^{-r\tau}N(-d_2)$
	<i>Call & Put</i>	$S\sqrt{\tau}N'(d_1)$
Rho (ρ)	<i>Call</i>	$K\tau e^{-r\tau}N(d_2)$
	<i>Put</i>	$-K\tau e^{-r\tau}N(-d_2)$

Tabela 1.3: Gregas para opções europeias

Para ver como as gregas funcionam, vamos supor uma opção de compra europeia com os seguintes parâmetros: $S = 100$, $K = 100$, $\sigma = 0.25$ ao ano, $r = 0.06$ ao ano e $\tau = 2$ meses.

$$\Delta = 0,559$$

A cada R\$ 1 que a ação sobe, a opção sobe 55,9 centavos.

$$\Gamma = 0,039$$

A cada R\$ 1 que a ação sobe, o delta ganha 3,9 pontos percentuais.

$$\Theta = -0,0415$$

A cada dia corrido que passa, se nenhum outro parâmetro sofrer alteração, a opção desvaloriza 4,15 centavos.

$$\nu = 16,11$$

Se a volatilidade aumentar 1 ponto percentual, a opção valoriza 16,11 centavos.

$$\rho = 8,56$$

Se a taxa de juros aumentar 1 ponto percentual, a opção valoriza 8,56 centavos.

Capítulo 2

Volatilidade Implícita e o Modelo de Heston

Apesar da revolução criada no mercado financeiro pelo modelo de Black-Scholes, ele apresenta limitações. Algumas hipóteses cruciais, apesar de facilitarem a obtenção de uma solução analítica, mostraram-se questionáveis na prática. Por exemplo, no contexto de Black-Scholes, o *delta-hedging* (estratégia que visa neutralizar o risco envolvendo mudanças nos preços do ativo subjacente) é um processo contínuo, mas, na prática, é um processo discreto: o portfólio é rebalanceado mais ou menos uma vez por semana, dependendo do tipo de derivativo. Além disso, os custos de transações não são contabilizados (o que, definitivamente, não condiz com a realidade). Entretanto, talvez a hipótese mais irreal delas seja a de volatilidade constante.

A fórmula de Black-Scholes, por inversão numérica, nos fornece uma maneira de mapear univocamente cada preço de mercado de *calls* ou *puts* a uma volatilidade, o que chamamos de *volatilidade implícita*. Se o modelo descrevesse fielmente a realidade do mercado, a volatilidade implícita, quando calculada para diversos *strikes*, deveria ser sempre a mesma (já que, supostamente, ela é constante). Entretanto, isso não se verifica no mundo real.

Se essa hipótese fosse verdadeira, o gráfico dos *strikes* pela volatilidade implícita seria uma reta horizontal, com valor igual a σ . Entretanto, o que se observa no mercado é que a volatilidade implícita varia com o preço de exercício. Mais ainda, ela é uma função convexa do *strike*. Por essa característica, essa curva recebe o nome de *volatility smile*, como mostra a figura 2.1.

E esse é o maior problema com o modelo de Black-Scholes: não conseguir reproduzir os *smiles* observados no mercado. Assim sendo, é necessário utilizar modelos mais complexos, que enfraquecem a hipótese de volatilidade constante, assumindo que ela seja uma função do tempo ou até um processo aleatório.

Queremos mais: vamos supor que o quadrado da volatilidade (sua variância) seja governado por um processo de Itô. Mais especificamente, estudaremos o famoso modelo de Heston, que possui diversas propriedades interessantes e fórmulas "quase fechadas", o que explica ele ser tão amplamente estudado e utilizado.

Principais referências utilizadas ao longo deste capítulo: Oosterlee e Grzelak (2020), Gatheral (2006), Rouah (2013) e Saporito (2023).

2.1 Volatilidade implícita

Considere r a taxa de juros livre de risco, um vencimento T e preço de exercício K sobre o ativo S . Denote por C^{BS} o preço de uma call no modelo Black-Scholes e C^{mkt} o preço de mercado dessa call (similarmente para *puts*).

Definição 2.1. Definimos a **volatilidade implícita** (de Black-Scholes) como sendo o número σ_{imp} para o qual:

$$C^{BS}(t_0, S, K, T, \sigma_{imp}, r) = C^{mkt}(K, T), \tag{2.1}$$

onde $t_0 = 0$. Em outras palavras, σ_{imp} é a volatilidade - inserida como parâmetro na solução de Black-Scholes - que reproduz o preço de mercado da opção no tempo $t_0 = 0$.

Observação 2.1. É comum referir-nos à volatilidade implícita a partir da frase de Riccardo Rebonato: "é o número errado a ser colocado na fórmula errada para obter o preço certo".

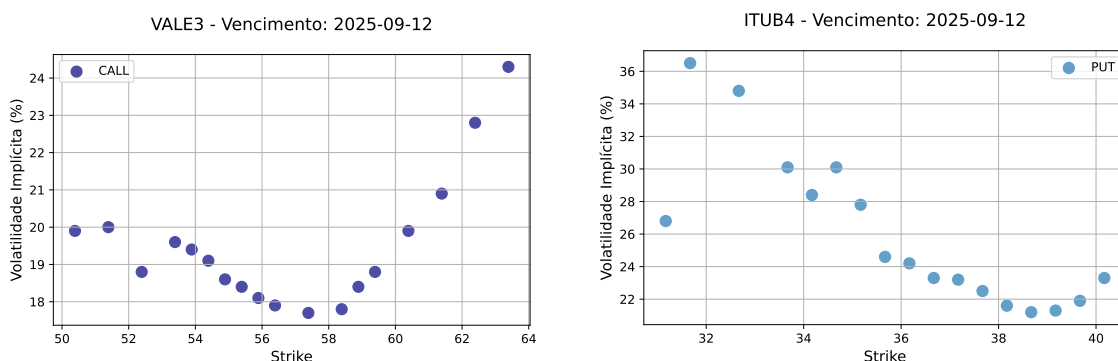


Figura 2.1: Gráficos das curvas *smile* de *calls* e *puts* para dados de mercado da VALE e ITAÚ, com o mesmo vencimento.

A importância da volatilidade implícita decorre do fato de o mercado, por vezes, quotar as opções europeias em termos de suas volatilidades implícitas, em vez dos preços de compra e venda da opção. Isso porque o preço monetário das opções depende fortemente da unidade do ativo subjacente e do *moneyness* dela, enquanto que a volatilidade implícita varia muito menos.

Em outras palavras, a volatilidade implícita pode ser vista como a *linguagem* na qual as opções são expressas.

2.1.1 Implementação computacional

Como não existe uma expressão fechada para a volatilidade implícita em termos dos preços da opção, para calculá-la precisamos recorrer a métodos numéricos. A equação (2.1) pode ser reformulada no seguinte problema de cálculo de raízes:

$$g(\sigma_{imp}) := C^{mkt}(K, T) - C^{BS}(t_0, S, K, T, \sigma_{imp}, r) = 0. \quad (2.2)$$

Assim, podemos utilizar vários métodos existentes para resolver esse problema, como o da bisseção ou o de Brent. Entretanto, um método bastante utilizado para essa tarefa é o de *Newton-Raphson*, devido à sua taxa de convergência quadrática.

Funciona assim: dado um chute inicial para a volatilidade implícita, $\sigma_{imp}^{(0)}$, e a derivada de $g(\sigma_{imp})$ com respeito a σ_{imp} , podemos encontrar as próximas aproximações, $\sigma_{imp}^{(k)}$, $k = 1, 2, \dots$, utilizando a recursão

$$\sigma_{imp}^{(k+1)} := \sigma_{imp}^{(k)} - \frac{g(\sigma_{imp}^{(k)})}{g'(\sigma_{imp}^{(k)})}, \text{ para } k \geq 0. \quad (2.3)$$

Note que, no contexto de Black-Scholes, os preços das opções europeias, bem como de suas derivadas, podem ser obtidos analiticamente, donde segue que:

$$g'(\sigma) = -\frac{\partial C}{\partial \sigma}(t_0, S_0, K, T, \sigma, r) = -Ke^{-r(T-t_0)}N'(d_2)\sqrt{T-t_0}, \quad (2.4)$$

isto é, o *vega* $g'(\sigma)$ de Black-Scholes. Entretanto, em algum caso mais complicado, onde uma expressão analítica para a derivada não for conhecida, aproximações adicionais são necessárias para aplicarmos o método.

Exemplo 2.2. A volatilidade implícita de Black-Scholes.

Para ilustrar o uso do método de Newton-Raphson no cálculo da volatilidade implícita, vamos considerar uma opção de compra com as seguintes características: preço atual da ação $S_0 = 100$, taxa de juros $r = 0.05$ a.a., maturidade em 1 ano, preço de exercício $K = 120$.

Vamos utilizar um chute inicial $\sigma_{imp}^{(0)} = 0.1$ e precisamos resolver o seguinte problema de cálculo de raízes:

$$g(\sigma_{imp}) = 2 - C^{BS}(0, 100, 120, 1, \sigma_{imp}, 0.05) = 0.$$

O programa nos retorna o valor $\sigma_{imp} \approx 0.16148272884139$. Se, agora, utilizarmos esse valor como parâmetro em C^{BS} , obteremos um valor para *call* de aproximadamente 2. \square

Ao observarmos os dados do mercado, fica claro que a volatilidade implícita depende fortemente não só do preço de exercício, mas também da sua maturidade. Isso nos leva a considerar o seguinte conjunto:

Definição 2.3. Seja $T^* > 0$ fixo e considere a família de preços $C_t = \{C^{mkt}(K, T) : K > 0, T \in (t, T^*]\}$. A **superfície de volatilidade implícita** (SVI) para esses preços é o conjunto

$$\hat{\sigma}_t = \{(T, K, \sigma_{imp}) : K > 0, T \in (t, T^*]\}. \quad (2.5)$$

É difícil sobrestimar a importância da superfície de volatilidade implícita, uma vez que ela é a representação mais completa de como o mercado, em consenso, precifica o risco. Ela contém toda a informação que nos interessa sobre preços de opções europeias.

Sua estrutura ilustra o comportamento presente e futuro dos dados daquele ativo: sua inclinação mostra como o mercado precifica o risco de queda e o potencial de alta (*puts* e *calls* fora do dinheiro, respectivamente); sua forma ao longo do vencimento (chamada **estrutura a termo**) mostra a expectativa do mercado sobre a volatilidade futura.

Além disso, a SVI é o retrato que os modelos de volatilidade tentam reproduzir: já que o Black-Scholes falha, o objetivo dos modelos que surgiram depois é tentar se ajustar a essa superfície através do processo de calibração. Um modelo bem sucedido pode, então, ser usado para precificar derivativos exóticos, por exemplo, ou realizar a gestão de risco de forma consistente com os dados observados no mercado.

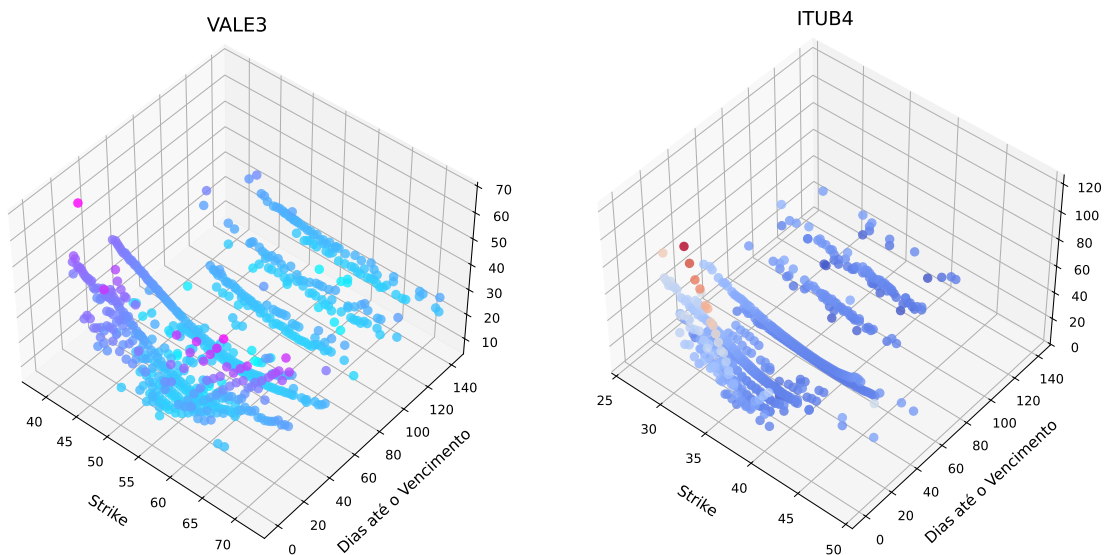


Figura 2.2: Superfícies de volatilidade implícita (SVI) para os mesmos dados de mercado da figura 2.1

2.2 Preliminares

2.2.1 A decomposição de Cholesky

Definição 2.4. Um processo estocástico $\{X_t\}_{t \geq 0}$ é dito ser um **processo gaussiano** se, para quaisquer instantes t_1, \dots, t_n , o vetor $(X_{t_1}, \dots, X_{t_n})$ tem distribuição normal multivariada.

Primeiramente, é importante notar que os incrementos do movimento browniano são um processo gaussiano. Portanto, podemos definir:

Definição 2.5. Dois movimentos brownianos \tilde{W}_i, \tilde{W}_j são ditos **independentes** se

$$\text{cov}[\tilde{W}_i(t), \tilde{W}_j(t)] = \mathbb{E}[\tilde{W}_i(t)\tilde{W}_j(t)] = \begin{cases} 0, & i \neq j, \\ t, & i = j, \end{cases} \quad (2.6)$$

para $i, j = 1, \dots, n$.

Observação 2.2. Independência em processos gaussianos.

A propriedade gaussiana do processo é essencial na definição acima. Isso porque, de forma geral, duas variáveis aleatórias serem não correlacionadas **não** implica que elas sejam independentes. Porém, no caso das distribuições normais multivariadas, essa implicação é sim verdadeira. Mais detalhes em Ash e Doléans-Dade (2000).

Agora, um problema de motivação.

Exemplo 2.6. Construindo movimentos brownianos correlacionados.

Sejam W_1, W_2 movimentos brownianos independentes e $-1 \leq \rho \leq 1$ um número dado. Para $0 \leq t \leq T$, defina o seguinte processo:

$$Z(t) := \rho W_1(t) + \sqrt{1 - \rho^2} W_2(t).$$

Em cada t , Z é uma combinação linear de normais independentes, de modo que também é normalmente distribuído, além de herdar a independência dos incrementos. Mais ainda, Z é um movimento browniano.

De fato, é imediato que $Z(0) = 0$. Além disso, a continuidade das trajetórias de Z segue diretamente da continuidade de W_1 e W_2 .

Agora,

$$\mathbb{E}Z(t) = \mathbb{E}[\rho W_1(t) + \sqrt{1 - \rho^2} W_2(t)] = 0$$

e também

$$\begin{aligned} \text{var } Z(t) &= \text{var}[\rho W_1(t) + \sqrt{1 - \rho^2} W_2(t)] \\ &= \text{var}[\rho W_1(t)] + \text{var}[\sqrt{1 - \rho^2} W_2(t)], \text{ pela independência} \\ &= \rho^2 t + (1 - \rho^2)t, \text{ pelas propriedades da variância} \\ &= t. \end{aligned}$$

Por fim, considere

$$\begin{aligned} Z(t + u) - Z(t) &= [\rho W_1(t + u) + \sqrt{1 - \rho^2} W_2(t + u)] - [\rho W_1(t) + \sqrt{1 - \rho^2} W_2(t)] \\ &= \rho(W_1(t + u) - W_1(t)) + \sqrt{1 - \rho^2}(W_2(t + u) - W_2(t)). \end{aligned}$$

Assim,

$$\begin{aligned} \text{var}[Z(t + u) - Z(t)] &= \text{var}[\rho(W_1(t + u) - W_1(t))] \\ &\quad + \text{var}[\sqrt{1 - \rho^2}(W_2(t + u) - W_2(t))] \\ &= \rho^2 u + (1 - \rho^2)u \\ &= u. \end{aligned}$$

Logo, vemos que a variância não depende de t . Portanto, Z é um movimento browniano.

Agora, vamos avaliar a correlação entre os processos Z e W_1 :

$$\begin{aligned}
 \text{corr}[Z(t), W_1(t)] &= \frac{\text{cov}[Z(t), W_1(t)]}{\sqrt{\text{var } Z(t)}\sqrt{\text{var } W_1(t)}} \\
 &= \frac{1}{t} \text{cov}[Z, W_1] \\
 &= \frac{1}{t} \text{cov}[\rho W_1 + \sqrt{1 - \rho^2} W_2, W_1] \\
 &= \frac{1}{t} \left(\rho \text{cov}[W_1, W_1] + \sqrt{1 - \rho^2} \text{cov}[W_2, W_1] \right) \\
 &= \frac{1}{t} \rho \text{var } W_1(t) \\
 &= \rho.
 \end{aligned}$$

Portanto, construímos dois movimentos brownianos *correlacionados* pelo fator ρ , para todo t , a partir de dois movimentos brownianos independentes. Essa construção é canônica na teoria que se segue. \square

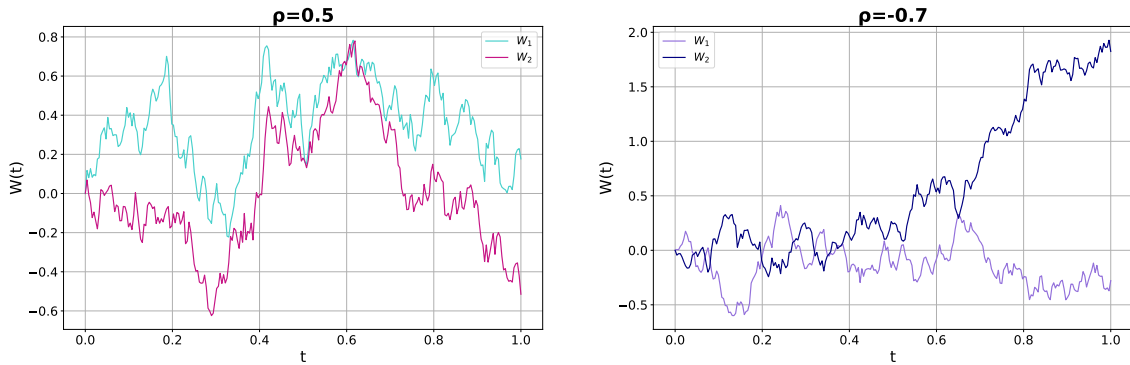


Figura 2.3: Simulação de dois movimentos brownianos correlacionados. Na primeira figura, temos uma correlação positiva, $\rho = 0.5$, enquanto que, na segunda, as trajetórias estão negativamente correlacionadas: $\rho = -0.7$.

A base para transformar sistemas de EDEs com movimentos brownianos correlacionados em sistemas com movimentos brownianos independentes, e vice-versa, é a conhecida decomposição de Cholesky, realizada na matriz de correlação.

Proposição 2.7. *Toda matriz simétrica e definida positiva C possui uma única fatoração, chamada decomposição de Cholesky, da forma*

$$C = LL^T, \quad (2.7)$$

onde L é uma matriz triangular inferior, com entradas diagonais positivas.

O exemplo a seguir fundamenta a escolha do processo $Z(t)$ no exemplo 2.6.

Exemplo 2.8. Decomposição de Cholesky para uma matriz de correlação.
Considere C uma matriz de correlação dada por

$$C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

A ideia da decomposição de Cholesky é encontrar uma matriz L da forma

$$C = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix},$$

de modo que $C = LL^T$. Assim,

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \cdot \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{pmatrix}.$$

Resolvendo o sistema, obtemos:

$$\begin{cases} a = 1 \\ b = \rho \\ c = \sqrt{1 - \rho} \end{cases}.$$

Portanto,

$$L = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho} \end{pmatrix} \quad (2.8)$$

é a matriz procurada pela decomposição de Cholesky. \square

Exemplo 2.9. Correlacionando movimentos brownianos independentes, revisitado.

No exemplo 2.6, vimos como construir movimentos brownianos correlacionados a partir de movimentos brownianos independentes. Agora, vamos reescrever esse procedimento em termos da decomposição de Cholesky.

Realizamos a multiplicação de matrizes:

$$\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix} \cdot \begin{pmatrix} \tilde{W}_1(t) \\ \tilde{W}_2(t) \end{pmatrix} = \begin{pmatrix} \tilde{W}_1(t) \\ \rho\tilde{W}_1(t) + \sqrt{1 - \rho^2}\tilde{W}_2(t) \end{pmatrix} \quad (2.9)$$

e definimos $W_1(t) := \tilde{W}_1(t)$ e $W_2(t) := \sqrt{1 - \rho^2}\tilde{W}_2(t)$. Da mesma forma, determinamos a covariância entre os movimentos brownianos correlacionados:

$$\text{cov}[W_1(t), W_2(t)] = \rho_{12}t,$$

de modo que o coeficiente de correlação entre W_1 e W_2 é ρ . \square

O procedimento acima pode ser generalizado diretamente para dimensões maiores, mas as contas matriciais são mais complicadas. Felizmente, na prática, essa rotina já está implementada nas linguagens de programação, de modo que nos resta apenas entender e saber aplicar o que está sendo feito.

2.2.2 Orstein-Uhlenbeck e Cox-Ingersoll-Ross

Uma das equações diferenciais estocásticas mais antigas a ser registrada foi postulada por Leonard Ornstein e George Eugene Uhlenbeck, em 1930, para descrever a aceleração de uma partícula de pólen suspensa num líquido, a qual estava sujeita a bombardeamento de moléculas. Nesse modelo, X_t representa a velocidade da partícula, dX_t sua variação com o tempo e W_t um "ruído aleatório" causado pelos bombardeamentos.

Quase 50 anos depois, essa EDE foi utilizada para modelar a taxa de juros no chamado modelo de Vasicek, onde, agora, dX_t representa a taxa de juros instantânea, como apresenta Wiersema (2008).

Definição 2.10. O processo de Orstein-Uhlenbeck (OU) é a solução da EDE

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t, \quad (2.10)$$

onde $\kappa, \sigma > 0$ e $\theta \in \mathbb{R}$ são chamados **taxa de reversão à média, volatilidade e média a longo prazo**, respectivamente.

Exemplo 2.11. Resolvendo a EDE de OU.

Considere a função $f(t, x) = xe^{\kappa t}$, que se comportará analogamente ao fator integrante das EDOs clássicas. Suas derivadas são: $\frac{\partial f}{\partial t} = \kappa xe^{\kappa t}$, $\frac{\partial f}{\partial x} = e^{\kappa t}$ e $\frac{\partial^2 f}{\partial x^2} = 0$.

Agora, a fórmula de Itô para $f(t, X_t)$ nos dá:

$$d(X_t e^{\kappa t}) = \left\{ \frac{\partial f}{\partial t} + [\kappa(\theta - X_t)] \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} dt + \sigma \frac{\partial f}{\partial x} dW_t.$$

Substituindo as derivadas na expressão acima resulta em:

$$d(X_t e^{\kappa t}) = \kappa \theta e^{\kappa t} dt + \sigma e^{\kappa t} dW_t.$$

Integrando de ambos os lados de 0 a t , obtemos:

$$\begin{aligned} X_t e^{\kappa t} - X_0 &= \int_0^t \kappa \theta e^{\kappa s} ds + \int_0^t \sigma e^{\kappa s} dW_s \\ &= X_0 + \theta(e^{\kappa t} - 1) + \sigma \int_0^t e^{\kappa s} dW_s. \end{aligned}$$

Reorganizando a expressão anterior, concluímos que a solução é dada por

$$X_t = \theta + e^{-\kappa t}(X_0 - \theta) + \sigma \int_0^t e^{-\kappa(t-s)} dW_s.$$

Um fato interessante é que o processo OU possui uma distribuição estacionária. Para provar isso, primeiramente, note que X_t é um processo gaussiano, pois é a combinação linear de um termo determinístico com uma integral de Itô de um integrando também determinístico (e que, portanto, tem distribuição normal). Assim, sua média e variância determinam completamente sua distribuição.

Vamos então calcular os primeiros dois momentos condicionais ao valor inicial:

$$\begin{aligned}\mathbb{E}[X_t|X_0] &= \mathbb{E}\left[\theta + e^{-\kappa t}(X_0 - \theta) + \sigma \int_0^t e^{-\kappa(t-s)} dW_s \mid X_0\right] \\ &= \theta + e^{-\kappa t}(X_0 - \theta) + \mathbb{E}\left[\sigma \int_0^t e^{-\kappa(t-s)} dW_s\right] \\ &= \theta + e^{-\kappa t}(X_0 - \theta).\end{aligned}$$

Para a variância condicional:

$$\begin{aligned}\text{var}[X_t|X_0] &= \mathbb{E}\left[\left(X_t - \mathbb{E}[X_t|X_0]\right)^2 \mid X_0\right] \\ &= \mathbb{E}\left[\left(\theta + e^{-\kappa t}(X_0 - \theta) + \sigma \int_0^t e^{-\kappa(t-s)} dW_s - \mathbb{E}[X_t|X_0]\right)^2 \mid X_0\right] \\ &= \mathbb{E}\left[\left(\sigma \int_0^t e^{-\kappa(t-s)} dW_s\right)^2 \mid X_0\right], \text{ pois } \mathbb{E}[X_t|X_0] = \theta + e^{-\kappa t}(X_0 - \theta) \\ &= \sigma^2 \int_0^t e^{-2\kappa(t-s)} ds, \text{ pela isometria de Itô} \\ &= \sigma^2 \left[-\frac{1}{2\kappa} e^{-2\kappa(t-s)}\right]_0^t \\ &= \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}).\end{aligned}$$

Vemos que esses momentos são funções do tempo. Entretanto, tomando $t \rightarrow \infty$, encontramos:

$$\begin{aligned}\lim_{t \rightarrow \infty} \mathbb{E}[X_t|X_0] &= \lim_{t \rightarrow \infty} \theta + e^{-\kappa t}(X_0 - \theta) = \theta \\ \lim_{t \rightarrow \infty} \text{var}[X_t|X_0] &= \lim_{t \rightarrow \infty} \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}) = \frac{\sigma^2}{2\kappa}.\end{aligned}$$

Como X_t é gaussiano e seus dois primeiros momentos convergem para limites independentes de X_0 , segue, pelo Teorema da Continuidade de Lévy, que X_t converge fracamente para uma distribuição normal estacionária. \square

Observação 2.3. Reversão à média.

O processo de OU possui uma característica interessante, chamada **reversão à média**. Quando o processo está acima da média a longo prazo θ , o *drift* é negativo, "puxando" o processo de volta à média θ . Similarmente, quando o processo está abaixo da média, seu *drift* o empurra de volta para cima, como podemos observar na figura 2.4. A intensidade com que o processo retorna é dada pelo parâmetro κ .

Exemplo 2.12. O processo OU-exponencial.

Considere $\{X_t\}$ um processo OU. Então, como vimos no exemplo anterior,

$$X_t|X_0 \sim N\left(\theta + e^{-\kappa t}(X_0 - \theta), \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})\right).$$

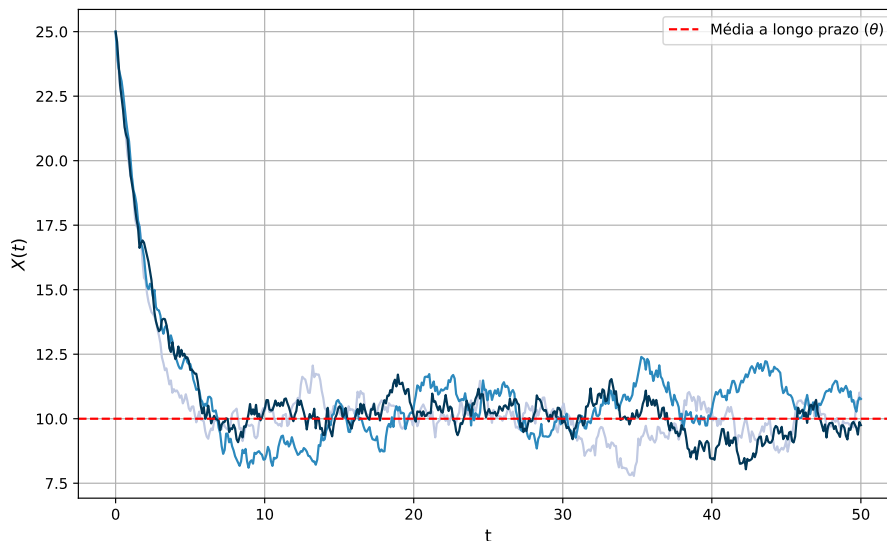


Figura 2.4: Simulação de 3 trajetórias do processo de Orstein-Uhlenbeck com parâmetros: $\kappa = 0.5$, $\theta = 10$, $\sigma = 0.8$ e $X_0 = 25$.

Agora, se considerarmos $S_t = e^{X_t}$, obtemos um modelo log-normal, chamado **OU-exponencial**. Esse modelo é útil para descrever ativos positivos que possuem a propriedade de reversão à média, como *commodities*, taxas de juros, entre outros. Além disso, S satisfaz a EDE

$$dS_t = \kappa \left(\theta + \frac{\sigma^2}{2\kappa} - \ln S_t \right) S_t dt + \sigma S_t dW_t.$$

□

O processo OU é normalmente distribuído e possui diversas propriedades bem conhecidas. Entretanto, uma dessas propriedades é indesejada: a possibilidade do processo tomar valores **negativos**. Sendo assim, ele não é adequado para modelarmos a volatilidade. Precisamos de outras opções.

Exemplo 2.13. Cox-Ingersoll-Ross.

O modelo de Cox-Ingersoll-Ross (CIR) pode ser visto como uma extensão do processo OU, onde a volatilidade é uma função de X_t . Mais especificamente, esse modelo possui dinâmica:

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t. \tag{2.11}$$

O coeficiente de difusão $\sigma\sqrt{X_t}$ impede que o processo assumira valores negativos. Além disso, também é um processo de reversão à média. Daí seu uso na modelagem de taxas de juros.

De maneira semelhante ao que fizemos no exemplo 2.11, podemos calcular:

$$\begin{aligned} \mathbb{E}[X_t|X_0] &= X_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \\ \text{var}[X_t|X_0] &= X_0 \frac{\sigma^2}{\kappa} (e^{-\kappa t} - e^{-2\kappa t}) + \theta \frac{\sigma^2}{2\kappa} (1 - e^{-\kappa t})^2. \end{aligned}$$

Seu comportamento limite é dado por:

$$\begin{aligned}\lim_{t \rightarrow \infty} \mathbb{E}[X_t | X_0] &= \lim_{t \rightarrow \infty} X_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) = \theta \\ \lim_{t \rightarrow \infty} \text{var}[X_t | X_0] &= \lim_{t \rightarrow \infty} X_0 \frac{\sigma^2}{\kappa} (e^{-\kappa t} - e^{-2\kappa t}) + \theta \frac{\sigma^2}{2\kappa} (1 - e^{-\kappa t})^2 = \frac{\theta \sigma^2}{2\kappa}.\end{aligned}$$

Portanto, assim como o processo OU, o modelo CIR converge para uma distribuição estacionária. Neste caso, porém, a distribuição limite é uma Gama com parâmetros de forma $\alpha = \frac{2\kappa\theta}{\sigma^2}$ e escala $\beta = \frac{\sigma^2}{2\kappa}$. \square

Observação 2.4. A condição de Feller.

Apesar de evitados os valores negativos, o processo CIR ainda pode assumir o valor zero (subsequentemente voltando para seus valores positivos, devido a reversão à média). Se quisermos garantir que o processo seja sempre positivo, precisamos da *condição de Feller*:

$$2\kappa\theta \geq \sigma^2. \quad (2.12)$$

Se essa condição não for garantida, o comportamento da função de distribuição de X_t , perto da origem, torna-se mais delicado. Em particular, a possibilidade de a variância atingir o valor zero introduz grandes dificuldades numéricas.

Na prática, entretanto, a condição de Feller é frequentemente violada quando se trabalha com dados financeiros reais. Nesses casos, ao ocorrer $v_t = 0$, costuma-se recorrer a técnicas alternativas de precificação ou a esquemas numéricos específicos, desenvolvidos para lidar adequadamente com a degenerescência do processo na fronteira. \square

2.3 O modelo de Heston

Em 1993, Steven L. Heston propôs uma formulação para o preço de opções europeias por meio de uma técnica inovadora: uma solução baseada na função característica do log-preço. Nesse trabalho, ele também mostrou, por meio de simulações, que a correlação entre o preço do ativo e sua volatilidade desempenha um papel fundamental na explicação da assimetria (*skewness*) dos retornos e dos vieses observados nos preços das opções no modelo de Black–Scholes.

2.3.1 Descrição matemática do modelo

No modelo de Heston (HSV), temos duas EDEs: uma para o ativo subjacente, $S(t)$, e outra para a variância, $v(t) = \sigma^2(t)$, de modo que, sob a medida \mathbb{Q} , temos:

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t dW_t^S, & S_{t_0} = S_0 > 0 \\ dv_t = \kappa(\bar{v} - v_t) dt + \gamma \sqrt{v_t} dW_t^v, & v_{t_0} = v_0 > 0 \end{cases} \quad (2.13)$$

onde $dW_t^S dW_t^v = \rho dt$. Os parâmetros κ , \bar{v} e γ são chamados **velocidade de reversão à média**, **média a longo prazo** do processo v e **volatilidade da volatilidade** (muitas vezes chamado de *vol-vol*), respectivamente.

Exemplo 2.14. Propriedades estatísticas do modelo CIR.

Sendo v_t um processo CIR, suas propriedades são bem conhecidas. Por exemplo, ele é um processo com reversão à média, cuja velocidade de reversão é determinada pelo parâmetro κ .

Além disso, como vimos no exemplo 2.13, dados $s < t$, temos:

$$\mathbb{E}[v_t|v_s] = \bar{v} + (v_s - \bar{v})e^{-\kappa(t-s)}, \quad (2.14)$$

$$\text{var}[v_t|v_s] = \frac{v_s \gamma^2 e^{-\kappa(t-s)}}{\kappa} (1 - e^{-\kappa(t-s)}) + \frac{\bar{v} \gamma^2}{2\kappa} (1 - e^{-\kappa(t-s)})^2. \quad (2.15)$$

□

Embora, ao longo deste trabalho, adotemos diretamente a medida neutra ao risco para o desenvolvimento da teoria do modelo de Heston (uma vez que nosso objetivo principal é a precificação de derivativos), bem como sua estimação de parâmetros, é pedagógico compreender como essa medida pode ser obtida a partir da medida física.

Essa derivação tem o papel de explicitar o efeito da mudança de medida sobre a dinâmica do modelo, em particular sobre os coeficientes da equação da variância, além de constituir um bom exercício de aplicação do Teorema de Girsanov.

Exemplo 2.15. De \mathbb{P} para \mathbb{Q} no modelo de Heston.

Suponha que o HSV tenha sua dinâmica descrita a partir da medida real \mathbb{P} :

$$\begin{cases} dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_S^{\mathbb{P}} \\ dv_t = \kappa(\bar{v}^{\mathbb{P}} - v_t) dt + \gamma \sqrt{v_t} dW_v^{\mathbb{P}} \end{cases}, \quad (2.16)$$

com correlação $dW_S^{\mathbb{P}}(t)dW_v^{\mathbb{P}}(t) = \rho dt$. Através da decomposição de Cholesky, podemos reescrever o modelo em termos de movimentos brownianos independentes:

$$\begin{cases} dS_t = \mu S_t dt + \sqrt{v_t} S_t d\tilde{W}_S^{\mathbb{P}} \\ dv_t = \kappa(\bar{v}^{\mathbb{P}} - v_t) dt + \gamma \sqrt{v_t} \left(\rho d\tilde{W}_S^{\mathbb{P}} + \sqrt{1 - \rho^2} d\tilde{W}_v^{\mathbb{P}} \right). \end{cases}$$

Agora, estamos procurando um processo Ξ_t em

$$d\tilde{W}_S^{\mathbb{Q}} = \Xi_t dt + d\tilde{W}_S^{\mathbb{P}},$$

de modo que $\bar{S}_t = S_t/B_t$ seja uma martingale. A fórmula de Itô aplicada em \bar{S}_t fornece:

$$\frac{d\bar{S}_t}{\bar{S}_t} = (\mu - r) dt + \sqrt{v_t} d\tilde{W}_S^{\mathbb{P}}. \quad (2.17)$$

Portanto, encontramos a seguinte transformação de medidas:

$$\begin{cases} d\tilde{W}_S^{\mathbb{P}} = d\tilde{W}_S^{\mathbb{Q}} - \frac{\mu - r}{\sqrt{v_t}} dt, \\ d\tilde{W}_v^{\mathbb{P}} = d\tilde{W}_v^{\mathbb{Q}}. \end{cases} \quad (2.18)$$

Note que é necessário satisfazer $v_t > 0$.

Assim, sob \mathbb{Q} , a dinâmica do modelo de Heston é dada por

$$dS_t = rS_t dt + \sqrt{v_t} S_t d\tilde{W}_S^{\mathbb{Q}},$$

e o processo de variância por:

$$\begin{aligned} dv_t &= \kappa(\bar{v}^{\mathbb{P}} - v_t)dt + \gamma\sqrt{v_t}\left(\rho d\tilde{W}_S^{\mathbb{P}} + \sqrt{1-\rho^2}d\tilde{W}_v^{\mathbb{P}}\right) \\ &= \kappa\left[\bar{v}^{\mathbb{P}} - \frac{\rho\gamma(\mu-r)}{\kappa} - v_t\right]dt + \gamma\sqrt{v_t}\left[\rho d\tilde{W}_S^{\mathbb{Q}} + \sqrt{1-\rho^2}d\tilde{W}_v^{\mathbb{Q}}\right]. \end{aligned}$$

Tomando $\bar{v} := \bar{v}^{\mathbb{P}} - \rho\gamma(\mu-r)/\kappa$ e retornando aos movimentos brownianos correlacionados, obtemos a EDE (2.13) para a variância na medida \mathbb{Q} . \square

A EDP de precificação

Em modelos de difusão contínua, como os modelos de Black-Scholes e Heston, o preço de não arbitragem de um derivativo europeu pode ser visto como uma função determinística do tempo e das variáveis de estado do modelo.

Mais especificamente, ao supor que o preço do derivativo possa ser representado por uma função suficientemente regular $V(t, S_t, v_t)$, a exigência de que seu valor descontado seja uma martingale sob \mathbb{Q} impõe que essa função satisfaça uma equação diferencial parcial. Essa EDP expressa de forma precisa a condição de ausência de arbitragem do modelo.

Assim, para obtermos a EDP de Heston, considere, sob \mathbb{Q} , o seguinte problema de precificação:

$$V(t, S, v) = B_t \mathbb{E}\left[\frac{V(T, S, v)}{B_T} \middle| \mathcal{F}_t\right], \quad (2.19)$$

com V sendo o **valor presente descontado**, $\mathcal{F}_t = \sigma(s, S, v : s \leq t)$ e B_t nosso ativo livre de risco usual.

Temos

$$\mathbb{E}\left[\frac{V(T, S, v)}{B_T} \middle| \mathcal{F}_t\right] = \frac{V(t, S, v)}{B_t}. \quad (2.20)$$

Suponha que nossa função preço descontado V/B seja suficientemente diferenciável. Ela também deve ser uma martingale. Sua dinâmica é dada por:

$$d\left(\frac{V}{B}\right) = \frac{1}{B}dV - r\frac{V}{B}dt.$$

De (2.13) e da fórmula de Itô, segue:

$$\begin{aligned} dV &= \left(\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \kappa(\bar{v} - v)\frac{\partial V}{\partial v} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\gamma^2v\frac{\partial^2 V}{\partial v^2} + \rho\gamma Sv\frac{\partial^2 V}{\partial S\partial v}\right)dt \\ &\quad + S\sqrt{v}\frac{\partial V}{\partial S}dW_S + \gamma\sqrt{v}\frac{\partial V}{\partial v}dW_v. \end{aligned}$$

Para que a propriedade martingale valha, os termos com dt na dinâmica devem zerar, de modo que, quando substituirmos na dinâmica de V/B ,

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \kappa(\bar{v} - v)\frac{\partial V}{\partial v} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\gamma^2v\frac{\partial^2 V}{\partial v^2} + \rho\gamma Sv\frac{\partial^2 V}{\partial S\partial v} - rV = 0, \quad (2.21)$$

obtemos a EDP procurada.

Agora, apliquemos esse resultado a um caso mais concreto.

Exemplo 2.16. O preço de uma *call* europeia no modelo de Heston.

Primeiramente, com o objetivo de transformar a equação numa EDP homogênea, façamos a seguinte mudança de variáveis: Considere $\tau = T - t$, $x = \ln(S/K) + r\tau$ e $V(t, S, v) = e^{-r\tau}C(\tau, x, v)$. Aplicando as regras da cadeia e produto, obtemos:

$$\frac{\partial V}{\partial t} = e^{-r\tau} \left(-\frac{\partial C}{\partial \tau} - r \frac{\partial C}{\partial x} + rC \right),$$

$$\frac{\partial V}{\partial S} = e^{-r\tau} \frac{1}{S} \frac{\partial C}{\partial x},$$

$$\frac{\partial V}{\partial v} = e^{-r\tau} \frac{\partial C}{\partial v}.$$

Derivando novamente,

$$\frac{\partial^2 V}{\partial S^2} = e^{-r\tau} \frac{1}{S^2} \left(\frac{\partial^2 C}{\partial x^2} - \frac{\partial C}{\partial x} \right),$$

$$\frac{\partial^2 V}{\partial S \partial v} = e^{-r\tau} \frac{1}{S} \frac{\partial^2 C}{\partial x \partial v},$$

$$\frac{\partial^2 V}{\partial v^2} = e^{-r\tau} \frac{\partial^2 C}{\partial v^2}.$$

Substituindo essas expressões na EDP original, obtemos a seguinte equação homogênea:

$$-\frac{\partial C}{\partial \tau} + \frac{1}{2}v \frac{\partial^2 C}{\partial x^2} - \frac{1}{2}v \frac{\partial C}{\partial x} + \frac{1}{2}\gamma^2 v \frac{\partial^2 C}{\partial v^2} + \rho\gamma v \frac{\partial^2 C}{\partial x \partial v} - \kappa(v - \bar{v}) \frac{\partial C}{\partial v} = 0. \quad (2.22)$$

Por analogia ao modelo de Black-Scholes, propomos uma solução da seguinte forma:

$$C(t, v, \tau) = K[e^x P_1(x, v, \tau) - P_0(x, v, \tau)]. \quad (2.23)$$

Substituindo nossa solução proposta na nova equação, por exemplo em

$$\begin{aligned} \frac{\partial C}{\partial x} &= \frac{\partial}{\partial x} [Ke^x P_1(x, v, \tau) - KP_0(x, v, \tau)] \\ &= \frac{\partial}{\partial x} [Ke^x P_1(x, v, \tau)] - \frac{\partial}{\partial x} [KP_0(x, v, \tau)] \\ &= \left(Ke^x P_1 + Ke^x \frac{\partial P_1}{\partial x} \right) - K \frac{\partial P_0}{\partial x}, \end{aligned}$$

implica que P_0 e P_1 devem satisfazer:

$$-\frac{\partial P_j}{\partial \tau} + \frac{1}{2}v \frac{\partial^2 P_j}{\partial x^2} - \left(\frac{1}{2} - j \right) v \frac{\partial P_j}{\partial x} + \frac{1}{2}\gamma^2 v \frac{\partial^2 P_j}{\partial v^2} + \rho\gamma v \frac{\partial^2 P_j}{\partial x \partial v} + (a - b_j v) \frac{\partial P_j}{\partial v} = 0, \quad (2.24)$$

para $j = 0, 1$, onde $a = \kappa \bar{v}$ e $b_j = \kappa - j\rho\gamma$ estão sujeitos à condição terminal

$$\lim_{\tau \rightarrow 0} P_j(x, v, \tau) = 1_{\{x > 0\}}, \quad (2.25)$$

que decorre diretamente do pagamento da opção de compra europeia.

Com efeito, no vencimento ($\tau = 0$), temos $x_T = \ln(S_T/K)$, de modo que o evento $\{x > 0\}$ é equivalente a $\{S_T > K\}$. Como o *payoff* da opção de compra é nulo fora dessa região e positivo em seu interior, a condição terminal das funções P_j deve reproduzir exatamente essa estrutura. Consequentemente, no limite $\tau \rightarrow 0$, P_j coincide com a indicadora do conjunto $\{x > 0\}$.

Agora, resolveremos a equação (2.24) sujeita à condição (2.25) utilizando o método da *transformada de Fourier* com o intuito de converter as EDPs em EDOs. Essa técnica funciona bem aqui pois o modelo de Heston pertence a uma classe maior chamada *processos de difusão afins* - ver Osterlee, Grzelak (2020) - para a qual a função característica do processo está bem definida.

Seguindo a convenção mais usual em probabilidade/finanças, a *transformada de Fourier* de P_j é definida por

$$\hat{P}_j(k, v, \tau) = \int_{-\infty}^{+\infty} e^{-ikx} P(x, v, \tau) dx,$$

de modo que

$$\hat{P}_j(k, v, 0) = \int_{-\infty}^{+\infty} e^{-ikx} 1_{\{x > 0\}} dx = \frac{1}{-ik} e^{-ikx} \Big|_0^{+\infty} = \frac{1}{ik}.$$

Já a transformada inversa é:

$$P(x, v, \tau) = \int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{ikx} \hat{P}(k, v, \tau) dk. \quad (2.26)$$

Das propriedades da transformada, vemos que as derivadas de \hat{P}_j com respeito à x correspondem à multiplicação por ik . Assim, aplicando a transformada em (2.24), resulta em:

$$-\frac{\partial \hat{P}_j}{\partial \tau} + i\rho\gamma kv \frac{\partial \hat{P}_j}{\partial v} - \frac{1}{2} k^2 v \hat{P}_j + \frac{1}{2} \gamma^2 v \frac{\partial^2 \hat{P}_j}{\partial v^2} - i \left(\frac{1}{2} - j \right) kv \hat{P}_j + (a - b_j v) \frac{\partial \hat{P}_j}{\partial v} = 0. \quad (2.27)$$

Rearranjando, obtemos:

$$v \left(\alpha_j \hat{P}_j - \beta_j \frac{\partial \hat{P}_j}{\partial v} + \eta \frac{\partial^2 \hat{P}_j}{\partial v^2} \right) + a \frac{\partial \hat{P}_j}{\partial v} - \frac{\partial \hat{P}_j}{\partial \tau} = 0, \quad (2.28)$$

onde

$$\alpha_j = -\frac{k^2}{2} - ik \left(\frac{1}{2} - j \right), \quad \beta_j = b_j - ik\rho\gamma, \quad \eta = \frac{\gamma^2}{2}.$$

Agora, mais uma vez propomos uma solução. Dessa vez,

$$\begin{aligned}\hat{P}_j(k, v, \tau) &= \exp\{C_j(k, \tau)\bar{v} + D_j(k, \tau)v\}\hat{P}_j(k, v, 0) \\ &= \frac{1}{ik} \exp\{C_j(k, \tau)\bar{v} + D_j(k, \tau)v\}.\end{aligned}\quad (2.29)$$

Tomando as seguintes derivadas de (2.29):

$$\frac{\partial \hat{P}_j}{\partial \tau} = \left[\bar{v} \frac{\partial C_j}{\partial \tau} + v \frac{\partial D_j}{\partial \tau} \right] \hat{P}_j, \quad \frac{\partial \hat{P}_j}{\partial v} = D_j \hat{P}_j, \quad \frac{\partial^2 \hat{P}_j}{\partial v^2} = D_j^2 \hat{P}_j,$$

e inserindo-as de volta em (2.28) nos dá:

$$v(\alpha_j - \beta_j D_j + \eta D_j^2) + a D_j - \left[\bar{v} \frac{\partial C_j}{\partial \tau} + v \frac{\partial D_j}{\partial \tau} \right] = 0. \quad (2.30)$$

Isso implica que as duas equações abaixo devem ser satisfeitas:

$$\begin{cases} \frac{\partial C_j}{\partial \tau} = \kappa D_j \\ \frac{\partial D_j}{\partial \tau} = \alpha_j - \beta_j D_j + \eta D_j^2 \end{cases}, \quad (2.31)$$

onde a EDO para D_j é conhecida como *equação de Riccati*.

As raízes da equação quadrática r_-^j e r_+^j são dadas por:

$$r_{\pm}^j = \frac{\beta_j \pm \sqrt{\beta_j^2 - 4\alpha_j\eta}}{2\eta} = \frac{b_j - ik\rho\gamma \pm d_j}{\gamma^2},$$

onde

$$d_j = \sqrt{(b_j - ik\rho\gamma)^2 + \gamma^2[k^2 - 2ik(j - 1/2)]}.$$

Integrando o sistema de equações diferenciais ordinárias (2.31) em relação à τ , com as condições iniciais $C(k, 0) = 0 = D(k, 0)$, resulta em:

$$D_j(k, \tau) = r_-^j \left(\frac{1 - e^{-d_j\tau}}{1 - g_j e^{-d_j\tau}} \right) \quad (2.32)$$

e

$$C_j(k, \tau) = \frac{\kappa}{\gamma^2} \left[r_-^j \gamma^2 \tau - 2 \ln \left(\frac{1 - g_j e^{-d_j\tau}}{1 - g_j} \right) \right], \quad (2.33)$$

onde definimos $g_j := r_-^j / r_+^j$.

Por fim, resta aplicar a transformada inversa para recuperarmos P_j . Isso pode ser feito através do laborioso cálculo direto da integral complexa, ou ainda, utilizar o resultado proposto por Gil-Pelaez (1951). O resultado final é bem conhecido na literatura e é dado por

$$P_j(x, v, \tau) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left\{ \frac{\exp\{C_j(k, \tau)\bar{v} + D_j(k, \tau)v + ikx\}}{ik} \right\} dk. \quad (2.34)$$

Os detalhes técnicos desse desenvolvimento podem ser encontrados em Sapozito (2023), Rouah (2013) ou Osterlee, Grzelak (2020). \square

Observação 2.5. Sobre os termos P_0 e P_1 .

No exemplo 2.16, surgem os coeficientes P_0 e P_1 na expressão para o preço da *call*. Veremos, agora, que eles possuem significados probabilístico e financeiro reais. De fato, primeiramente perceba que a solução proposta em (2.23) é uma representação do *valor futuro* da opção. Note que, sendo $x = \ln(S/K) + r\tau$, podemos escrever (2.23) como:

$$C = S_t e^{r\tau} P_1 - K P_0.$$

Podemos, ainda, reescrevê-la em termos do *valor presente*, descontando-a por $e^{-r\tau}$:

$$C = [S_t e^{r\tau} P_1 - K P_0] e^{-r\tau} = S_t P_1 - K e^{-r\tau} P_0. \quad (2.35)$$

Agora, considere o preço da opção (definido em termos do valor esperado do *payoff*):

$$\begin{aligned} C(K) &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}}[(S_T - K)^+] \\ &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}}[(S_T - K) \mathbf{1}_{\{S_T > K\}}] \\ &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}}[S_T \mathbf{1}_{\{S_T > K\}}] - K e^{-r\tau} \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_{\{S_T > K\}}]. \end{aligned} \quad (2.36)$$

Analisemos essas esperanças separadamente. Note que, por definição,

$$\mathbb{E}_{\mathbb{Q}}[\mathbf{1}_{\{S_T > K\}}] = \mathbb{Q}(S_T > K).$$

Para avaliar a outra esperança, considere \mathbb{Q}^S e a derivada de Radon-Nikodym:

$$\frac{d\mathbb{Q}}{d\mathbb{Q}^S} = \frac{B_T/B_t}{S_T/S_t} = \frac{e^{r(T-t)} S_t}{S_T}. \quad (2.37)$$

Daí, segue que:

$$\begin{aligned} e^{-r\tau} \mathbb{E}_{\mathbb{Q}}[S_T \mathbf{1}_{\{S_T > K\}}] &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}^S} \left[S_T \mathbf{1}_{\{S_T > K\}} \frac{d\mathbb{Q}}{d\mathbb{Q}^S} \right] \\ &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}^S} \left[S_T \mathbf{1}_{\{S_T > K\}} \frac{e^{r(T-t)} S_t}{S_T} \right] \\ &= e^{-r\tau} \mathbb{E}_{\mathbb{Q}^S} [\mathbf{1}_{\{S_T > K\}} e^{r\tau} S_t] \\ &= S_t \mathbb{E}_{\mathbb{Q}^S} [\mathbf{1}_{\{S_T > K\}}] \\ &= S_t \mathbb{Q}^S(S_T > K). \end{aligned}$$

Dessa forma, podemos escrever

$$C(K) = S_t \mathbb{Q}^S(S_T > K) - K e^{-r\tau} \mathbb{Q}(S_T > K). \quad (2.38)$$

Comparando a última equação com (2.35), concluímos que P_0 é a probabilidade, na medida neutra ao risco, da opção vencer *ITM*, enquanto P_1 representa a probabilidade de exercício da opção na medida \mathbb{Q}^S , utilizando o próprio preço do ativo como numerário. \square

A distribuição do log-preço

Agora, estamos em condições de apresentar a função caracterísitca de Heston. Por definição, a f.c. condicional de uma v.a. x é dada por

$$\phi_T(u) = \mathbb{E}[e^{iux_T} | x_t = 0].$$

Como vimos, a probabilidade do log-preço final x_T ser maior que o preço de exercício, isto é, a probabilidade de a opção valer alguma coisa, é dada por

$$\begin{aligned} \Pr(x_T > x) &= P_0(x, v, \tau) \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left\{ \frac{\exp\{C_0(u, \tau)\bar{v} + D_0(u, \tau)v + iux\}}{iu} \right\} du, \end{aligned}$$

com $x = \ln(S_t/K)$ e $\tau = T - t$. Defina também o *log-strike* k por $k := \ln(K/S_t)$. Assim, a densidade de $k = -x$ é dada por:

$$\begin{aligned} f(k) &= -\frac{\partial P_0}{\partial k} \\ &= \frac{1}{2\pi} \int_{-\infty}^\infty \exp\{C_0(u', \tau)\bar{v} + D_0(u', \tau)v - iu'k\} du'. \end{aligned} \tag{2.39}$$

Dessa forma, a função característica (para a medida \mathbb{Q}) no modelo de Heston possui a seguinte forma:

$$\begin{aligned} \phi_T(u) &= \int_{-\infty}^{+\infty} e^{iuk} f(k) dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\{C_0(u', \tau)\bar{v} + D_0(u', \tau)v\} du' \int_{-\infty}^{+\infty} e^{i(u-u')k} du \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\{C_0(u', \tau)\bar{v} + D_0(u', \tau)v\} 2\pi \delta(u - u') du' \\ &= \int_{-\infty}^{+\infty} \exp\{C_0(u', \tau)\bar{v} + D_0(u', \tau)v\} \delta(u - u') du' \\ &= \exp\{C_0(u, \tau)\bar{v} + D_0(u, \tau)v\}, \end{aligned} \tag{2.40}$$

onde δ é a *delta de Dirac* e, na última igualdade, utilizamos sua propriedade de filtragem.

2.3.2 Implementação numérica

Agora, com a função característica descontada em mãos, podemos resolver o problema de precificação neutra ao risco.

Até o presente momento, temos utilizado a discretização de Euler-Maruyama para simular nossos processos estocásticos. Entretanto, ela já não nos serve tão bem pois, se simplesmente fizermos

$$v_{i+1} = v_i + \kappa(\bar{v} - v_i)\Delta t + \gamma\sqrt{v_i}\sqrt{\Delta t}Z,$$

com $Z \sim N(0, 1)$, podemos atingir variâncias negativas com probabilidade não desprezível, dependendo dos parâmetros.

Existem diversas maneiras de corrigir esse método, mas podemos contornar esse problema de maneira simples implementando um esquema de *discretização de Milstein*. Mais especificamente, acrescentando mais uma ordem superior na expansão de Itô-Taylor de $v(t + \Delta t)$, resulta na seguinte discretização para o processo de variância:

$$v_{i+1} = v_i + \kappa(\bar{v} - v_i)\Delta t + \gamma\sqrt{v_i}\sqrt{\Delta t}Z + \frac{\gamma^2}{4}\Delta t(Z^2 - 1). \quad (2.41)$$

Mais ainda, podemos acrescentar, por exemplo, uma *hipótese de absorção*, isto é, tomar $v_i^+ = \max\{0, v_i\}$ para garantir a não-negatividade da volatilidade.

Nesse contexto, deve-se utilizar a seguinte discretização para o processo do ativo, que é consequência da fórmula de Itô para $x_t = \ln S_t$:

$$x_{i+1} = x_i - \frac{v_i}{2}\Delta t + \sqrt{v_i}\Delta tW, \quad (2.42)$$

com $x_i := \ln(S_i/S_0)$, $W \sim N(0, 1)$ e $\mathbb{E}[ZW] = \rho$.

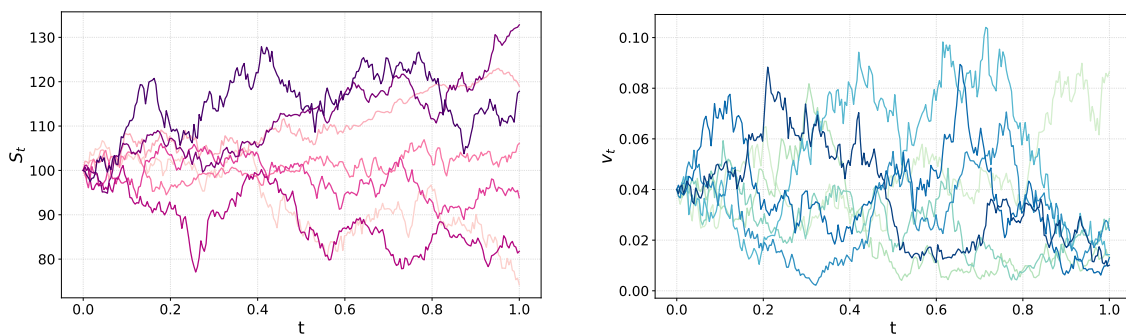


Figura 2.5: Simulação das trajetórias do preço do ativo S_t e do processo de variância v_t , respectivamente, no HSV. Os parâmetros utilizados foram: $S_0 = 100$, $v_0 = 0.04$, $r = 0.05$, $\kappa = 2$, $\bar{v} = 0.04$ e $\gamma = 0.03$, para um coeficiente de correlação $\rho = -0.7$ entre os movimentos brownianos.

Observação 2.6. Alavancagem e o sinal de ρ .

O chamado **efeito de alavancagem** refere-se à tendência observada no mercado de os retornos de um ativo estarem negativamente correlacionados à volatilidade do mesmo. Isto é, tipicamente, um aumento no preço dos ativos está acompanhado de uma queda na volatilidade, e vice-versa. Entretanto, está bem documentado que esse efeito é assimétrico, no sentido de que quedas no preço do ativo são acompanhadas por maiores aumentos na volatilidade do que o contrário.

Essa interpretação econômica para esse fenômeno foi introduzida em 1976 por Fischer Black em seu artigo "*Studies in stock price volatility changes*". Apesar de amplamente aceita, ela não é a única. Uma explicação alternativa é a hipótese de retorno de volatilidade (*volatility feedback*), defendida por autores como Campbell e Hentschel (1992) e French, Schwert e Stambaugh (1987). Para esses autores, a relação é oposta: se a volatilidade aumenta, o mercado passa a exigir um prêmio de risco maior, o que pressiona o preço do ativo para baixo.

2.3.3 A SVI no modelo de Heston

Estudo de parâmetros para a curva de volatilidade implícita

Discutiremos agora os padrões de volatilidade implícita que podem ser gerados pelo modelo de Heston. Como sabemos, a variância $v_t = \sigma_t^2$ é governada por um processo de reversão à média CIR, e cada parâmetro possui um efeito específico na curva gerada pela dinâmica proposta.

Entender o efeito dos parâmetros do modelo pode ser útil, por exemplo, durante a estimação de parâmetros ou a calibração, onde um conjunto inicial de parâmetros precisa ser prescrito a priori, antes das iterações começarem, para ajustar os parâmetros do HSV, $\{\rho, v_0, \bar{v}, \kappa, \gamma\}$, para corresponder aos dados de mercado.

Seguindo Oosterlee e Grzelak (2020), para analisarmos numericamente os efeitos dos parâmetros, usaremos os seguintes parâmetros de referência:

$$S_0 = 100, \tau = 2, r = 0.05, v_0 = 0.05, \kappa = 0.1, \bar{v} = 0.1, \rho = -0.75, \gamma = 0.1.$$

O estudo numérico foi realizado variando individualmente diversos parâmetros, enquanto os outros foram mantidos constantes. Para cada conjunto de parâmetros, foram calculados preços de *call* no HSV, para uma grade de *strikes*, utilizando as funções que implementamos. Cada um desses preços foi utilizado como "preço de mercado" na função que calcula a volatilidade implícita (utilizando Newton-Raphson) e, então, as curvas *smile* foram plotadas.

Correlação e vol-vol

Começamos com os parâmetros que mais intefere no formato da curva de volatilidade implícita: ρ e γ . A figura 2.6 mostra que, conforme ρ decresce, a inclinação da *smile* aumenta, o que chamamos de **volatility skew** e está relacionado ao terceiro momento - *skewness* - da distribuição dos retornos que está implícita nos preços das opções. Já o parâmetro γ , relacionado ao quarto momento - *kurtosis* - afeta proporcionalmente a curvatura da curva *smile*.

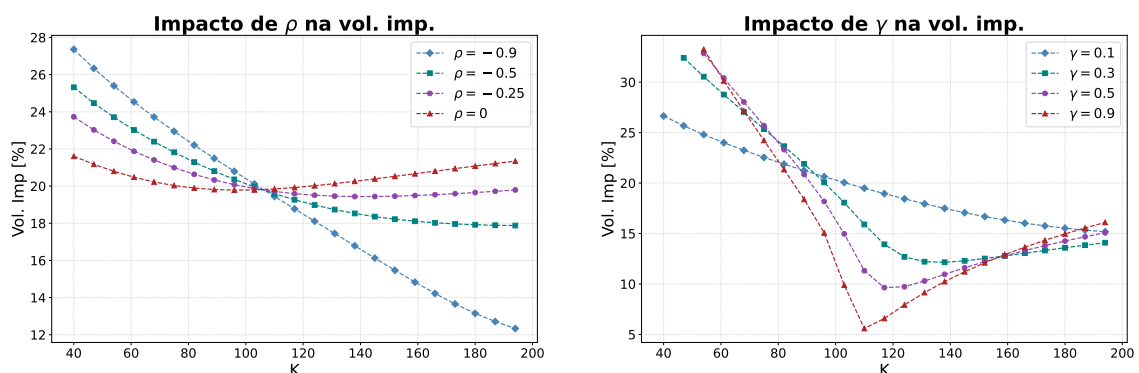


Figura 2.6: Impacto da variação da correlação ρ e da vol-vol γ na volatilidade implícita, como uma função do preço de exercício K .

Reversão à média

Na Figura 2.7, observa-se que a taxa de reversão à média κ exerce influência limitada sobre o formato do *smile* de volatilidade implícita. Em contrapartida, seu efeito é mais pronunciado na estrutura a termo da volatilidade: valores mais elevados de κ promovem uma convergência mais rápida da volatilidade implícita para o nível de longo prazo $\sqrt{\bar{v}}$, especialmente para opções ATM. Do ponto de vista visual, isso se manifesta pelo achatamento mais rápido da estrutura a termo da volatilidade implícita.

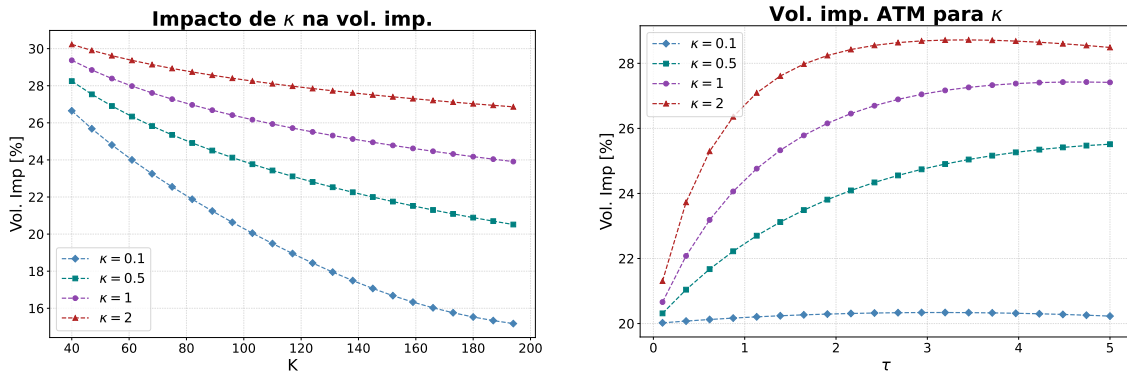


Figura 2.7: Impacto da variação da velocidade de reversão à média κ como uma função do preço de exercício K e impacto da variação de κ na volatilidade implícita ATM, como uma função da maturidade τ .

Variância inicial

Por fim, a Figura 2.8 mostra que a variância inicial v_0 e a média a longo prazo, que estão relacionados ao segundo momento, possuem efeito parecido na curva de volatilidade implícita: controlam o nível da curva. Intuitivamente, elas "puxam" o gráfico para cima conforme v_0 e \bar{v} crescem. Entretanto, seu efeito é limitado, cerca de 1% a 2% apenas. Além disso, ambos dependem de κ .

Observação 2.7. Vantagens e desvantagens dos modelos de volatilidade estocástica.

A principal vantagem desses modelos é a capacidade de replicar, ao menos de forma aproximada, as curvas *smile* e a dinâmica dos dados observados no mercado. Como vimos, algumas características desse comportamento são: caudas mais pesadas na distribuição dos retornos, reversão à média de longo prazo para a volatilidade e o efeito de alavancagem.

Outra vantagem de alguns modelos é a disponibilidade de fórmulas (quase) fechadas para os preços das opções europeias, como no modelo de Heston, e até para a volatilidade implícita.

Entre as desvantagens, destaca-seo fato de que tais modelos, em geral, não conseguem reproduzir exatamente a superfície de volatilidade implícita observada

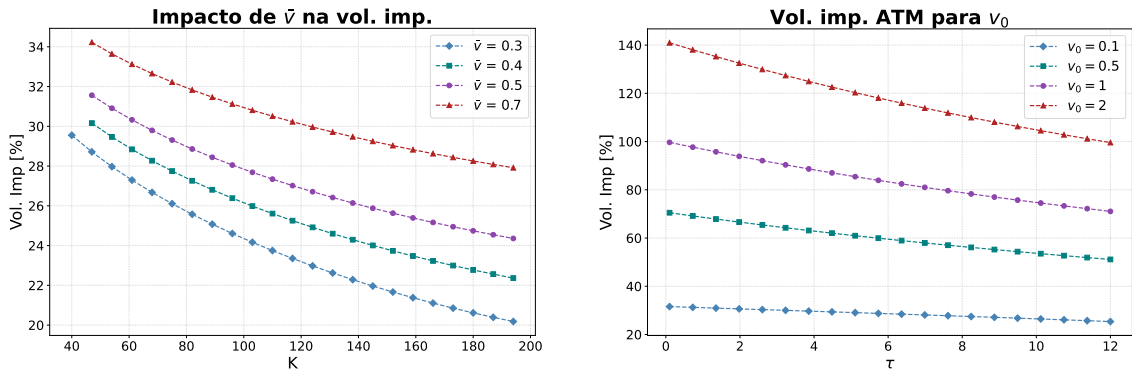


Figura 2.8: Impacto das mudanças na média a longo prazo \bar{v} como função do *strike* K e impacto da variação de v_0 na volatilidade implícita ATM, como função de τ .

no mercado. Além disso, a precificação de derivativos exóticos (como opções asiáticas, de barreira, bermudas, cestas, etc.) tende a ser computacionalmente custosa, seja por métodos numéricos baseados em equações diferenciais parciais. No escopo deste trabalho, contudo, essas limitações não são centrais, uma vez que nos restringimos ao modelo de Heston e à precificação de opções *vanilla*.

Capítulo 3

Estimação de Parâmetros

Principais referências utilizadas ao longo deste capítulo: Oosterlee e Grzelak (2020), Saporito (2023) e Seber e Wild (2003) e Press et al. (2007).

3.1 Fundamentos

Na literatura financeira, o termo *calibração* é frequentemente utilizado de forma ampla para descrever o ajuste de modelos a dados observados. Neste trabalho, empregamos um procedimento no qual os parâmetros do modelo são obtidos a partir da resolução de um único problema de otimização, resultando em um vetor fixo de parâmetros.

Por essa razão, e a fim de evitar ambiguidades terminológicas, ao longo deste capítulo adotaremos a expressão *estimação de parâmetros* para nos referirmos ao procedimento utilizado.

Formulamos essa tarefa como um problema de regressão não linear. Dentre as diversas abordagens disponíveis (máxima verossimilhança, método dos momentos generalizados e até redes neurais) optamos pelo método dos mínimos quadrados, escolha que se justifica tanto por sua simplicidade teórica quanto por sua ampla utilização na prática financeira, em particular no ajuste de superfícies de volatilidade implícita.

3.1.1 O Método dos Mínimos Quadrados

Considere n observações (x_i, y_i) , $i = 1, \dots, n$, de um modelo de regressão não linear com relação funcional f . Podemos representar isso como:

$$\begin{aligned} y_i &= \text{modelo teórico} + \text{ruído} \\ &= f(x_i; \theta) + \varepsilon_i, \end{aligned} \tag{3.1}$$

onde ε é um erro aleatório, com $\mathbb{E}\varepsilon_i = 0$, x_i é um vetor de variáveis independentes e θ um vetor de parâmetros. No caso do modelo de Heston, $\theta = (\rho, \gamma, \bar{v}, \kappa, v_0) \in \mathbb{R}^5$.

O método dos mínimos quadrados consiste em buscar o estimador $\hat{\theta}$ que

minimiza a soma dos erros quadráticos, definida por

$$L(\theta) = \sum_{i=1}^n [y_i - f(x_i; \theta)]^2. \quad (3.2)$$

Na literatura, essa função L é conhecida como *loss* ou *target function*

Portanto, o **estimador de mínimos quadrados**, $\hat{\theta}$, é o argumento que minimiza $L(\theta)$:

$$\hat{\theta} = \arg \min_{\theta} L(\theta). \quad (3.3)$$

Observação 3.1. Alguns comentários.

Deve-se notar que, ao contrário do caso linear, $L(\theta)$ pode assumir diversos mínimos locais além do mínimo global. Além disso, supondo que os ε_i são i.i.d., pode-se mostrar que, sob certas condições de regularidade, $\hat{\theta}$ é um estimador consistente para θ .

Trazendo para o nosso contexto, considere um conjunto de m maturidades distintas τ_1, \dots, τ_m . Para cada τ_i , temos um conjunto de n_i preços de exercício K_{i1}, \dots, K_{in_i} , $j = 1, \dots, n_i$, e, para cada par (τ_i, K_{ij}) , temos um preço de mercado, $V_{ij}^{mkt}(\tau_i, K_{ij})$, e um preço teórico correspondente, $V(\tau_i, K_{ij}; \theta)$, gerado pelo modelo de Heston. Esses preços possuem volatilidades implícitas correspondentes dadas por $\sigma_{imp}^{mkt}(\tau_i, K_{ij})$ e $\sigma_{imp}(\tau_i, K_{ij}; \theta)$, respectivamente.

Associado à diferença, é possível ainda atribuir pesos w_{ij} para refinar o processo de otimização. Dependendo do caso, pode-se tomar os pesos todos iguais 1 e, após diversos testes, optamos por seguir essa abordagem. Entretanto, Saporito (2023), por exemplo, sugere outras escolhas, como $w_{ij} = 1/(mn_i)$ ou $w_{ij} = \text{vega}^{BS}(\tau_i, K_{ij}, \sigma_{imp}(\tau_i, K_{ij}))$.

Naturalmente, não existe uma única função custo para o nosso problema de estimação. Entretanto, mesmo que as definições sejam diferentes, usualmente ou são baseadas nas diferenças entre os preços ou na diferença entre as volatilidades implícitas, as mais comuns sendo:

$$L_1(\theta) := \sum_{i=1}^m \sum_{j=1}^{n_i} w_{ij} [V^{mkt}(\tau_i, K_{ij}) - V(\tau_i, K_{ij}; \theta)]^2 \quad (3.4)$$

e

$$L_2(\theta) := \sum_{i=1}^m \sum_{j=1}^{n_i} w_{ij} [\sigma_{imp}^{mkt}(\tau_i, K_{ij}) - \sigma_{imp}(\tau_i, K_{ij}; \theta)]^2. \quad (3.5)$$

Embora a função L_1 seja computacionalmente mais rápida e de implementação mais simples, pois evita a etapa de inversão numérica necessária para obter a volatilidade implícita a cada iteração, optamos por trabalhar com L_2 .

Essa escolha justifica-se principalmente por possuir maior estabilidade numérica, além de conter maior relevância financeira, sendo, frequentemente, considerada o "padrão da indústria".

Observação 3.2. A notação χ^2 .

É comum encontrarmos na literatura de otimização ou estatística a notação χ^2 para a função custo em vez de L ou qualquer outra letra.

Essa escolha está relacionada ao fato de que, sob algumas hipóteses sobre os erros de observação (em particular, normalidade e independência) a soma dos resíduos ao quadrado pode ser interpretada como uma estatística com distribuição do tipo χ^2 .

Com efeito, se $Z_1, \dots, Z_n \sim N(0, 1)$, são variáveis aleatórias independentes, então a soma dos quadrados

$$Q = \sum_i^n Z_i^2$$

segue uma distribuição χ^2 com n graus de liberdade.

Assim, ao considerarmos uma hipótese sobre a *natureza* dos erros, mais especificamente, a de que $\varepsilon_i \sim N(0, \sigma_i^2)$, onde σ_i representa a incerteza sobre a medição de y_i , imbuímos uma interpretação estatística ao problema de otimização, pois agora a função custo, com pesos $w_{ij} = 1/\sigma_i^2$ é, por definição, uma estatística χ^2 , dada por

$$\chi^2 = \sum_{i=1}^n \left(\frac{y_i - f(x_i; \theta)}{\sigma_i} \right)^2,$$

onde $Z_i = \frac{y_i - f(x_i; \theta)}{\sigma_i} \sim N(0, 1)$.

Consequentemente, dentro da hipótese dos erros gaussianos e independentes, a notação χ^2 não somente indica a distribuição da função custo, como também sinaliza que o método de otimização é, na verdade, uma implementação do princípio da *máxima verossimilhança*.

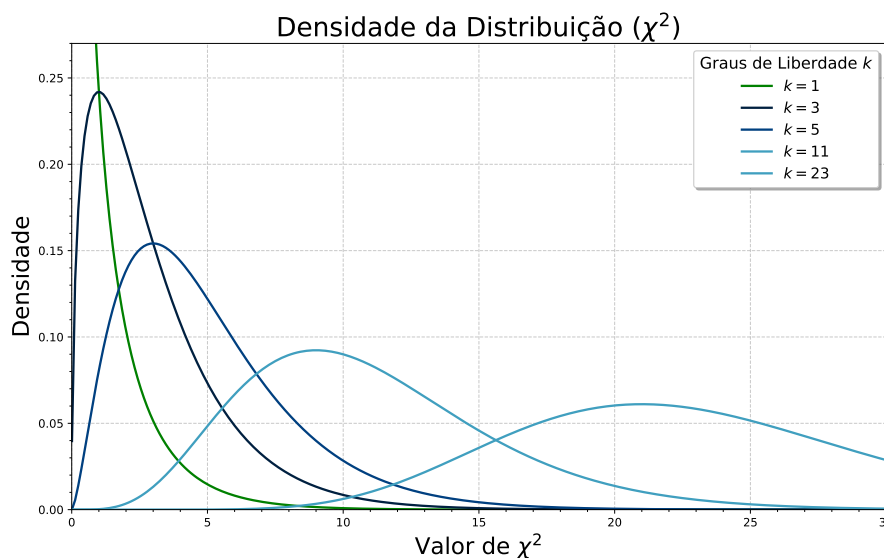


Figura 3.1: Comportamento da distribuição χ^2 conforme o valor de k aumenta.

3.1.2 Implicações da escolha da função custo

Como nossa escolha de função custo foi motivada por sua robustez e estabilidade, é natural esperar que o procedimento de otimização seja penalizado de alguma forma. A escolha de L_2 vem acompanhada do fato de essa ser uma minimização consideravelmente mais lenta, uma vez que, a cada iteração, precisamos primeiro obter o preço teórico $V(\tau_i, K_{ij}; \theta)$ e só então utilizarmos algum algoritmo de cálculo de raízes (optamos pelo Newton-Raphson, que possui convergência quadrática) para extraírmos $\sigma_{imp}(\tau_i, K_{ij}; \theta)$ e, por fim, calcularmos o quadrado da diferença entre os valores teóricos e de mercado.

A escolha da função L_1 mostrou-se inadequada neste contexto. Em primeiro lugar, como os prêmios das opções variam drasticamente, a função objetivo acaba sendo dominada pelos erros dos contratos mais caros, mal condicionando o problema de otimização.

Além disso, o cálculo do gradiente por diferenças finitas, implementado em `scipy.optimize.approx_fprime`, introduz erros numéricos adicionais, inerentes à aproximação discreta das derivadas.

Por fim, o algoritmo de otimização empregado mostrou-se altamente sensível a esse tipo de irregularidade, não sendo capaz de convergir de forma estável diante da combinação de disparidade de escalas e ruído numérico no cálculo do gradiente.

Após diversos testes numéricos, a função custo mais adequada para o nosso conjunto de dados foi:

$$L(\theta) := \sum_{i=1}^m \sum_{j=1}^{n_i} \left[\sigma_{imp}^{mkt}(\tau_i, K_{ij}) - \sigma_{imp}(\tau_i, K_{ij}; \theta) \right]^2. \quad (3.6)$$

3.1.3 Desafios da estimação

Uma das maiores dificuldades no ajuste do modelo de Heston é que o conjunto $H = \{\rho, \gamma, \bar{v}, \kappa, v_0\}$ possui 5 parâmetros a serem estimados cujo efeito no formato da superfície de volatilidade implícita pode ser bastante similar. Por esse motivo, pode acontecer de encontrarmos diversos mínimos locais durante o procedimento de otimização.

Entender como esses parâmetros afetam a volatilidade implícita, como fizemos anteriormente, pode nos dar um bom ponto de partida, ou seja, um bom conjunto inicial de parâmetros.

A situação não é tão ruim quanto parece, tendo em vista que podemos acelerar a determinação dos parâmetros reduzindo o conjunto H a ser otimizado. Ao compararmos o impacto a longo prazo da reversão à média κ e da curvatura γ , vemos que esses dois parâmetros possuem um efeito similar no formato da SVI. É, portanto, prática comum fixarmos κ (com um valor, por exemplo, $\kappa = 0.5$) e otimizarmos γ . Assim, nosso conjunto H reduz-se a 4 parâmetros.

Outro parâmetro que pode ser determinado a priori heurísticamente é o valor inicial da variância, v_0 . Para maturidades muito curtas, isto é, $\tau \rightarrow 0$, é de se esperar que o preço do ativo se comporte de maneira muito próxima ao caso de

Black-Scholes. O impacto da variância estocástica do modelo de Heston deve se reduzir a zero, no limite $\tau \rightarrow 0$.

Para opções com maturidades muito curtas, o processo de preço pode ser aproximado por outro com seguinte forma:

$$dS_t = rS_t dt + \sqrt{v_0} S_t dW_x.$$

Isso sugere que, para v_0 , podemos considerar uma boa aproximação o quadrado da volatilidade implícita ATM de uma opção com maturidade muito curta, $v_0 \approx \sigma_{imp}^2(\tau_{min}, K_{ATM})$, para $\tau \rightarrow 0$. Logo, reduzimos novamente nosso conjunto H .

Outro problema que se apresenta, antes mesmo da etapa de minimização, diz respeito à qualidade dos dados de mercado. Em qualquer contexto empírico, dados reais exigem algum tipo de pré-processamento. No caso de opções, os preços observados podem apresentar ruídos decorrentes, por exemplo, de baixa liquidez, *bid-ask spread*, erros de registro, descompassos temporais entre ativos correlacionados ou violações pontuais das condições de não-arbitragem.

Consequentemente, se eles não forem devidamente tratados antes de serem otimizados, o processo retornará parâmetros instáveis e com pouco significado econômico, que não se adequam a qualquer modelo, independentemente de quão robusta seja a metodologia empregada.

Levando isso em consideração, optamos por seguir algumas recomendações descritas por Bardgett et al. (2019) e remover opções:

- * com vencimento muito curto ou muito longo, evitando possíveis distorções ou opções esparsamente negociadas;
- * com volume de negociações zero, para evitar o uso de preços defasados;
- * ITM, por possuírem menor liquidez e informações sobre a *smile*;
- * e cujo *moneyness* está abaixo de 75% ou acima de 125%, para eliminar *outliers*.

Por fim, vale notar que outra fonte de complexidade na otimização da função custo $L(\theta)$ é que, geralmente, não se trata de um problema convexo. Isso ocorre porque a relação de V ou σ_{imp} com θ é extremamente não linear. Logo, a escolha de um bom algoritmo de otimização não-linear deve ser feita levando em consideração todas as particularidades do modelo. No nosso caso, estamos diante de um problema de regressão não-linear, multidimensional e com restrições.

Vários algoritmos podem ser utilizados aqui, cada um com suas vantagens e desvantagens, tanto de otimização local quanto global. Optamos pela otimização local por serem mais rápidos e mais simples de compreender. Em troca, criamos uma forte dependência de um bom chute inicial para garantirmos a convergência. Existem diversas possibilidades de algoritmos que poderíamos utilizar aqui, alguns deles: Nelder-Mead (busca direta, sem uso de gradientes), Trust Region Reflective (trabalha com uma região de confiança ao redor do ponto atual), SLSQP (projetado para problemas com restrições mais gerais), L-BFGS-B (quasi-Newton).

Entretanto, optamos pelo método do *Gradiente Projetado* (PG), que pode ser visto como uma extensão do *Gradiente Descendente* para lidar com problemas com restrições sobre um conjunto convexo, como as restrições de caixa (*box constraints*).

A escolha desse método foi fundamentada numa série de fatores:

- * sua simplicidade conceitual e de implementação;
- * sua velocidade (possivelmente mais rápido que o Nelder-Mead, que também é conceitualmente simples);
- * sua estabilidade numérica;
- * e por lidar nativamente com restrições, como as presentes no modelo de Heston:
 - restrições de caixa: $-1 \leq \rho \leq 1, \gamma, \bar{v}, \kappa, v_0 > 0$
 - condição de Feller: $2\kappa\bar{v} \geq \gamma^2$.

Apesar de ser um algoritmo pensado para otimização convexa, Lan, Li e Xu (2024) forneceram uma nova e rigorosa análise sobre os limites de complexidade para esse método em problemas não convexos, mostrando ser um algoritmo adequado para a situação.

Além disso, os autores forneceram uma melhoria significativa no método ao propor o *Auto-conditioned projected gradient* (AC-PG), que elimina a necessidade de um tamanho de passo fixo - um parâmetro difícil de ser ajustado - ao estimá-lo dinamicamente a cada iteração com base na curvatura local da função objetivo.

3.1.4 Otimização não convexa e o algoritmo AC-PG

A ideia por trás do método é muito simples: realizar um "passo de gradiente", projetá-lo - isto é, procurar o ponto mais próximo no conjunto de parâmetros viáveis (*feasible set*) - para satisfazer as restrições e, por fim, o tamanho do próximo passo é refinado com as informações atuais. Para que a projeção possa ser feita efetivamente, é preciso que essas restrições sejam simples (é o nosso caso, com exceção da condição de Feller).

Formalmente, considere o seguinte problema geral de otimização, possivelmente não-convexo:

$$L^* = \min_{\theta \in \Theta} L(\theta), \quad (3.7)$$

onde $\Theta \subset \mathbb{R}^n$ é um conjunto compacto e convexo, e $L : \Theta \rightarrow \mathbb{R}$ é uma função continuamente diferenciável, mas possivelmente não convexa. Suponha também que L seja suave, com gradiente Lipschitz contínuo, isto é, para algum $M \geq 0$,

$$\|\nabla L(x) - \nabla L(y)\| \leq M\|x - y\|, \quad \forall x, y \in \Theta. \quad (3.8)$$

Temos, assim, que M é o limite superior da curvatura da função L :

$$L(x) - L(y) - \langle \nabla L(y), x - y \rangle \leq \frac{M}{2}\|x - y\|^2, \quad \forall x, y \in \Theta. \quad (3.9)$$

Adicionalmente, assumimos que L é fracamente convexa com um parâmetro de curvatura inferior $m \in [0, M]$, isto é,

$$L(x) - L(y) - \langle \nabla L(y), x - y \rangle \geq -\frac{m}{2}\|x - y\|^2, \quad \forall x, y \in \Theta. \quad (3.10)$$

O algoritmo PG pode ser resumido em seu passo de atualização, dado por:

$$\theta_t = \arg \min_{\theta \in \Theta} \left\{ \langle \nabla L(\theta_{t-1}), \theta \rangle + \frac{1}{2\alpha_t} \|\theta_{t-1} - \theta\|^2 \right\}, \quad (3.11)$$

onde α_t é o tamanho do passo na iteração t .

Vale notar que a expressão de θ_t em (3.11) é equivalente ao passo do gradiente seguido da projeção no conjunto possível Θ . Para enxergar isso, basta realizar algumas manipulações algébricas, expandindo o termo quadrático, agrupando valores, completando quadrados, etc. O resultado é a definição de projeção euclidiana de um ponto sobre Θ .

Finalmente, consideraremos uma variante do PG, chamado *Auto-conditioned projected gradient*, que não requer nenhum conhecimento prévio acerca das curvaturas M e m . Além disso, é um algoritmo livre de qualquer procedimento de busca linear ou *backtracking*. O método é dado por:

Entrada: Ponto inicial $\theta_0 \in \Theta$ e estimativa $M_0 > 0$

1 para $t \leftarrow 1$ até k faça

2 Definir α_t tal que $\alpha_t^{-1} = \max\{M_0, \dots, M_{t-1}\}$

3 $\theta_t \leftarrow \arg \min_{\theta \in \Theta} \left\{ \langle \nabla L(\theta_{t-1}), \theta \rangle + \frac{1}{2\alpha_t} \|\theta_{t-1} - \theta\|^2 \right\}$

4 $M_t \leftarrow \frac{2(L(\theta_t) - L(\theta_{t-1}) - \langle \nabla L(\theta_{t-1}), \theta_t - \theta_{t-1} \rangle)}{\|\theta_t - \theta_{t-1}\|^2}$

5 fim

Algorithm 1: Método AC-PG

O teorema a seguir estabelece a convergência do algoritmo AC-PG.

Teorema 3.1. *Sejam $\{\theta_t\}$ a sequência gerada pelo algoritmo 1 e considere $q \leq \lfloor \log_{3/2}(M/M_0) \rfloor + 1$ o número de incrementos na estimativa da constante de Lipschitz. Sob essas hipóteses, tem-se:*

$$\min_{t \in [k]} \|g_{\Theta,t}\|^2 \leq \frac{48\alpha_k^{-2} D_{\Theta}^2}{(k-2q)^2} + \frac{12\alpha_k^{-1} D_{\Theta}^2 \cdot \left(\sum_{t=0}^{k-1} \max\{0, m_t\} \right)}{(k-2q)^2}, \quad (3.12)$$

onde denotamos $g_{\Theta,t} := \alpha_t^{-1}(\theta_{t-1} - \theta_t)$ o gradiente projetado, D_{Θ} o diâmetro do conjunto Θ e $m_t \leq m$ o parâmetro de curvatura inferior local.

Observação 3.3. A escolha de M_0 .

É importante notar também que, no algoritmo acima, a quantidade M_0 pode ser determinada simplesmente escolhendo dois pontos em Θ e calculando o fator de suavidade local usando a mesma fórmula de M_t . Nas iterações seguintes ele será adaptado sequencialmente.

3.2 Metodologia

3.2.1 Descrição dos dados

Os dados utilizados na minimização da função custo foram obtidos no dia 28/08/2025, de maneira gratuita, pelo site <<https://opcoes.net.br/>>. São plani-

lhas *xlsx* com cotações de opções de compra e venda de ações do Banco do Brasil (BBAS3), Banco Bradesco (BBDC4), Itaú Unibanco (ITUB4), Petrobras (PETR4) e Vale (VALE3).

Cada linha do conjunto de dados corresponde a uma opção e inclui as seguintes variáveis principais:

- * **Ticker:** código de negociação da opção;
- * **Vencimento:** data de vencimento do contrato;
- * **Dias úteis até o vencimento:** tempo restante até o vencimento, em dias úteis;
- * **Tipo:** indica se a opção é uma *call* ou uma *put*;
- * **Mod.:** modalidade de exercício (americana ou europeia)
- * **Strike:** preço de exercício da opção;
- * **A/I/OTM:** classificação do *moneyness* (ITM, ATM ou OTM);
- * **Distância percentual do strike:** diferença percentual entre o preço de exercício e o preço à vista da ação;
- * **Último preço:** preço mais recente negociado no mercado;
- * **Volume, Negócios, Abertura, Fechamento, Máxima e Mínima:** indicadores de negociação do ativo derivativo;
- * **Delta, Gamma, Theta (\$), Theta (%) e Vega:** gregas calculadas pelo mercado para cada opção;
- * **IQ:** indicador de liquidez;
- * **Coberto, Travado, Descoberto, Titulares e Lançadores:** posições agregadas do mercado associadas a cada contrato.

Os dados contemplam opções com diferentes prazos até o vencimento, variando aproximadamente 16 a 60 dias úteis e com preços de exercício distribuídos em torno do *spot no momento da coleta*.

3.2.2 Limpeza dos dados

Como esperado, os dados, da maneira como vieram, não estavam adequados ao procedimento. Para garantir a qualidade das informações que seriam utilizadas na otimização, foi desenvolvida uma função de limpeza, chamada *calibration_cleaning*, que realiza uma série de transformações, filtragens e padronizações dos dados, retornando um *DataFrame* apenas com as observações adequadas para o processo de estimação, o que reduziu drasticamente o tamanho do conjunto utilizado.

A seguir, um resumo das principais etapas do procedimento de limpeza:

1. **Seleção das colunas relevantes:** o arquivo original é muito carregado e a maioria das informações são irrelevantes para a estimação. Apenas as colunas de interesse foram mantidas e renomeadas.

São elas: K , τ (dias), V_{mkt} , volatilidade implícita, vega, tipo, trades, mod.

Essas variáveis representam, respectivamente, o preço de exercício, o tempo até o vencimento, o preço de mercado da opção, a volatilidade implícita, o vega, o tipo (se é *call* ou *put*), o número de negociações e a modalidade de exercício (americana ou europeia).

2. **Padronização das unidades:** algumas variáveis precisaram ser reescaladas para manter a consistência das magnitudes:

- * K foi dividido por 100;
- * volatilidade implícita foi dividida por 1000;
- * vega foi dividido por 10000.

3. **Filtragem:** o procedimento de filtragem das opções descrito em 3.1.3 foi aplicado da seguinte maneira:

- * Primeiramente, selecionamos apenas as opções europeias no conjunto de dados (coluna *mod* = 'E'), uma vez que o modelo de Heston é formulado sob essa hipótese;
- * Definimos o mínimo de negociações aceitas: $\text{min_trades} = 10$;
- * Estipulamos como horizonte temporal o prazo entre 21 e 504 dias úteis;
- * Estabelecemos o preço mínimo de uma opção por $\text{min_price} = 0.01$;
- * Para cada opção, calculamos seu *moneyness*, $m = S_0/K$, e mantivemos no *dataset* apenas *calls* com $0.75 \leq m \leq 1$ e *puts* com $1 \leq m \leq 1.25$.

4. **Adição de colunas:**

- * O tempo até o vencimento foi convertido para anos, utilizando o padrão de 252 dias úteis por ano. Essa medida foi necessária pois "ao ano (a.a.)" é a unidade padrão utilizada no mercado;
- * Para podermos estimar apenas em termos de *calls*, as opções de venda tiveram seus preços equivalentes em opção de compra calculados através da *paridade put-call* (1.54), utilizando o vencimento τ em anos.

3.2.3 Estratégia de estimação

O procedimento de estimação de parâmetros foi estruturado em duas etapas principais:

1. Pré-processamento dos dados de mercado;
2. Otimização dos parâmetros do modelo.

Vamos detalhar cada um dos passos.

Etapa 1: Para cada ativo, os dados foram limpos e organizados por meio da função `calibration_cleaning`, eliminando contratos com inconsistências de preços, liquidez insuficiente ou ausência de informações relevantes.

Em seguida, o valor inicial da variância, v_0 , foi estimado a partir das volatilidades implícitas observadas, como descrito em 3.1.3. Para essa tarefa, definimos a função `v0_estimator`.

Com a função `df_converter`, os dados processados foram convertidos em um formato mais simples de se percorrer durante as iterações: um dicionário de listas, onde cada lista representa uma opção.

Etapa 2: A otimização foi conduzida separadamente, um ativo por vez. A estratégia de busca de parâmetros consistiu em um *grid search* sobre diferentes valores do parâmetro κ , que não estava sendo minimizado. Foram testados os seguintes valores:

$$\kappa \in \{0.15, 0.5, 1, 1.5, 2, 3, 5, 7\}.$$

Para cada valor de κ , os demais parâmetros (\bar{v} , γ e ρ) foram estimados a partir do algoritmo 1.

Durante todo o procedimento de estimação, tomamos como vetor inicial de parâmetros

$$\theta_0 = (\bar{v}_0, \gamma_0, \rho_0) = (0.08, 0.1, -0.8).$$

Além de θ_0 , outras configurações constantes foram feitas: definimos três dicionários, `col_loss`, `otim` e `funcs` contendo, respectivamente, as colunas necessárias como argumento de `df_converter`, os parâmetros da otimização (como valor de M_0 , número de iterações e tolerância) e as funções que são chamadas dentro da implementação do algoritmo 1.

O algoritmo AC-PG foi implementado na função `ac_pg` e, dentro dele, foram chamadas as diversas funções desenvolvidas especificamente para este trabalho: `loss_vol` (que implementa nossa função custo (3.6)), `approx_grad` (que calcula o gradiente numérico de L) e `proj` (que projeta os parâmetros de volta dentro dos limites pré estabelecidos, garantindo que: $\bar{v} \in [0.08, 5]$, $\gamma \in [0.1, 7]$ e $\rho \in [-0.999, 0.999]$).

A condição de Feller (2.12) foi verificada logo após a etapa de projeção, ajustando as possibilidades de parâmetros que a violassem para o máximo permitido, garantindo que apenas parâmetros que satisfizessem a condição fossem utilizados nos cálculos subsequentes. Mais especificamente, caso a condição fosse violada ($2\kappa\bar{v} < \gamma^2$), o valor de γ era ajustado para o seu limite máximo permitido, $\sqrt{2\kappa\bar{v}}$, antes de prosseguir com a iteração.

Essa projeção garante a consistência matemática dos parâmetros, mas, como veremos, cria uma dependência direta entre κ , \bar{v} e o máximo que γ pode assumir, influenciando forte e diretamente os resultados da minimização.

Diversos testes numéricos mostraram que uma tolerância de 10^{-4} sobre a variação do valor da função custo ou um limite de 200 iterações eram critérios de parada mais que suficientes para nosso conjunto de dados.

Ao final da otimização, para cada ativo obtivemos um conjunto de parâmetros que minimizou o erro total dentro do *grid* de valores κ . O resultado foi armaze-

nado numa tabela resumo contendo, para cada ativo, o respectivo *ticker*, os valores de κ , o erro total e as estimações de \bar{v} , γ e ρ .

3.3 Resultados

Utilizando a metodologia descrita anteriormente, foram estimados os parâmetros do modelo de Heston para os cinco ativos selecionados. A tabela 3.1 apresenta os resultados ótimos encontrados para cada ativo, apresentando o valor de κ que minimizou a função custo (erro) e o valor de v_0 estimado para o ativo, juntamente com os parâmetros \bar{v} , γ e ρ correspondentes.

ticker	v_0	κ	erro	\bar{v}	γ	ρ
BBAS3	0.1116	3.0	0.011985	0.098573	0.769048	-0.762411
BBDC4	0.0520	1.0	0.002041	0.320517	0.800646	-0.343435
ITUB4	0.0488	1.0	0.000831	0.080000	0.400000	-0.886510
PETR4	0.1391	7.0	0.243781	0.080000	0.993631	0.334658
VALE3	0.0420	3.0	0.003968	0.080000	0.692820	-0.420339

Tabela 3.1: Resumo da minimização: parâmetros ótimos por ativo, com sua respectiva variância inicial estimada v_0 .

Análise dos parâmetros

Os resultados individuais do processo de otimização podem ser encontrados nas tabelas 3.2 - 3.6. De modo geral, observa-se consistência entre os parâmetros estimados e o comportamento esperado do modelo, sem valores extremos ou inconsistentes, à exceção de PETR4, que apresentou um comportamento atípico.

Começando com a velocidade de reversão à média. Os ativos BBAS3 e VALE3 apresentaram valores ótimos em torno de $\kappa = 3$, indicando uma reversão relativamente rápida da volatilidade em direção ao nível de longo prazo \bar{v} . Já para BBDC4 e ITUB4, o menor valor da função custo foi obtido com $\kappa = 1$, sugerindo uma dinâmica da variância mais persistente e com maior liberdade de flutuação ao redor da média de longo prazo.

O parâmetro \bar{v} , associado à variância de longo prazo, permaneceu fixado no valor 0,08 para ITUB4 e VALE3. Esse comportamento decorre da restrição de caixa imposta ao modelo, que, ao forçar \bar{v} ao seu limite inferior admissível, permitiu maior flexibilidade no ajuste dos parâmetros responsáveis pela curvatura e inclinação da superfície de volatilidade.

Quanto a γ , que controla a volatilidade da volatilidade, os valores estimados situaram-se entre 0,4 e 0,8 para BBAS3, BBDC4, ITUB4 e VALE3. Esses valores mostraram-se adequados para reproduzir *smiles* de volatilidade moderadamente acentuados, sem introduzir instabilidades numéricas relevantes durante o processo de otimização.

Por fim, a correlação ρ foi negativa para todos os ativos que apresentaram convergência satisfatória, confirmando a presença do efeito de alavancagem. A

magnitude dessa correlação variou entre os ativos, com valores mais pronunciados para BBAS3 e ITUB4 e mais moderados para VALE3, sugerindo que o impacto da alavancagem difere substancialmente mesmo entre ativos pertencentes a um mesmo setor econômico.

O ativo PETR4 apresentou comportamento atípico ao longo do processo de otimização. O menor valor da função custo foi obtido para $\kappa = 7$, ainda assim associado a um erro elevado, da ordem de 10^{-1} (aproximadamente 24%), indicando um ajuste ruim. A análise da Tabela 3.5 sugere que valores ainda maiores de κ poderiam reduzir o erro, mas tais valores implicariam meias-vidas excessivamente curtas, com pouco significado financeiro, razão pela qual foram excluídos pela metodologia adotada.

Além disso, os parâmetros \bar{v} e γ atingiram seus limites admissíveis, com $\bar{v} = 0.08$ e $\gamma = 0.999$, indicando que o algoritmo buscou compensar a dificuldade de ajuste por meio de parâmetros extremos. Esse comportamento evidencia a sensibilidade do procedimento à qualidade dos dados de mercado e delimita os limites práticos da metodologia empregada no contexto desse ativo específico.

Em suma, a análise dos parâmetros estimados indica que o modelo foi capaz de capturar, de forma coerente, características relevantes das dinâmicas de volatilidade para a maior parte dos ativos considerados, com valores economicamente plausíveis e boa qualidade de ajuste. As diferenças observadas nos parâmetros refletem tanto particularidades intrínsecas de cada ativo quanto restrições impostas pela metodologia de otimização e pelos dados disponíveis. O desempenho atípico verificado em PETR4 reforça a sensibilidade do procedimento à qualidade dos dados de mercado e evidencia os limites práticos do modelo e da abordagem de estimação adotada, sem comprometer, contudo, a consistência dos resultados obtidos para os demais ativos.

κ	erro	\bar{v}	γ	ρ
0.15	0.038493	0.159993	0.219084	-0.999000
0.50	0.029130	0.087061	0.295061	-0.999000
1.00	0.020222	0.080995	0.402481	-0.999000
1.50	0.016052	0.088443	0.515100	-0.985913
2.00	0.014429	0.093429	0.611324	-0.874857
3.00	0.011985	0.098573	0.769048	-0.762411
5.00	0.013568	0.102912	0.708548	-0.915999
7.00	0.014023	0.104692	0.699088	-0.980934

Tabela 3.2: Resultados da minimização para BBAS3 ($v_0 = 0.1116$)

κ	erro	\bar{v}	γ	ρ
0.15	0.002599	1.664328	0.706610	-0.347662
0.50	0.002213	0.575281	0.758473	-0.346154
1.00	0.002041	0.320517	0.800646	-0.343435
1.50	0.004731	0.241300	0.537511	-0.421183
2.00	0.007447	0.205697	0.266017	-0.638540
3.00	0.007697	0.160197	0.217149	-0.737322
5.00	0.007675	0.123400	0.204644	-0.796642
7.00	0.007612	0.107604	0.198329	-0.809505

Tabela 3.3: Resultados da minimização para BBDC4 ($v_0 = 0.0520$)

κ	erro	\bar{v}	γ	ρ
0.15	0.001533	0.247830	0.272670	-0.999000
0.50	0.001192	0.092277	0.303772	-0.999000
1.00	0.000831	0.080000	0.400000	-0.886510
1.50	0.000876	0.080000	0.482270	-0.749205
2.00	0.001287	0.080000	0.441137	-0.794155
3.00	0.002142	0.080000	0.389914	-0.828743
5.00	0.003384	0.080000	0.495819	-0.568573
7.00	0.005613	0.080000	0.380825	-0.627005

Tabela 3.4: Resultados da minimização para ITUB4 ($v_0 = 0.0488$)

κ	erro	\bar{v}	γ	ρ
0.15	0.459304	0.080000	0.154919	0.999000
0.50	0.432523	0.080000	0.282843	0.999000
1.00	0.402619	0.080000	0.400000	0.999000
1.50	0.377889	0.080000	0.489898	0.999000
2.00	0.356899	0.080000	0.565685	0.999000
3.00	0.322809	0.080000	0.692820	0.892704
5.00	0.271649	0.080000	0.894427	0.717995
7.00	0.243781	0.080000	0.993631	0.334658

Tabela 3.5: Resultados da minimização para PETR4 ($v_0 = 0.1391$)

Análise dos ajustes

Apesar da otimização ter gerado erros pequenos, isso não necessariamente garantiu um bom ajuste para os ativos, como observa-se nas figuras 3.2 - 3.4. Os gráficos mostram as smiles de mercado e os ajustes realizados pelos parâmetros ótimos obtidos na minimização da função custo, permitindo comparar o formato e a curvatura de ambas as curvas. Essa comparação possibilita avaliarmos a capacidade do modelo de capturar a forma da *smile* observada no mercado, que é nosso objetivo principal.

κ	erro	\bar{v}	γ	ρ
0.15	0.005926	0.405138	0.348628	-0.713604
0.50	0.005331	0.150305	0.387692	-0.669014
1.00	0.004643	0.096319	0.438905	-0.623465
1.50	0.004060	0.080000	0.489898	-0.586692
2.00	0.004481	0.080000	0.507846	-0.648257
3.00	0.003968	0.080000	0.692820	-0.420339
5.00	0.007479	0.080000	0.894427	-0.294550
7.00	0.012228	0.080000	1.058301	-0.208665

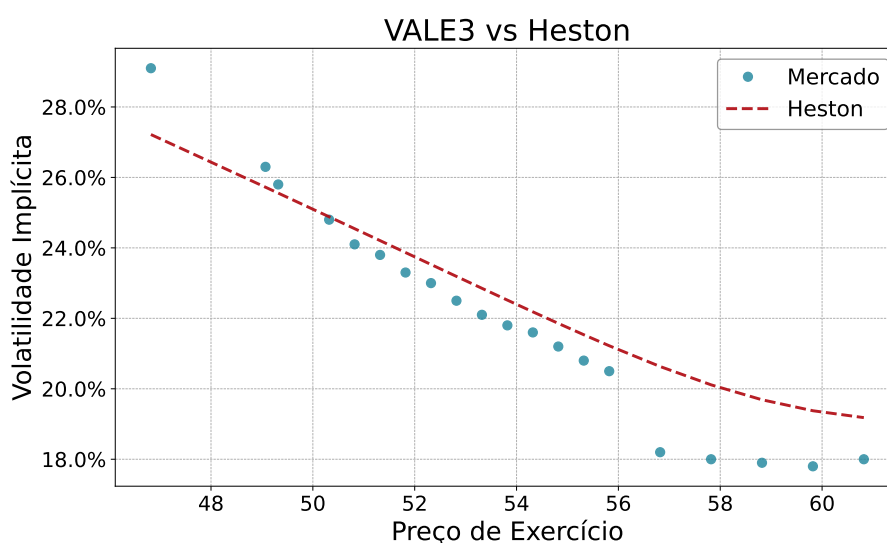
Tabela 3.6: Resultados da minimização para VALE3 ($v_0 = 0.0420$)

Figura 3.2: Ajuste do modelo de Heston para os dados de VALE3, com vencimento em aproximadamente 36 dias.

Pode-se afirmar que VALE3 apresentou o ajuste visual mais consistente dentre os expostos. O maior problema com BBAS3 foi a pouca curvatura presente na curva gerada pelo Heston, apesar de a inclinação estar bem alinhada com os dados do ativo. Em BBDC4 e ITUB4 percebe-se distâncias muito pequenas entre o modelo e os dados. Essa dispersão é, contudo, enganosa pois decorre, em grande parte, da escassez de dados após a aplicação de todos os filtros nos *datasets* originais.

Evidentemente, PETR4 representou um caso de falha de ajuste. O erro numérico de 24% é visualmente confirmado pela figura 3.4, que ilustra a curva do modelo de Heston não conseguindo reproduzir o comportamento do mercado. Isso é uma consequência direta dos parâmetros incoerentes resultantes do procedimento, notadamente o sinal positivo de ρ . Essa divergência reforça que o método é sensível a dados ruidosos e a superfícies pouco informativas.

Assim sendo, os resultados obtidos indicam que a metodologia atual é adequada, até certo ponto, para superfícies bem comportadas e com dados abundantes, mas perde eficiência diante de dados muito dispersos ou com muito ruído.

Em outras palavras, ativos com superfícies mais completas e estáveis, como BBAS3 e VALE3, apresentaram melhor aderência ao modelo, enquanto aqueles com menor liquidez e maior ruído, como BBDC4, ITUB4 e PETR4, realçaram as limitações práticas do modelo de Heston.

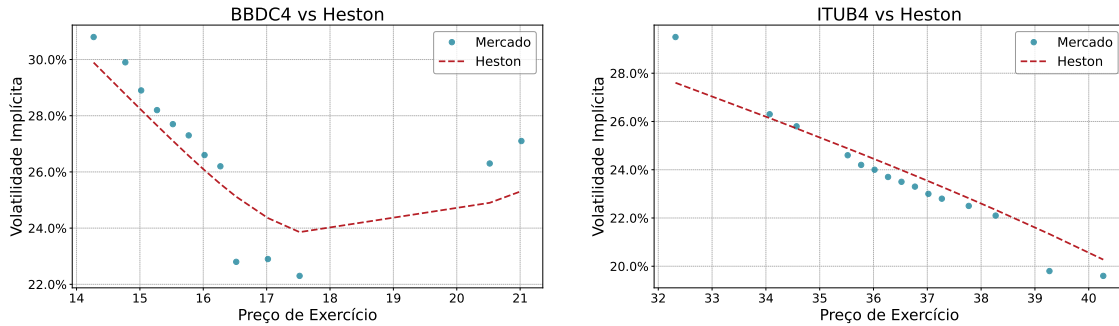


Figura 3.3: Ajuste do modelo de Heston para os dados de BBDC4 e ITUB4, com vencimento em aproximadamente 36 dias.

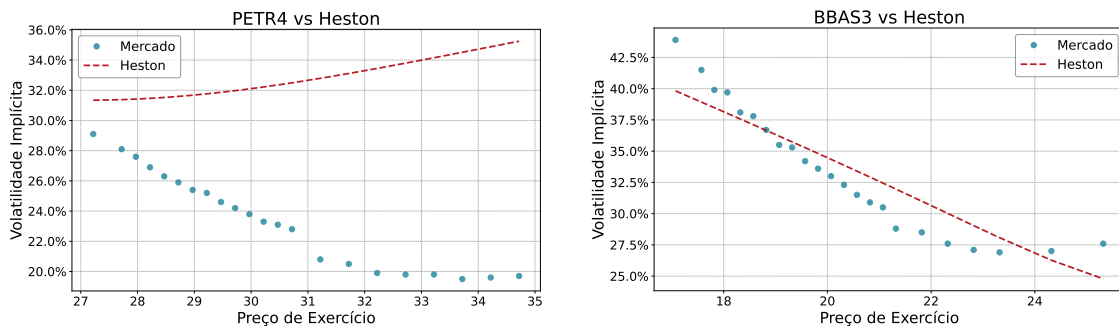


Figura 3.4: Ajuste do modelo de Heston para os dados de PETR4 e BBAS3, com vencimento em aproximadamente 36 dias.

3.3.1 Verificação com *Multi-Start*

Uma pergunta natural surge ao avaliarmos a metodologia adotada para minimizar a função custo: e se a busca realizada estiver presa em mínimos locais? Tal possibilidade é plausível, pois, como vimos, algoritmos de otimização local dependem fortemente do ponto inicial da busca (que mantivemos constante para todos os dados ao longo do procedimento de estimação).

Para investigar essa hipótese, realizamos o seguinte experimento numérico, conhecido como **otimização *multi-start***:

1. Fixamos um valor de κ para os dados;
2. Para cada ativo, executamos o loop de otimização `n_starts = 20` vezes;
3. Em cada iteração, um vetor inicial foi gerado aleatoriamente, com limites ligeiramente mais restritos nos parâmetros, a fim de garantir estabilidade numérica, utilizando `np.random.uniform`;
4. O conjunto de parâmetros que produziu o menor erro foi selecionado como o resultado ótimo correspondente ao κ fixado.

Primeiramente, conduzimos os testes nos κ ótimos encontrados na seção anterior: $\kappa = 1$ (para BBDC4 e ITUB4) e $\kappa = 3$ (para BBAS3 e VALE3). Os resultados são apresentados nas tabelas 3.7 e 3.8.

Para manter a seção concisa, apresentamos aqui apenas os resultados consolidados de cada ativo. As tabelas completas do procedimento encontram-se no Apêndice B, contendo informações detalhadas dos parâmetros e erros obtidos em cada iteração para todos os ativos analisados.

ticker	κ	erro	\bar{v}	γ	ρ
BBDC4	1.0	0.002041	0.320521	0.800651	-0.343581
ITUB4	1.0	0.000831	0.080000	0.400000	-0.886568

Tabela 3.7: Resultados do teste *multi-start* para $\kappa = 1.0$ (BBDC4 e ITUB4).

ticker	κ	erro	\bar{v}	γ	ρ
BBAS3	3.0	0.011984	0.098595	0.769135	-0.758754
VALE3	3.0	0.003968	0.080000	0.692820	-0.418450

Tabela 3.8: Resultados do teste *multi-start* para $\kappa = 3.0$ (BBAS3 e VALE3).

Comparando essas tabelas com a tabela resumo 3.1, observa-se que o procedimento *multi-start* confirmou os resultados obtidos na busca em grade com vetor inicial fixo, tanto os erros quanto os parâmetros ótimos, validando a consistência e estabilidade locais da estimação e sugerindo que o *grid search* foi suficientemente preciso na identificação dos mínimos estáveis, mesmo diante de diferentes pontos iniciais.

Porém, é importante ressaltar que esse comportamento não pode ser generalizado: em outros testes - especialmente para PETR4 - o método mostrou-se sensível à escolha do ponto inicial, produzindo erros significativamente maiores quando o vetor θ_0 se afastava muito do valor utilizado na otimização principal.

Por fim, realizamos alguns testes adicionais para PETR4, com o intuito de verificar se o problema estava relacionado aos pontos iniciais escolhidos ou à inadequação do modelo para descrever os dados. A tabela 3.9 mostra que trata-se da segunda opção: o modelo de Heston simplesmente não conseguiu representar a superfície para esse ativo.

Esses resultados reforçam a necessidade de aprimorar o procedimento de estimação, tema que discutiremos a seguir.

κ	erro	\bar{v}	γ	ρ
0.5	0.432523	0.08	0.282843	0.999000
1.0	0.402619	0.08	0.400000	0.999000
3.0	0.322809	0.08	0.692820	0.892709
7.0	0.235360	0.08	1.058301	0.614394

Tabela 3.9: Resultados do teste *multi-start* para PETR4.

Observação 3.4. Sobre a qualidade dos dados e o objetivo deste trabalho.

É importante destacar que parte das limitações refletidas nos resultados pode ser atribuída à qualidade e ao tamanho do conjunto de dados que foi utilizado no procedimento de minimização. Após a aplicação de diversos filtros, todos os *datasets* tornaram-se reduzidos, o que, embora tenha diminuído os erros numéricos da otimização, comprometeu a riqueza de informações das superfícies analisadas e, consequentemente, o ajuste do modelo.

Ressaltamos também que o objetivo principal deste texto não foi alcançar o melhor ajuste possível para o modelo, mas sim compreender o processo de estimação de parâmetros, o papel da volatilidade implícita no mercado, o comportamento do modelo de Heston e suas limitações diante de dados reais do mercado brasileiro.

3.4 Melhorando a performance

A seguir, discutimos possíveis aprimoramentos no processo de determinação de parâmetros do modelo de Heston, com o objetivo de aumentar a velocidade, a eficiência e a robustez numérica do procedimento. As soluções apresentadas são comentadas de forma sucinta, juntamente com suas principais limitações e desafios de implementação.

3.4.1 Preços

O cálculo dos preços no modelo de Heston, tal como foi implementado, torna-se ineficiente conforme avançamos no *loop* de otimização. Isso ocorre porque cada iteração requer diversas integrações numéricas, conforme o esquema a seguir:

- * Para calcular o preço de uma opção, duas integrais são avaliadas numericamente via *Quadratura Adaptativa*, utilizando `scipy.integrate.quad`;
- * Para cada iteração, a função custo calcula a diferença entre o preço de mercado e o preço teórico de todas as opções no conjunto de dados. Assim, para n opções, são computadas $2n$ integrais;
- * Se o procedimento de minimização tiver m iterações, o total de integrações numéricas sobe para $2nm$, representando um esforço computacional considerável.

Para contornar essa limitação, algumas técnicas se destacam na literatura, como a *Transformada Rápida de Fourier* (FFT) e o *Método COS*.

O FFT é um algoritmo altamente eficiente para calcular transformadas discretas de Fourier e suas inversas. Carr e Madan (1999) mostraram que, quando a função característica dos retornos é analiticamente conhecida, os preços de opções podem ser obtidos por meio dessa técnica. A principal desvantagem é a necessidade de reformular a estrutura de precificação, o que alteraria significativamente o desenvolvimento teórico adotado neste trabalho.

Já o método COS, proposto por Fang e Oosterlee (2008), aproxima a densidade do log-preço do ativo no vencimento por uma expansão em cossenos de Fourier, resultando em um ganho expressivo de performance, pois substitui a integração numérica por uma soma finita sobre termos da série. Mais detalhes a respeito do método podem ser encontrados em Oosterlee e Grzelak (2020). A principal desvantagem é a necessidade de um bom domínio da teoria de séries de Fourier e da escolha adequada do intervalo de truncamento da densidade.

3.4.2 Otimização

O processo de otimização também apresenta gargalos de performance. Isso decorre da combinação de gradientes numéricos, obtidos a partir de dados ruidosos, com o uso de um otimizador local sensível a esse ruído. A seguir, discutimos algumas alternativas para tornar o processo de estimação mais eficiente e robusto.

Primeiramente, o método das diferenças finitas é computacionalmente caro. A maneira mais eficiente de acelerar essa etapa seria através do **cálculo exato do gradiente** da função custo. A melhoria seria substancial, tanto na velocidade quanto na estabilidade numérica, uma vez que eliminaríamos os ruídos decorrentes das aproximações. Entretanto, sua implementação é muito desafiadora e suscetível a erros, uma vez que nossa função custo é baseada na volatilidade implícita $\sigma_{imp}(\theta)$. Assim, para calcular o gradiente $\nabla L(\theta)$, seria necessário aplicarmos ferramentas avançadas, como os **métodos adjuntos** (*adjoint methods*), conforme descrito em Clevenhaus et al. (2023).

Resolvido o problema do gradiente numérico, podemos utilizar um algoritmo de otimização local mais potente como o **L-BFGS-B**, um método *quasi-Newton* extremamente rápido e eficiente, pois utiliza o histórico de gradientes para construir uma aproximação da Hessiana, convergindo em poucas iterações.

Entretanto, a possibilidade de convergência para mínimos locais permanece. Assim, propõe-se uma *estratégia híbrida*, que combine a eficiência dos métodos

locais com a capacidade exploratória dos globais. Um exemplo é o algoritmo Dual Annealing, disponível em `scipy.optimize.dual_annealing`, que integra a busca global do Simulated Annealing com um refinamento local interno. Esse algoritmo pode, inclusive, incorporar o gradiente analítico obtido por métodos adjuntos, tornando-se simultaneamente rápido e preciso, resolvendo o problema do ruído e acelerando a convergência.

Dessa forma, a combinação de melhorias no cálculo dos preços e no processo de otimização tende a produzir estimações consideravelmente mais rápidas, eficientes e numericamente robustas.

3.5 Limitações do modelo de Heston

Como discutido nas seções anteriores, o ajuste do modelo de Heston apresenta dificuldades numéricas relevantes. Além dessas, há limitações estruturais nesse modelo, amplamente reconhecidas na literatura, que restringem sua aplicabilidade conforme o tipo de dado e o fenômeno que se deseja reproduzir. Mencionamos algumas delas a seguir.

- * **Processo de variância:** O processo de Cox–Ingersoll–Ross, utilizado no Heston para modelar a variância, tende a produzir trajetórias demasiadamente suaves em comparação aos dados empíricos. Evidências recentes - Gatheral, Jaisson e Rosenbaum (2018) - indicam que a volatilidade de mercado exibe comportamento mais irregular, caracterizado por propriedades de “rugosidade” que o modelo de Heston não reproduz.
- * **Ausência de saltos:** O modelo de Heston padrão não captura saltos, que ocorrem frequentemente nos retornos de ativos financeiros. Consequentemente, o *smile* gerado tende a dificultar a reprodução das superfícies observadas no mercado, especialmente para maturidades muito curtas. Gatheral (2006) discute em detalhes a importância de incorporar saltos a modelos de volatilidade e o impacto disso no formato da superfície.
- * **Parâmetros constantes:** O modelo de Heston possui uma fórmula fechada somente sob a hipótese de parâmetros constantes (ou constantes por partes). Na prática, entretanto, esses parâmetros variam no tempo, o que reduz a capacidade do modelo de reproduzir superfícies dinâmicas de volatilidade. Benhamou, Gobet e Miri (2010) propõem uma generalização com parâmetros dependentes do tempo e derivam uma formulação analítica aproximada, com ganhos expressivos em precisão e velocidade de otimização.
- * **Instabilidade na otimização:** Como resultado, a minimização pode ser instável, resultando em parâmetros pouco razoáveis, além de ser computacionalmente custosa, mesmo com as melhorias de performance discutidas neste trabalho. Ademais, o procedimento de estimação costuma ser diário, e os valores ótimos podem variar significativamente entre dias, dependendo das condições iniciais e dos mínimos locais encontrados, como relatam Guillaume e Schoutens (2010)).

Portanto, embora o modelo de Heston represente um avanço em relação ao de Black–Scholes, suas hipóteses fundamentais ainda são insuficientes para capturar com precisão e eficiência a complexidade do mercado financeiro, como demonstrado neste trabalho. Essas limitações motivaram o desenvolvimento de outros modelos, além de generalizações do Heston, que buscam representar de maneira mais realista o comportamento dos preços e da volatilidade observados empiricamente.

3.6 Alternativas ao modelo de Heston

O desenvolvimento de modelos alternativos de volatilidade não necessariamente seguiu a mesma abordagem do modelo de Heston. Algumas extensões introduzem saltos no processo de preço, outras modificam a dinâmica da variância ou consideram taxas de juros estocásticas. A seguir, apresentamos alguns modelos amplamente utilizados na literatura e na prática de mercado.

Bates

A generalização do modelo de Heston proposta por Bates (1996) incorpora saltos no processo de preço S_t . Esse modelo é descrito pelo seguinte sistema de EDEs:

$$\begin{cases} dS_t/S_t = (r - \lambda \mathbb{E}[e^J - 1])dt + \sqrt{v_t}dW_x + (e^J - 1)dN_t, \\ dv_t = \kappa(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dW_v, \end{cases}$$

onde N_t é um processo de Poisson com intensidade λ e $J \sim N(\mu_J, \sigma_J^2)$ representa o tamanho dos saltos. O processo N_t é independente tanto dos movimentos brownianos quanto dos tamanhos dos saltos, o que permite calcular analiticamente a função característica do log-preço, de forma análoga ao modelo de Heston. (Ver Haugh (2013))

Esse modelo explica melhor o comportamento das caudas e o smile para maturidades curtas, sem perder a estrutura afim do Heston.

SABR

O modelo SABR (de Stochastic Alpha, Beta, Rho), proposto por Hagan et al. (2002), é amplamente utilizado na precificação e calibração de produtos de renda fixa, commodities e câmbio. Sua dinâmica é dada por:

$$\begin{cases} dS_t = rS_t dt + \sigma_t S_t^\beta dW_x, \\ d\sigma_t = \gamma \sigma_t dW_\sigma, \end{cases}$$

onde $\beta \in [0, 1]$ controla o grau de elasticidade da volatilidade em relação ao preço e $d\langle W_x, W_\sigma \rangle = \rho dt$, com $\rho \in [-1, 1]$.

A força do modelo está na aproximação analítica da volatilidade implícita proposta pelos autores, que oferece excelente ajuste para smiles observados em

mercado. Para o caso $\beta = 1$, isto é, o modelo log-normal, a aproximação é dada por:

$$\sigma_{imp}(\tau, K) \approx \sigma_t \frac{y}{\psi(y)} \left[1 + \left(\frac{1}{4} \rho \gamma \sigma_t + \frac{2 - 3\rho^2}{24} \gamma^2 \right) \tau \right],$$

onde

$$y := \frac{\gamma}{\sigma_t} \log \left(\frac{S e^{r\tau}}{K} \right)$$

e

$$\psi(y) := \log \left(\frac{\sqrt{1 - 2\rho y + y^2} + y - \rho}{1 - \rho} \right).$$

Essa aproximação é válida especialmente quando $\gamma^2 \tau \ll 1$, e contribuiu para a ampla adoção do modelo na indústria financeira.

Heston SLV

O modelo de *Heston Stochastic-Local Volatility* surge como uma proposta híbrida, que combina as dinâmicas de *smile* realistas do modelo de Heston à estimação exata garantida pelos métodos de volatilidade local. Mais ainda, Stoep, Grzelak e Oosterlee (2014) mostraram que este esquema permite uma precificação rápida e consistente que são sensíveis ao *skew* de volatilidades futuras.

Sua dinâmica é dada por:

$$\begin{cases} dS_t/S_t = r dt + \sigma(t, S_t) \psi(v_t) dW_x, \\ dv_t = a_v(t, v_t) dt + b_v(t, v_t) dW_v, \end{cases} \quad (3.13)$$

onde σ é o componente de variância local, $\psi(v_t)$ controla a volatilidade estocástica, os parâmetros a_v e b_v determinam o *drift* e a difusão do processo de variância, respectivamente, r é taxa de juros constante e os movimentos brownianos são correlacionados com parâmetro ρ .

Perceba, ainda, que esse modelo SLV geral descrito pela primeira equação da dinâmica acima pode colapsar num modelo de volatilidade puramente local ou estocástica, dependendo da escolha dos parâmetros. Por exemplo, tanto o modelo de Heston como o de Schobel-Zhu (que supõe a variância como sendo um processo OU, não um CIR) podem ser explicados a partir da dinâmica 3.13. Essa abordagem SLV é mais comum na prática do que utilizar apenas métodos locais ou estocásticos.

3.6.1 O paradigma da Rough Volatility

É comum, no mundo de derivativos, modelar o log-preço de um ativo como uma semimartingale. Isto é, para um ativo com log-preço Y_t , sua dinâmica é geralmente descrita por

$$dY_t = \mu_t dt + \sigma_t dW_t,$$

onde W_t é um movimento browniano padrão unidimensional e σ_t representa a volatilidade instantânea.

Como vimos, no contexto de Black-Scholes, $\sigma_t \equiv \sigma$ constante ou uma função determinística do tempo. No paradigma da volatilidade local, tem-se $\sigma_t = \sigma(t, Y_t)$, uma função determinística do tempo e do preço. Em ambos os casos, a dinâmica é pouco realista e gera superfícies de volatilidade completamente diferentes das observadas no mercado. Por outro lado, nos chamados *modelos de volatilidade estocástica*, como Hull-White, Heston e SABR, a volatilidade σ_t é ela própria uma semi-martingale contínua, permitindo uma dinâmica mais realista. Ainda assim, os preços gerados por esses modelos não reproduzem de forma consistente os preços de opções europeias observados empiricamente.

Uma mudança de paradigma foi sugerida por Comte e Renault (1998) ao propor o uso do movimento browniano fracionário (fBm) para modelar a log-volatilidade. Anos depois, ao revisitar a questão da suavidade do processo de volatilidade utilizando dados de alta frequência, Gatheral, Jaisson e Rosenbaum (2018) mostraram que, de fato, a log-volatilidade se comporta essencialmente como um fBm com expoente de Hurst $H \in (0, \frac{1}{2})$. Essa descoberta levou ao desenvolvimento do chamado paradigma da *Rough Volatility*, que propõe que a volatilidade exibe uma estrutura temporal muito mais irregular, ou “rugosa”, do que a prevista pelos modelos clássicos.

Para compreender essa abordagem, precisamos de alguns conceitos novos. Para detalhes, ver Dupret (2020).

Definição 3.2. Um processo estocástico $\{X_t\}_{t \geq 0}$ é chamado **autossimilar** se $\exists H > 0$ tal que, para todo $c > 0$,

$$\{X_{ct}\}_{t \geq 0} \stackrel{D}{=} \{c^H X_t\}_{t \geq 0}, \quad (3.14)$$

onde $\stackrel{D}{=}$ representa igualdade em distribuição e H é chamado **expoente de Hurst** do processo.

Intuitivamente, a propriedade de autossimilaridade significa que um processo escalado no tempo é estatisticamente o mesmo do processo original quando reescalado adequadamente no espaço.

Observação 3.5. Um processo autossimilar não pode ser estritamente estacionário, mas pode possuir incrementos estritamente estacionários. Vale lembrar que um processo possui incrementos estritamente estacionários se, para qualquer $h \in \mathbb{R}^+$,

$$\{X_{t+h} - X_t\}_{t \geq 0} \stackrel{D}{=} \{X_t - X_0\}_{t \geq 0}.$$

Além disso, se X_t é gaussiano com média 0 e $X_0 = 0$, então a propriedade de incrementos estacionários equivale a

$$\mathbb{E}(X_t - X_s)^2 = \mathbb{E}(X_{t-s})^2, \forall s, t \in \mathbb{R}^+.$$

Definição 3.3. Seja $0 < H < 1$. Um processo gaussiano a valores reais $\{B_t^H\}_{t \geq 0}$ é chamado um **movimento browniano fracionário** (fBm) se $\mathbb{E}B_t^H = 0$ e

$$\text{cov}(B_t^H, B_s^H) = \mathbb{E}(B_t^H B_s^H) = \frac{1}{2} \{t^{2H} + s^{2H} - |t - s|^{2H}\} \mathbb{E}(B_1^H)^2. \quad (3.15)$$

Pode-se provar que, quando $0 < H < 1$, o movimento browniano fracionário possui a seguinte representação integral em termos de movimentos brownianos padrão:

$$B_t^H = C_H \left\{ \int_{-\infty}^0 \left[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right] dW_s + \int_0^t (t-s)^{H-\frac{1}{2}} dW_s \right\}, \quad (3.16)$$

onde W_t é um movimento browniano padrão e C_H é dado por

$$C_H = \left[\int_{-\infty}^0 \left((1-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right) ds + \frac{1}{2H} \right]^{-1/2}.$$

Agora, seguimos o desenvolvimento de Cheridito, Kawaguchi e Maejima (2003) para definir:

Definição 3.4. O processo de **Ornstein-Uhlenbeck fracionário** (fOU) $\{Y_t^H\}_{t \geq 0}$ é dado por

$$Y_t^H = \nu \int_{-\infty}^t e^{-\alpha(t-u)} dB_u^H + m, \quad t \in \mathbb{R}, \quad (3.17)$$

para $\alpha, \nu > 0$ e $m \in \mathbb{R}$.

O processo Y_t^H é estacionário e gaussiano, e pode ser interpretado como uma versão fracionária do Ornstein-Uhlenbeck, preservando reversão à média, mas com correlações de longo alcance determinadas por H .

Observação 3.6. A influência de H .

Para o caso $0 < H < 1/2$, o artigo mostra que o processo fOU apresenta incrementos consecutivos do fBm negativamente correlacionados, o que resulta num processo muito menos suave do que no caso $1/2 < H < 1$ e com comportamento muito mais irregular.

Agora, podemos apresentar o modelo introduzido por Gatheral, Jaisson e Rosenbaum (2018):

O modelo RFSV

O modelo que melhor se ajustou aos dados analisados pelos autores, denominado Rough Fractional Stochastic Volatility (RFSV), é definido por

$$\sigma_t = \exp\{Y_t^H\}, \quad t \in [0, T], \quad (3.18)$$

onde Y_t^H satisfaz a equação (3.17) para alguns parâmetros $\alpha, \nu > 0$, $m \in \mathbb{R}$ e $H < 1/2$, que controla a "rugosidade" da volatilidade. O modelo é estacionário e, quando $\alpha \ll 1/T$, a log-volatilidade se comporta localmente como um fBm, como afirma o resultado abaixo:

Proposição 3.5. *Sejam B^H um fBm e Y^α definido por (3.17) para um $\alpha > 0$ dado. Conforme $\alpha \rightarrow 0$,*

$$\mathbb{E} \sup_{t \in [0, T]} |Y_t^\alpha - Y_0^\alpha - \nu B_t^H| \rightarrow 0.$$

Gatheral, Jaisson e Rosenbaum (2018) mostraram que o RFSV reproduz com alta precisão as propriedades empíricas da volatilidade, além de oferecer um arcabouço promissor para previsão da variância e calibração de modelos mais gerais, como o Rough Heston.

Rough Heston

Para finalizar, apresentamos a adaptação do modelo de Heston para o contexto da volatilidade rugosa, proposta por Euch e Rosenbaum (2016). Os autores acrescentaram ao modelo de Heston o *kernel* de Volterra $(t - s)^{H-\frac{1}{2}}$, que tem um papel central na dinâmica do fBm para $H < 1/2$, da seguinte maneira:

$$\begin{cases} dS_t = S_t \sqrt{v_t} dW_t \\ v_t = v_0 + \frac{1}{\Gamma(H + \frac{1}{2})} \int_0^t (t - s)^{H-\frac{1}{2}} \kappa (\bar{v} - v_s) ds + \frac{1}{\Gamma(H + \frac{1}{2})} \int_0^t (t - s)^{H-\frac{1}{2}} \gamma \sqrt{v_s} dW_s^* \end{cases}$$

onde Γ é a função gama, os parâmetros κ , \bar{v} e γ desempenham o mesmo papel do modelo de Heston tradicional e $d\langle W_t, W_t^* \rangle = \rho dt$.

Os autores também derivam uma expressão para a função característica do log-preço para o modelo, e mostram que ela exibe a mesma estrutura da obtida para o modelo clássico:

$$\mathbb{E} e^{iaX_t} = \exp\{g_1(a, t) + v_0 g_2(a, t)\},$$

onde

$$\begin{aligned} g_1(a, t) &= \kappa \bar{v} \int_0^t h(a, s) ds, \\ g_2(a, t) &= \frac{1}{\Gamma(1/2 - H)} \int_0^t (t - s)^{-1/2-H} h(a, s) ds, \end{aligned}$$

e $h(a, t)$ é a solução da seguinte equação de Riccati fracionária:

$$\frac{1}{\Gamma(1/2 - H)} \frac{d}{dt} \int_0^t (t - s)^{-1/2-H} h(a, s) ds = \frac{1}{2} (-a^2 - ia) + (ia\rho\gamma - \kappa)h(a, t) + \frac{\gamma^2}{2} h(a, t)^2.$$

No limite $H \rightarrow 1/2$, a dinâmica converge para o modelo de Heston original. Por outro lado, para $H < 1/2$, a equação assume a forma de uma equação integral fracionária do tipo Volterra, que não admite solução em forma fechada, mas pode ser tratada numericamente por meio de esquemas apropriados.

Apêndice A

Disponibilidade dos códigos

Aqui, apresentamos o repositório onde estão os códigos em Python desenvolvidos como parte fundamental do trabalho aqui apresentado. O objetivo é garantir a transparência e a reprodutibilidade dos resultados, permitindo que outros estudantes possam replicar, validar e estender os experimentos aqui descritos.



Apêndice B

Resultados do *multi-start*

A seguir, as tabelas com os resultados nos experimentos de *multi-search*, separados por κ e, para PETR4, o resultado de todos os testes com os diferentes κ realizados.

ticker	erro	\bar{v}	γ	ρ	\bar{v}_0	γ_0	ρ_0
BBDC4	2.04111e-03	0.32052	0.80065	-0.34344	0.58491	2.17405	-0.45712
BBDC4	2.00000e+10	0.46366	0.96297	-0.99890	0.46366	2.18894	-0.99890
BBDC4	2.04112e-03	0.32052	0.80064	-0.34338	0.48112	0.17519	-0.50485
BBDC4	2.04112e-03	0.32052	0.80064	-0.34338	0.58673	2.15363	-0.73748
BBDC4	2.04147e-03	0.32048	0.80060	-0.34199	0.96967	1.68697	-0.12456
BBDC4	2.04112e-03	0.32052	0.80064	-0.34339	0.28423	2.62512	-0.81316
BBDC4	2.04146e-03	0.32048	0.80060	-0.34201	0.90143	1.06274	-0.26072
BBDC4	2.04112e-03	0.32052	0.80064	-0.34338	0.15020	2.36176	-0.60487
BBDC4	2.04147e-03	0.32048	0.80060	-0.34200	0.88356	2.90877	-0.21759
BBDC4	2.04111e-03	0.32052	0.80065	-0.34343	0.08954	1.55544	-0.55330
BBDC4	2.04112e-03	0.32052	0.80064	-0.34337	0.78961	0.16018	-0.42935
BBDC4	2.04112e-03	0.32052	0.80065	-0.34340	0.24585	0.15648	-0.58257
BBDC4	2.04111e-03	0.32052	0.80065	-0.34340	0.22183	2.24614	-0.76228
BBDC4	2.04147e-03	0.32048	0.80060	-0.34199	0.79549	0.78887	-0.25797
BBDC4	2.04147e-03	0.32048	0.80060	-0.34197	0.55283	2.34218	-0.21649
BBDC4	2.04111e-03	0.32052	0.80065	-0.34343	0.86091	0.61880	-0.95013
BBDC4	2.04112e-03	0.32052	0.80064	-0.34337	0.28543	1.61717	-0.50392
BBDC4	2.04111e-03	0.32052	0.80065	-0.34343	0.35109	1.63870	-0.82682
BBDC4	2.04148e-03	0.32048	0.80060	-0.34195	0.67834	1.56581	-0.20914
BBDC4	2.04111e-03	0.32052	0.80065	-0.34341	0.16973	2.30762	-0.77700
ITUB4	8.33528e-04	0.08000	0.40000	-0.87552	0.58491	2.17405	-0.45712
ITUB4	8.31227e-04	0.08000	0.40000	-0.88650	0.46366	2.18894	-0.99890
ITUB4	8.31702e-04	0.08000	0.40000	-0.88295	0.48112	0.17519	-0.50485
ITUB4	8.31345e-04	0.08000	0.40000	-0.88541	0.58673	2.15363	-0.73748
ITUB4	8.44743e-04	0.08000	0.40000	-0.85467	0.96967	1.68697	-0.12456
ITUB4	8.31220e-04	0.08000	0.40000	-0.88656	0.28423	2.62512	-0.81316
ITUB4	8.38990e-04	0.08000	0.40000	-0.86333	0.90143	1.06274	-0.26072
ITUB4	8.31426e-04	0.08000	0.40000	-0.88476	0.15020	2.36176	-0.60487
ITUB4	8.41612e-04	0.08000	0.40000	-0.85908	0.88356	2.90877	-0.21759
ITUB4	8.31515e-04	0.08000	0.40000	-0.88413	0.08954	1.55544	-0.55330
ITUB4	8.32479e-04	0.08000	0.40000	-0.87921	0.78961	0.16018	-0.42935
ITUB4	8.31299e-04	0.08000	0.40000	-0.88580	0.24585	0.15648	-0.58257
ITUB4	8.31223e-04	0.08000	0.40000	-0.88653	0.22183	2.24614	-0.76228
ITUB4	8.37377e-04	0.08000	0.40000	-0.86629	0.79549	0.78887	-0.25797
ITUB4	8.37930e-04	0.08000	0.40000	-0.86524	0.55283	2.34218	-0.21649
ITUB4	8.31222e-04	0.08000	0.40000	-0.88655	0.86091	0.61880	-0.95013
ITUB4	8.32038e-04	0.08000	0.40000	-0.88116	0.28543	1.61717	-0.50392
ITUB4	8.31223e-04	0.08000	0.40000	-0.88653	0.35109	1.63870	-0.82682
ITUB4	8.38896e-04	0.08000	0.40000	-0.86349	0.67834	1.56581	-0.20914
ITUB4	8.31222e-04	0.08000	0.40000	-0.88654	0.16973	2.30762	-0.77700

Tabela B.1: Resultados do *multi-start* para $\kappa = 1$.

ticker	erro	\bar{v}	γ	ρ	\bar{v}_0	γ_0	ρ_0
BBAS3	1.19840e-02	0.09860	0.76914	-0.75869	0.58491	2.17405	-0.45712
BBAS3	8.00000e+10	0.08000	0.69282	0.99900	0.46366	2.18894	-0.99890
BBAS3	1.19860e-02	0.09857	0.76903	-0.76306	0.48112	0.17519	-0.50485
BBAS3	1.19840e-02	0.09860	0.76914	-0.75869	0.58673	2.15363	-0.73748
BBAS3	1.19840e-02	0.09860	0.76914	-0.75868	0.96967	1.68697	-0.12456
BBAS3	1.19840e-02	0.09859	0.76913	-0.75876	0.28423	2.62512	-0.81316
BBAS3	1.19840e-02	0.09860	0.76914	-0.75837	0.90143	1.06274	-0.26072
BBAS3	1.19840e-02	0.09860	0.76914	-0.75861	0.15020	2.36176	-0.60487
BBAS3	1.19840e-02	0.09860	0.76914	-0.75869	0.88356	2.90877	-0.21759
BBAS3	1.19840e-02	0.09860	0.76914	-0.75870	0.08954	1.55544	-0.55330
BBAS3	1.19840e-02	0.09860	0.76914	-0.75847	0.78961	0.16018	-0.42935
BBAS3	1.19859e-02	0.09857	0.76904	-0.76295	0.24585	0.15648	-0.58257
BBAS3	1.19840e-02	0.09860	0.76914	-0.75872	0.22183	2.24614	-0.76228
BBAS3	1.19840e-02	0.09860	0.76914	-0.75839	0.79549	0.78887	-0.25797
BBAS3	1.19840e-02	0.09860	0.76914	-0.75865	0.55283	2.34218	-0.21649
BBAS3	1.19858e-02	0.09857	0.76904	-0.76282	0.86091	0.61880	-0.95013
BBAS3	1.19840e-02	0.09860	0.76914	-0.75869	0.28543	1.61717	-0.50392
BBAS3	1.19840e-02	0.09860	0.76914	-0.75866	0.35109	1.63870	-0.82682
BBAS3	1.19840e-02	0.09860	0.76914	-0.75866	0.67834	1.56581	-0.20914
BBAS3	1.19840e-02	0.09860	0.76914	-0.75870	0.16973	2.30762	-0.77700
VALE3	3.96762e-03	0.08000	0.69282	-0.41840	0.58491	2.17405	-0.45712
VALE3	3.96762e-03	0.08000	0.69282	-0.42033	0.46366	2.18894	-0.99890
VALE3	3.98373e-03	0.08000	0.69282	-0.43836	0.48112	0.17519	-0.50485
VALE3	3.96762e-03	0.08000	0.69282	-0.42031	0.58673	2.15363	-0.73748
VALE3	3.97620e-03	0.08000	0.69282	-0.40552	0.96967	1.68697	-0.12456
VALE3	3.96762e-03	0.08000	0.69282	-0.42033	0.28423	2.62512	-0.81316
VALE3	3.96771e-03	0.08000	0.69282	-0.41771	0.90143	1.06274	-0.26072
VALE3	3.96763e-03	0.08000	0.69282	-0.42039	0.15020	2.36176	-0.60487
VALE3	3.96762e-03	0.08000	0.69282	-0.41845	0.88356	2.90877	-0.21759
VALE3	3.96762e-03	0.08000	0.69282	-0.42036	0.08954	1.55544	-0.55330
VALE3	3.96787e-03	0.08000	0.69282	-0.42190	0.78961	0.16018	-0.42935
VALE3	5.38337e-03	0.08000	0.58372	-0.57772	0.24585	0.15648	-0.58257
VALE3	3.96762e-03	0.08000	0.69282	-0.42035	0.22183	2.24614	-0.76228
VALE3	3.96784e-03	0.08000	0.69282	-0.41700	0.79549	0.78887	-0.25797
VALE3	3.96762e-03	0.08000	0.69282	-0.41844	0.55283	2.34218	-0.21649
VALE3	3.96777e-03	0.08000	0.69282	-0.42144	0.86091	0.61880	-0.95013
VALE3	3.96763e-03	0.08000	0.69282	-0.41837	0.28543	1.61717	-0.50392
VALE3	3.96762e-03	0.08000	0.69282	-0.42034	0.35109	1.63870	-0.82682
VALE3	3.96762e-03	0.08000	0.69282	-0.41843	0.67834	1.56581	-0.20914
VALE3	3.96762e-03	0.08000	0.69282	-0.42031	0.16973	2.30762	-0.77700

Tabela B.2: Resultados do Multi-start para $\kappa = 3$.

κ	erro	\bar{v}	γ	ρ	\bar{v}_0	γ_0	ρ_0
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.58491	2.17405	-0.45712
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.46366	2.18894	-0.99890
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.48112	0.17519	-0.50485
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.58673	2.15363	-0.73748
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.96967	1.68697	-0.12456
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.28423	2.62512	-0.81316
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.90143	1.06274	-0.26072
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.15020	2.36176	-0.60487
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.88356	2.90877	-0.21759
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.08954	1.55544	-0.55330
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.78961	0.16018	-0.42935
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.24585	0.15648	-0.58257
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.22183	2.24614	-0.76228
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.79549	0.78887	-0.25797
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.55283	2.34218	-0.21649
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.86091	0.61880	-0.95013
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.28543	1.61717	-0.50392
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.35109	1.63870	-0.82682
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.67834	1.56581	-0.20914
0.5	4.32523e-01	0.08000	0.28284	0.99900	0.16973	2.30762	-0.77700
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.58491	2.17405	-0.45712
1.0	4.44276e-01	0.08000	0.10000	-0.67594	0.46366	2.18894	-0.99890
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.48112	0.17519	-0.50485
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.58673	2.15363	-0.73748
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.96967	1.68697	-0.12456
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.28423	2.62512	-0.81316
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.90143	1.06274	-0.26072
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.15020	2.36176	-0.60487
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.88356	2.90877	-0.21759
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.08954	1.55544	-0.55330
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.78961	0.16018	-0.42935
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.24585	0.15648	-0.58257
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.22183	2.24614	-0.76228
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.79549	0.78887	-0.25797
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.55283	2.34218	-0.21649
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.86091	0.61880	-0.95013
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.28543	1.61717	-0.50392
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.35109	1.63870	-0.82682
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.67834	1.56581	-0.20914
1.0	4.02619e-01	0.08000	0.40000	0.99900	0.16973	2.30762	-0.77700

Tabela B.3: Resultados do Multi-start para PETR4 ($\kappa = 0.5$ e $\kappa = 1.0$).

κ	erro	\bar{v}	γ	ρ	\bar{v}_0	γ_0	ρ_0
3.0	3.22809e-01	0.08000	0.69282	0.89265	0.58491	2.17405	-0.45712
3.0	3.22809e-01	0.08000	0.69282	0.89270	0.46366	2.18894	-0.99890
3.0	3.22819e-01	0.08000	0.69282	0.88070	0.48112	0.17519	-0.50485
3.0	3.22809e-01	0.08000	0.69282	0.89264	0.58673	2.15363	-0.73748
3.0	3.22809e-01	0.08000	0.69282	0.89271	0.96967	1.68697	-0.12456
3.0	3.22809e-01	0.08000	0.69282	0.89269	0.28423	2.62512	-0.81316
3.0	3.22809e-01	0.08000	0.69282	0.89271	0.90143	1.06274	-0.26072
3.0	3.22809e-01	0.08000	0.69282	0.89268	0.15020	2.36176	-0.60487
3.0	3.22809e-01	0.08000	0.69282	0.89265	0.88356	2.90877	-0.21759
3.0	3.22809e-01	0.08000	0.69282	0.89264	0.08954	1.55544	-0.55330
3.0	3.22809e-01	0.08000	0.69282	0.89238	0.78961	0.16018	-0.42935
3.0	3.43446e-01	0.08000	0.66709	0.15956	0.24585	0.15648	-0.58257
3.0	3.22809e-01	0.08000	0.69282	0.89264	0.22183	2.24614	-0.76228
3.0	3.22809e-01	0.08000	0.69282	0.89246	0.79549	0.78887	-0.25797
3.0	3.22809e-01	0.08000	0.69282	0.89270	0.55283	2.34218	-0.21649
3.0	3.22809e-01	0.08000	0.69282	0.89255	0.86091	0.61880	-0.95013
3.0	3.22809e-01	0.08000	0.69282	0.89267	0.28543	1.61717	-0.50392
3.0	3.22809e-01	0.08000	0.69282	0.89264	0.35109	1.63870	-0.82682
3.0	3.22809e-01	0.08000	0.69282	0.89271	0.67834	1.56581	-0.20914
3.0	3.22809e-01	0.08000	0.69282	0.89264	0.16973	2.30762	-0.77700
7.0	2.35360e-01	0.08000	1.05830	0.61436	0.58491	2.17405	-0.45712
7.0	2.35360e-01	0.08000	1.05830	0.61439	0.46366	2.18894	-0.99890
7.0	2.87098e-01	0.08000	0.10000	-0.32439	0.48112	0.17519	-0.50485
7.0	2.35360e-01	0.08000	1.05830	0.61434	0.58673	2.15363	-0.73748
7.0	2.35360e-01	0.08000	1.05830	0.61381	0.96967	1.68697	-0.12456
7.0	2.35360e-01	0.08000	1.05830	0.61433	0.28423	2.62512	-0.81316
7.0	2.35360e-01	0.08000	1.05830	0.61308	0.90143	1.06274	-0.26072
7.0	2.35360e-01	0.08000	1.05830	0.61439	0.15020	2.36176	-0.60487
7.0	2.35360e-01	0.08000	1.05830	0.61438	0.88356	2.90877	-0.21759
7.0	2.35360e-01	0.08000	1.05830	0.61437	0.08954	1.55544	-0.55330
7.0	2.86459e-01	0.08000	0.10109	-0.16619	0.78961	0.16018	-0.42935
7.0	2.87675e-01	0.08000	0.10000	-0.46640	0.24585	0.15648	-0.58257
7.0	2.35360e-01	0.08000	1.05830	0.61436	0.22183	2.24614	-0.76228
7.0	2.35360e-01	0.08000	1.05830	0.61282	0.79549	0.78887	-0.25797
7.0	2.35360e-01	0.08000	1.05830	0.61433	0.55283	2.34218	-0.21649
7.0	2.35362e-01	0.08000	1.05830	0.61011	0.86091	0.61880	-0.95013
7.0	2.35360e-01	0.08000	1.05830	0.61434	0.28543	1.61717	-0.50392
7.0	2.35360e-01	0.08000	1.05830	0.61435	0.35109	1.63870	-0.82682
7.0	2.35360e-01	0.08000	1.05830	0.61365	0.67834	1.56581	-0.20914
7.0	2.35360e-01	0.08000	1.05830	0.61435	0.16973	2.30762	-0.77700

Tabela B.4: Resultados do Multi-start para PETR4 ($\kappa = 3.0$ e $\kappa = 7.0$).

Referências Bibliográficas

ASH, R. B.; DOLÉANS-DADE, C. A. *Probability & Measure Theory*. 2nd. ed. San Diego: Academic Press, 2000. Disponível em: <<https://faculty.math.illinois.edu/~r-ash/PMT.html>>.

BATES, D. S. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, v. 9, n. 1, p. 69–107, 1996.

BENHAMOU, E.; GOBET, E.; MIRI, M. Time dependent heston model. *SIAM Journal on Financial Mathematics*, v. 1, n. 1, p. 289–325, 2010. Accurate analytical formula for the Heston model with time-dependent parameters. Disponível em: <<https://hal.science/hal-00370717>>.

CAMPBELL, J. Y.; HENTSCHEL, L. No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial economics*, Elsevier, v. 31, n. 3, p. 281–318, 1992.

CHERIDITO, P.; KAWAGUCHI, H.; MAEJIMA, M. Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, Institute of Mathematical Statistics, v. 8, n. 3, p. 1–14, 2003.

COMTE, F.; RENAULT, E. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, Wiley, v. 8, n. 4, p. 291–323, 1998.

DUPRET, J.-L. *Rough stochastic volatility modeling*. Dissertação (Mémoire présenté en vue de l'obtention du mastère en sciences actuarielles) — Institut de Statistique, Biostatistique et Actuariat (ISBA), Université Catholique de Louvain, Louvain-La-Neuve, ago. 2020.

EUCH, O. E.; ROSENBAUM, M. The characteristic function of rough Heston models. *arXiv preprint arXiv:1609.02108*, 2016. Disponível em: <<https://arxiv.org/abs/1609.02108>>.

FRENCH, K. R.; SCHWERT, G. W.; STAMBAUGH, R. F. Expected stock returns and volatility. *Journal of financial economics*, Elsevier, v. 19, n. 1, p. 3–29, 1987.

GATHERAL, J. *The Volatility Surface: A Practitioner's Guide*. Hoboken, NJ: Wiley, 2006. ISBN 978-0471792512.

GATHERAL, J.; JAISSON, T.; ROSENBAUM, M. Volatility is rough. *Quantitative Finance*, Taylor & Francis, v. 18, n. 6, p. 933–949, jun. 2018. Empirical evidence that volatility exhibits roughness inconsistent with smooth models like Heston-CIR.

GUILLAUME, F.; SCHOUTENS, W. *Model and calibration risks for the Heston model*. [S.l.], 2010. Disponível em: <<https://eurandom.tue.nl/reports/2010/026-report.pdf>>.

HAGAN, P. S.; KUMAR, D.; LESNIEWSKI, A. S.; WOODWARD, D. E. Managing smile risk. *Wilmott Magazine*, p. 84–108, jul. 2002.

HAUGH, M. *Local Volatility, Stochastic Volatility and Jump-Diffusion Models*. 2013. Lecture Notes, IEOR E4707: Financial Engineering: Continuous-Time Models. Columbia University. Disponível em: <https://www.columbia.edu/~mh2078/FoundationsFE/LocalStochVol_JumpDiff.pdf>.

ITÔ, K. My Sixty Years along the Path of Probability Theory. In: *The Inamori Foundation: Kyoto Prizes & Inamori Grants*. Kyoto: The Inamori Foundation, 1998. p. 142–177. Palestra Comemorativa do 14º Prêmio Kyoto (1998). Publicado em 25/12/1999. Texto em Japonês (p. 142) e Inglês (p. 147). Disponível em: <<https://www.inamori-f.or.jp/>>.

KUO, H.-H. *Introduction to Stochastic Integration*. [S.l.]: Springer, 2006.

MUSIELA, M.; RUTKOWSKI, M. *Martingale Methods in Financial Modelling*. 2nd. ed. [S.l.]: Springer, 2005.

OOSTERLEE, C. W.; GRZELAK, L. A. *Mathematical Modeling and Computation in Finance: With Exercises and Python and MATLAB Computer Codes*. [S.l.]: World Scientific Publishing Company, 2020. ISBN 978-1-78634-794-7.

PFÜTZENREUTER, E. *Investindo no Mercado de Opções*. 2nd. ed. [S.l.]: Novatec, 2021.

PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd. ed. [S.l.]: Cambridge University Press, 2007.

ROUAH, F. D. *The Heston Model and Its Extensions in MATLAB and C#*. Hoboken, NJ: Wiley, 2013. ISBN 978-1118548256.

SAPORITO, Y. F. *Modelos de volatilidade para derivativos*. Rio de Janeiro: Editora FGV, 2023. (Cadernos de Matemática). ISBN 978-65-5652-235-7.

SEBER, G. A. F.; WILD, C. J. *Nonlinear Regression*. Hoboken, New Jersey: Wiley-Interscience, 2003. (Wiley Series in Probability and Statistics). Reprint of the 1989 edition.

SHREVE, S. E. *Stochastic Calculus for Finance II: Continuous-Time Models*. [S.l.]: Springer, 2004.

STOEP, A. W. van der; GRZELAK, L. A.; OOSTERLEE, C. W. THE HESTON STOCHASTIC-LOCAL VOLATILITY MODEL: EFFICIENT MONTE CARLO SIMULATION. *International Journal of Theoretical and Applied Finance*, v. 17, n. 07, p. 1450045, 2014.

WIERSEMA, U. *Brownian Motion Calculus*. Chichester, UK: Wiley, 2008. ISBN 978-0470724623.

ØKSENDAL, B. *Stochastic Differential Equations: An Introduction with Applications*. 6th. ed. [S.l.]: Springer, 2003.