

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

FERNANDO GOUVEIA PASSOS

**MODELAGEM DA EVOLUÇÃO DINÂMICA DO SKEW
DE VOLATILIDADE IMPLÍCITA DE OPÇÕES VIA
FILTRO DE KALMAN**

SÃO PAULO

2019

FERNANDO GOUVEIA PASSOS

**MODELAGEM DA EVOLUÇÃO DINÂMICA DO SKEW
DE VOLATILIDADE IMPLÍCITA DE OPÇÕES VIA
FILTRO DE KALMAN**

Dissertação apresentada ao Programa de
Mestrado Profissional da Escola de Economia
de São Paulo da Fundação Getulio Vargas,
como requisito para a obtenção do título de
Mestre em Economia.

Área de concentração:
Engenharia Financeira.

Orientador:
Prof. Dr. Afonso de Campos Pinto

SÃO PAULO

2019

Passos, Fernando Gouveia.

Modelagem da evolução dinâmica do skew de volatilidade implícita de opções via Filtro de Kalman / Fernando Gouveia Passos. - 2019.

65 f.

Orientador: Prof. Dr. Afonso de Campos Pinto.

Dissertação (mestrado profissional MPFE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Kalman, Filtragem de. 2. Volatilidade (Finanças). 3. Teoria da previsão. 4. Mercado de opções. 5. Engenharia financeira. I. Pinto, Afonso de Campos. II. Dissertação (mestrado profissional MPFE) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 336.764.2

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

FERNANDO GOUVEIA PASSOS

MODELAGEM DA EVOLUÇÃO DINÂMICA DO SKEW DE VOLATILIDADE IMPLÍCITA DE OPÇÕES VIA FILTRO DE KALMAN

Dissertação apresentada ao Programa de
Mestrado Profissional da Escola de Economia
de São Paulo da Fundação Getúlio Vargas,
como requisito para a obtenção do título de
Mestre em Economia.

Área de concentração:
Engenharia Financeira.

Data da Aprovação: / /

Banca Examinadora:

Prof. Dr. Afonso de Campos Pinto
(Orientador)
EESP-FGV

Prof. Dr. Alessandro Martim Marques
EESP-FGV

**Prof. Dr. Luiz Henrique Moraes da
Silva**
EAESP-FGV

RESUMO

Este trabalho foca na implementação do algoritmo de Filtro de Kalman discreto e linear - modelo de espaço de estados - cujo processo de estimação recursivo baseia-se na incorporação das curvas empíricas de volatilidade implícita disponíveis no mercado. O objetivo do modelo é capturar a evolução estocástica dos coeficientes que direcionam o formato dos *skews* de volatilidade com uma única maturidade das opções de PETR4 e do Índice IBOVESPA, considerando as projeções de 1 dia e comparando a performance de predição com outros modelos *benchmark*. Evidenciamos, por meio dos indicadores estatísticos e análise das densidades preditivas da variação diária do valor dos portfólios de opções, que o modelo mostra performance superior vis-à-vis a dos outros modelos, o que é crucial para administração de risco, geração de valor econômico em estratégias de *trading* e apreçamento de opções exóticas. Ainda, exploramos o impacto dos movimentos passados dos fatores dos *skews* do Índice IBOVESPA no poder preditivo do modelo.

Palavras-chave: Volatilidade Implícita. Filtro de Kalman. Densidade preditiva.

ABSTRACT

The present dissertation is focused on the implementation of a discrete and linear Kalman filter algorithm - a state space model - based on an empirical skew updating recursive process in order to efficiently estimate and forecast the coefficients that captures the dynamics of implied volatility curves of PETR4 and IBOVESPA Index options on a daily basis. We also compare our model to two alternative models that will be further presented. From a risk management and economic value perspectives, we found that the model mostly outperformed the two alternative models by producing good density forecasts of daily returns on different option portfolios and also good statistical measures. Furthermore, considering the impact on the model predictability power, we explored the relationship between Index IVS and PETR4 options IVS.

Keywords: Implied volatility. Kalman filter. Density forecasting.

Lista de ilustrações

Figura 1 – Evolução diária da Volatilidade Implícita ATM como <i>proxy</i> da volatilidade de PETR4.	30
Figura 2 – Evolução dos coeficientes estimados pelo modelo KF_PETR4 e KF_IBOVESPA para o período de Jan_2015 a Dez_2017.	46
Figura 3 – Evolução do RMSE-IV para o 3 modelos nos períodos fora da amostra.	49
Figura 4 – Momentos das distribuições do PnL diário para <i>straddle</i> ATM.	52
Figura 5 – Momentos das distribuições do PnL diário para <i>risk reversal</i> com 25 Delta.	52
Figura 6 – Momentos das distribuições do PnL diário para <i>butterfly spread</i> com 50 e 25 Delta.	53
Figura 7 – Evolução do índice de Sharpe para <i>straddle</i> [25, 50 e 75 Delta].	55

Lista de tabelas

Tabela 1 – Interpretação das variáveis explicativas da equação de medida	23
Tabela 2 – Fluxos de caixa de posições <i>delta hedged</i> em <i>Call/Put</i> OTM.	39
Tabela 3 – Parâmetros estimados dos modelos KF e KF_IBOV	43
Tabela 4 – Matrizes de correlação dos modelos KF e KF_IBOV.	44
Tabela 5 – Correlação cruzada entre os coeficientes KF_PETR4 e KF_IBOVESPA.	45
Tabela 6 – Resumo das estatísticas para os parâmetros estimados pelo modelo II.	45
Tabela 7 – Valores médios das medidas estatísticas do erro de previsão.	47
Tabela 8 – Teste de acurácia de previsão (contra o Modelo KF).	49
Tabela 9 – Evolução dos valores da medida MDA para os 3 modelos.	50
Tabela 10 – Resumo das estatísticas dos resultados diários dos portfólios.	51
Tabela 11 – Resumo estatístico das estratégias de sinais com <i>straddle</i> [25, 50 e 75 Delta].	54

Sumário

1	Introdução	10
2	Revisão Bibliográfica	14
3	Fundamentação Teórica	19
3.1	O Modelo de Espaço de Estado	19
3.2	O Algoritmo Discreto e Linear do Filtro de Kalman	20
3.3	Os Modelos Propostos de Filtro de Kalman	22
3.3.1	Parametrização do Modelo Teórico de Filtro de Kalman	22
3.3.2	Método de Estimação	26
3.4	Modelo Benchmark	27
4	Metodologia	29
4.1	Tratamento dos Dados	29
4.2	Implementação dos modelos KF e KF_IBOV	32
4.3	Estimação dos Modelos	33
4.4	Implementação do Modelo II	35
4.5	Método de análise da performance de predição	36
4.5.1	Previsão do <i>skew</i> pelos modelos KF e KF_IBOV	36
4.5.2	Previsão do <i>skew</i> pelo modelo II	36
4.5.3	Indicadores Estatísticos de Desempenho	38
4.5.4	Estratégias de <i>Trading</i>	38
4.5.5	Valor Econômico das Estratégias de <i>Trading</i>	40
5	Resultados	43
5.1	Validação de performance dos modelos (<i>in sample</i>)	43
5.2	Análise dos Indicadores Estatísticos de Previsão (<i>out of sample</i>)	47
5.3	Resultados das estratégias de <i>Trading</i>	49
5.3.1	Análise das densidades preditivas	50
5.3.2	Análise do valor econômico das estratégias	53
6	Conclusão	57
	Referências	59
	Apêndices	61
	APÊNDICE A Variância assintótica do processo de Ornstein-Uhlenbeck	62
	APÊNDICE B Spline cúbico	63

1 Introdução

A despeito do modelo de Black & Scholes (BS) ser largamente aceito e utilizado no mercado para precificação de derivativos em virtude da sua simplicidade e tratabilidade matemática, algumas proposições, como assumir que o retorno do ativo objeto segue uma distribuição log-normal e a volatilidade do modelo é determinística, podem produzir preços enviesados vis-à-vis os preços de mercado. Com isso, há impactos no gerenciamento dos fatores de risco do portfólio, como, por exemplo, o vega e o delta. Ainda, sob a perspectiva de apreçamento de opções exóticas, que considera o *smile* da volatilidade, o modelo deve se ajustar perfeitamente à superfície de volatilidade implícita de modo a assegurar robustez aos preços dos derivativos.

Neste contexto, assumindo, por meio de investigações empíricas, que o retorno do ativo segue uma distribuição leptocúrtica, ou seja, com caudas mais pesadas em relação à distribuição normal e que a volatilidade varia em função de diferentes *strikes* e maturidades¹, modelos alternativos foram empregados a fim de capturar essas características da distribuição do ativo, como por exemplo, a inclusão de um processo com saltos nos preços em detrimento do movimento browniano geométrico (MBG) utilizado pelo BS, que possibilita a formação de uma distribuição com excesso de curtose como resultado do componente do salto ou da variância da própria volatilidade; e o processo estocástico para a volatilidade, que pode capturar uma distribuição com assimetria, derivada, dentre outros fatores, da correlação entre os termos de aleatoriedade dos processos do ativo e da volatilidade.

Como exemplo destes modelos, podem-se citar: Heston (1993) propôs um modelo de volatilidade estocástica (SVM), especificando a dinâmica do ativo objeto sob uma medida neutra ao risco, e adicionando um processo estocástico para a volatilidade do ativo; Bates (1996) e Kou (2002) propuseram modelos de *jump-diffusion* ou a combinação deste com o SVM a fim de capturar o *smile* da curva para diferentes maturidades; Derman e Kani (1994) propuseram a volatilidade como uma função determinística do ativo subjacente e o tempo, cujo processo é análogo à derivação da equação diferencial parcial que origina a equação de BS. Embora os dados do modelo sejam bem ajustados à superfície empírica de volatilidade implícita em um determinado dia, a evolução dinâmica dos parâmetros estimados pode tornar o modelo menos ajustado em outro dia, gerando necessidade de constante re-calibragem, o que o torna computacionalmente intensivo e pouco eficiente.

Os modelos mencionados caracterizam-se pela modelagem da volatilidade *spot* do

¹ Há evidências na análise das superfícies de volatilidade implícita, bem como uma extensa literatura sobre modelos ARCH e GARCH, que evidenciam a heterocedasticidade da variância (BOLLERSLEV; KRONER, 1992)

ativo objeto, que é uma variável não observável; ao passo que o mercado precifica as opções utilizando diretamente as volatilidades implícitas. Portanto, os parâmetros estimados por estes modelos capturam de forma menos aderente o formato da superfície empírica de volatilidade. Ainda, os autores evidenciaram instabilidade temporal nos parâmetros estimados, especificamente nos modelos combinados SVM e *jump-diffusion*, com deficiência para capturar a evolução da estrutura a termo da superfície. Por último, há grande dificuldade para estimação dos parâmetros de modelos com saltos.

Notando a inabilidade destes modelos baseados apenas nas informações do ativo subjacente para descrever o comportamento dinâmico da superfície de volatilidade implícita (IVS)², alguns autores optaram por incorporar as propriedades estatísticas inerentes à dinâmica das opções e seu conteúdo informativo, ou seja, modelar de forma direta a IVS observável.

Ainda, além da variação da volatilidade para diferentes *strikes* e maturidades em um determinado instante de tempo, a volatilidade implícita (IV) varia, de forma estocástica, em relação a diferentes instantes de tempo para cada opção. Logo, pressupõe-se que a IVS também evolui no tempo, surgindo a necessidade de modelagem desta dinâmica.

Pode-se dividir estes *frameworks* em três vertentes:

- Modelos *Market-based* - O objetivo é modelar diretamente a IV baseado nos preços de mercado, com a imposição de restrições nos coeficientes estimados para assegurar condições de não arbitragem (SCHONBUCHER, 1999).
- Modelo Vega-Gamma-Vanna-Volga - Proposto por Carr e Wu (2016) que especifica apenas a dinâmica neutra ao risco da IV, sem definir o processo estocástico da volatilidade do ativo, como é feito nos *market-models*, assumindo que toda a superfície da IV é direcionada por um único fator de aleatoriedade (MBG). A partir do formato da superfície empírica aplicando a função de BS, obtém-se, por meio do Lema de Itô, uma equação diferencial parcial (EDP) que dispõe de condições de não arbitragem.
- Modelos *factor-based* - Surgiram como alternativa à literatura vigente, cujo foco recai ou sobre a modelagem das superfícies individuais (*cross-sectional*) da IV com parametrização estática, como, por exemplo, o modelo *Sticky-Moneyness*, ignorando as dimensões e propriedades das séries de tempo; ou sobre a modelagem da estrutura a termo de um *strike* específico, por exemplo, uma opção *at-the-money* (ATM), considerando as informações incrementais da serie temporal. O modelo assume que a evolução de toda a superfície da IV pode ser capturada por um número restrito de fatores. Para tanto, Cont e Fonseca (2002) e Villa, Fengler e Hardle (2003)

² Há estudos que corroboram que o poder preditivo de modelos baseados na volatilidade implícita são superiores ao de modelos baseados apenas na dinâmica do ativo objeto e sua variância (FLEMING, 2002)

utilizaram métodos não paramétricos, como análise dos componentes principais (PCA), que extrai os fatores para estudar as suas dinâmicas; e Gonçalves e Guidolin (2006) desenvolveram um método paramétrico, utilizando uma abordagem em 2 estágios: capturar a evolução dinâmica das superfícies individuais por meio de uma função linear, obtendo a estimação dos parâmetros para cada período; e a análise das informações das séries temporais destes parâmetros por meio do Vetor Auto regressivo³.

Neste âmbito, Bedendo e Hodges (2009) adotaram o filtro de Kalman para estimação simultânea dos parâmetros do modelo de forma mais robusta, sendo a primeira aplicação do filtro para atualização da curva de volatilidade. Os autores propõem a modelagem da dinâmica do *skew* de volatilidade, ou seja, a curva para uma maturidade específica, por meio de um modelo de espaço de estados, de modo a incorporar toda a informação disponível no mercado. Para cada instante de tempo, tem-se a função linear e discreta, definida por um polinômio cúbico, cujas variáveis independentes são as medidas de *moneyness*, que captura a superfície única da IV. O objetivo é estimar os coeficientes do polinômio, denominados variáveis de estado, e, por consequência, estimar o *skew* de volatilidade. O modelo utilizado como base deste trabalho seguiu esta abordagem.

Embora haja na literatura diversos estudos utilizando o modelo de filtro de Kalman no mercado de opções, especificamente para o Índice S&P500, há ainda escassez de trabalhos com foco na projeção da volatilidade das opções de ações individuais. No Brasil, ainda há um forte direcionamento para implementação deste método para modelagem no mercado de juros.

Neste contexto, o objetivo do trabalho é implementar o algoritmo do Filtro de Kalman, que busca uma estimação ótima dos parâmetros e eficiente projeção do *skew* da IV para 1 dia, e aplicá-lo nos dados disponíveis de opções de PETR4 do Índice IBOVESPA, bem como explorar a relação conjunta dos movimentos passados dos *skews* de volatilidade.

De forma intuitiva, assume-se que o processo de atualização do *skew* da IV da opção de ação individual é mais lento em função da frequência de negociação de uma variedade de *strikes*. Logo, se as opções do Índice forem mais líquidas do que as da ação, e houver evidência de correlação dinâmica entre os coeficientes, pressupõe-se que o conteúdo informacional dos *skews* diários do Índice possa explicar de forma parcial os movimentos da curva de volatilidade das opções da ação, e desta forma, aprimorar a estimativa dos coeficientes que conduzem o *skew* e elevar a performance de predição do modelo vis-à-vis a de modelos *benchmark*.

³ O autor evidencia a instabilidade dos parâmetros estimados por meio de testes da autocorrelação, o que corrobora a necessidade de explorar as informações das series. Dumas, Fleming e Whaley (1998) também obtiveram os mesmos resultados.

Evidenciamos, no entanto, que a inclusão dos *skews* passados das opções do Índice não elevaram o poder preditivo do modelo proposto para as opções de PETR4, a despeito da evidência de correlação entre curvas diárias (*cross sectional*).

Ainda, os resultados reportados evidenciam que as curvas diárias de volatilidade implícita das opções de PETR4 são estatisticamente previsíveis, refletindo a análise dos indicadores do erro de previsão e estratégias de *trading*, cujos resultados foram majoritariamente superiores ao retorno do ativo objeto (PETR4), utilizado como *benchmark*, para as diferentes janelas fora da amostra.

Além do modelo permitir potencial geração de valor econômico por meio dos portfólios de opções e gestão de risco em função da maior robustez e consistência ao longo dos períodos fora da amostra, pode-se empregar essa modelagem para apreçamento e *hedge* de opções exóticas.

Este trabalho está dividido em 6 capítulos. O capítulo dois trata da revisão bibliográfica contextualizando o trabalho em relação aos estudos e técnicas já implementadas nos mercados de opções externo e interno. O capítulo três apresenta o modelo de Filtro de Kalman linear como um modelo de espaço de estado, bem como o modelo *benchmark* e o método de estimação dos parâmetros. O quarto capítulo demonstra a metodologia de implementação e calibração dos modelos, o tratamento e obtenção dos dados e as técnicas estatística e econômicas de análise de predição. No capítulo cinco são apresentados os resultados da aplicação dos modelos no mercado de opções brasileiro e o capítulo seis encerra com a conclusão do trabalho, abordando potenciais limitações e possíveis extensões.

2 Revisão Bibliográfica

Este capítulo tem por objetivo revisar brevemente a literatura existente sobre modelos de volatilidade estocástica com a incorporação de saltos, bem como analisar de forma detalhada os estudos recentes sobre modelagem da dinâmica da IVS, sendo a base para elaboração deste trabalho.

Merton (1976) propôs uma extensão do BS com a incorporação do processo de Poisson para capturar variações anormais de preço, assumindo que a distribuição da frequência dos saltos é gaussiana. Kou (2002) estendeu o modelo de Merton ao adotar uma distribuição exponencial dupla para os saltos aleatórios.

Heston (1993) introduziu um modelo em que especifica uma dinâmica estocástica para a volatilidade do processo, cujo termo de difusão é a raiz quadrada da variância do ativo e o *drift* apresenta reversão à média. A correlação entre a aleatoriedade do processo da variância e do processo do retorno do ativo permite o ajuste do modelo às características assimétrica e leptocúrtica da distribuição do ativo. A forma analítica resulta da implementação da estratégia do portfólio replicante livre de arbitragem, ajustando as quantidades do ativo e opção visando obter posição neutra no delta e vega. Em seguida, deriva-se a equação diferencial parcial (PDE) ao aplicar o Lemma de Itô e a sua solução é obtida por meio da aplicação da transformada de Fourier.

Bates (1996) desenvolveu um método para precificação de opções americanas modeladas com processos combinados de SVM e *jump-diffusion*, como extensão à metodologia de precificação empregada por Heston (1993). O modelo mostra-se consistente com os dados de mercado da IV das opções de cambio do marco alemão, explicando as características assimétricas da distribuição. Nota-se, contudo, que há instabilidade dos parâmetros, principalmente em relação à estrutura a termo da IVS.

Como limitações adicionais dos modelos mencionados acima, pode-se citar o fator único que conduz a SVM na equação diferencial estocástica (SDE), sem a possibilidade de extensão para um modelo multifatorial; e a complexidade dos métodos para precificação de forma analítica, como por exemplo, o modelo de Heston.

Schonbucher (1999) desenvolveu um *market-based framework* para capturar a comportamento estocástico da IVS, cuja dinâmica do ativo é representada pelo movimento browniano lognormal sob uma medida neutra ao risco e o processo da IV é direcionado por um *drift*, um vetor dos termos de difusão que conduzem a IV e o termo de aleatoriedade do retorno do ativo, refletindo a correlação entre esses fatores por meio da matriz de covariância do movimento browniano. Em vez de derivar as condições de não arbitragem diretamente da IVS, a abordagem considera os dados observáveis da IVS corrente e impõe

restrições à dinâmica dos fatores do processo de modo a assegurar que os preços descontados sejam *martingale* sob uma medida neutra ao risco. Para tanto, deriva-se a PDE do preço da opção por meio do Lemma de Itô a partir das dinâmicas da IV e do ativo. Com isso, impõe-se que o valor esperado (*drift*) desta PDE seja zero, constituindo uma fórmula fechada para estimação dos parâmetros para cada IV individual. A possível limitação do modelo baseia-se na premissa de que o conhecimento do IVS corrente é suficiente para estimar de forma estável a dinâmica futura do *drift* e assegurar que o modelo seja livre de arbitragem dinâmica. Por exemplo, se o *smile* inicial tem inclinação negativa, a correlação entre os termos de difusão não pode ser positiva, o que restringe a escolha ou estimação do componente *martingale*.

O modelo Vega-Gamma-Vanna-Volga proposto por Carr e Wu (2016) define uma SDE contínua para o ativo e uma SDE para a IV com um único termo de difusão que conduz toda a superfície, assumindo uma medida de probabilidade neutra ao risco para ambas equações. Com o objetivo de impôr restrições para eliminar a arbitragem dinâmica, impõe-se que o termo do *drift* da PDE originada pelo Lema de Itô na função de BS seja zero, ou seja, eliminando o fator de aleatoriedade da equação, baseado na definição de que a volatilidade implícita é o valor colocado na formula de BS tal que o preço teórico é igual ao preço de mercado das opções, tal como: $C_t(K, T) = C_{BS}(t, \hat{\sigma}(K, T)_t, S_t)$. Como resultado, tem-se uma PDE que define uma relação linear entre os componentes theta, vega, gamma, vanna e volga¹. Ao aplicar a equação das gregas pela fórmula de BS, a PDE é reduzida à uma equação algébrica, cujos coeficientes são estimados por meio da solução da equação quadrática. Nota-se que o objetivo não é identificar uma solução para a PDE, que, neste caso, é a própria função de BS, mas identificar as variáveis estocásticas que satisfazem essa equação a fim de excluir a arbitragem e, desta forma, obter os parâmetros estimados. Tem-se que as restrições de não arbitragem encapsuladas na PDE conectam o formato corrente da superfície de IV com os níveis correntes dos coeficientes estocásticos dos processos de difusão do IV e do ativo. Portanto, não se pode assegurar que tais condições se aplicam à evolução dinâmica destes fatores, configurando uma possível deficiência do modelo.

Cont e Fonseca (2002) propuseram um modelo que descreve a evolução no tempo do formato da IVS capturada por um número restrito de fatores ortogonais. A abordagem baseia-se na análise das séries temporais das IV's e, para tanto, constrói-se a IVS suavizada para cada dia com *strikes* e maturidades arbitrárias por meio de um método não paramétrico². A fim de analisar a dinâmica conjunta da série de variações diárias das IV's, aplica-se a decomposição de Karhunen–Loeve, que é uma generalização do PCA para

¹ O vanna é a sensibilidade do preço da opção em relação à covariação da vol e o preço do ativo; o volga mede a sensibilidade do vega em relação à volatilidade, seguindo a mesma relação entre o gamma e o delta.

² Estimador de Nadaraya–Watson

dimensões maiores. Pelos resultados empíricos do comportamento das séries e restrições de valor, pode-se interpretar os principais componentes, que respondem por 98% da variação da superfície, como nível, inclinação (*slope*) e curvatura (*smile*), que podem ser modelados por um processo auto regressivo de primeira ordem com reversão à média³ de acordo com os testes de auto correlação parcial. Como limitações do modelo, além de ser computacionalmente intensivo, há restrições analíticas dos fatores obtidos, com significado apenas estatístico, como análise de tendências e correlações, o que limita a utilização para precificação.

Gonçalves e Guidolin (2006) adotaram uma abordagem que combina o ajuste da IVS às superfícies diárias das opções do índice S&P500 (*cross-sectional*) como uma função linear do *moneyness* e maturidade⁴, com a aplicação do modelo de vetor auto regressivo nas séries temporais multivariadas dos coeficientes estimados. O objetivo era aprimorar o poder preditivo ao modelar a evolução dinâmica da IVS, considerando a evidência de que os coeficientes diários estimados na regressão são instáveis, o que corrobora o argumento de que as superfícies são variantes no tempo. Critérios estatísticos mostraram que o modelo tem habilidade de predição superior aos modelos *random walk* e NGARCH(1,1)⁵. Sob a perspectiva de valor econômico, aplicou-se simulações de estratégias de *trading* delta-hedged no período fora da amostra, computando excesso de índice de Sharpe vis-à-vis a estratégia de manter apenas uma posição comprada no índice S&P500. Salienta-se, entretanto, que a inclusão de custos transacionais, aos quais o investidor no varejo está sujeito, torna este excesso de ganho estatisticamente nulo, o que rejeita a violação da hipótese de eficiência de mercado das opções. A limitação do modelo decorre da possibilidade de haver algum erro na parametrização do primeiro estágio, o que prejudicaria a especificação das séries temporais e, por conseguinte, a projeção dos fatores que direcionam a evolução da IVS. Além disso, a abordagem em 2 estágios torna a estimação dos parâmetros sub ótima se comparada com a estimação simultânea da dinâmica dos coeficientes diários por meio da implementação do filtro de Kalman.

De modo geral, o filtro de Kalman é empregado para estimar de forma recursiva o estado completo de um processo linear a partir de uma sequência de observações ruidosas com distribuição gaussiana e o conhecimento prévio do sistema dinâmico de modo a gerar de forma eficiente e aprimorada a média e covariância da densidade preditiva (*a posteriori*). Pode-se citar, dentre as suas diversas aplicações, a implementação realizada pela NASA de sistemas de navegação de mísseis e ônibus espaciais. Para tanto, essa abordagem de

³ AR(1)/Ornstein-Uhlenbeck.

⁴ Aplica-se a regressão linear nos coeficientes e não linear no *moneyness* e maturidade, utilizando o método dos Mínimos Quadrados Ordinários (MQO) para estimação, sendo a mesma abordagem utilizada por Dumas, Fleming e Whaley (1998).

⁵ Modelo proposto por Heston e Nandi (2000), que considera os coeficientes invariantes no tempo, a despeito das densidades neutras ao risco serem variantes no tempo. Portanto, é razoável assumir que sua performance preditiva é igual ou inferior à do modelo proposto.

estimação simultânea requer que o modelo seja representado em forma de espaço de estados (HARVEY, 1989)⁶.

O trabalho de Bedendo e Hodges (2009) focou na modelagem da evolução temporal do *skew* de volatilidade das opções do índice S&P500⁷ utilizando o algoritmo linear e discreto de filtro de Kalman para estimação mais robusta (via máxima verossimilhança) dos coeficientes estocásticos que direcionam a curva de volatilidade, visando incorporar de forma simultânea informações disponíveis no mercado dos novos negócios de opções individuais. O *skew* é modelado pela função polinomial cúbica do *moneyness*⁸, cuja dinâmica dos coeficientes estimados segue um processo de Ornstein-Uhlenbeck gaussiano, correlacionado com o processo EGARCH (1,1) do índice S&P500.

A análise da capacidade preditiva do modelo é fundamentada a partir de trajetórias simuladas de Monte Carlo bivariado para cada dia, computando, no período fora da amostra, as densidades preditivas das variáveis de estado do algoritmo de KF e do retorno do ativo, bem como as IVs projetadas, via polinômio cúbico, e o *moneyness*. Obtém-se, aplicando a formula de BS, as densidades da variação diária do valor de algumas estratégias de *trading* de modo a explorar as variações de diferentes naturezas na curva de volatilidade (i.e. estratégias como *straddle*, *risk reversal* e *butterfly*, são sensíveis às variações no nível, inclinação (assimetria) e curvatura (curtose) da curva (distribuição de retornos), respectivamente).

A performance do modelo proposto é comparada com 2 modelos existentes no mercado: O modelo *Sticky-Delta* que assume que a volatilidade é fixa em relação ao *moneyness* (ou, de forma equivalente, ao delta), sendo $\sigma_{k,t} = \sigma_0 - b(K - S_t)$, em que $b > 0$ direciona o *skew*. O único fator que afeta a curva de volatilidade é a evolução do preço do ativo; e o modelo Vega-Gamma (VG) que envolve a aplicação da expansão das séries de Taylor da função de BS em relação à variação do preço do ativo, volatilidade e tempo.

Para aplicação do modelo VG, os autores consideraram o σ_{ATM} , representando mudanças apenas no nível da curva e, neste caso, não incluíram o termo de segunda ordem da volatilidade na expansão de Taylor, que seria a inclinação da curva. Como os portfólios são delta-neutro e a variação de preço é diária, exclui-se o delta e o theta. Com isso, a evolução do preço da opção é representada pelas gregas (vega e gamma) e as dinâmicas da volatilidade ($\Delta\sigma_{ATM,t}$) e preço (ΔF_t), representadas por um processo com reversão à média com inovações gaussianas e o EGARCH(1,1), respectivamente.

⁶ Um modelo de espaço de estado (MEE) representa a descrição da evolução de um sistema, com múltiplas entradas e saídas, de ordem n por um conjunto de n equações diferenciais de primeira ordem.

⁷ Considera-se apenas o *skew* diário para uma determinada maturidade, e não a estrutura a termo da superfície, pois os autores se baseiam nos apontamentos feitos por Harvey (1989) sobre a implementação do Filtro de Kalman multidimensional (*skew*/maturidade) ser pouco consistente.

⁸ Por meio de investigações do formato dos componentes principais dos *skews* projetados pelo modelo, os fatores $\{X_i\}_{i=1,\dots,4}$ são interpretados como nível, inclinação, curvatura e assimetria, respectivamente.

Conclui-se que os primeiros momentos das densidades do modelo de filtro de Kalman apresentam maior aderência aos momentos empíricos vis-à-vis as estatísticas dos modelos *benchmark*, corroborando a necessidade de mais de um fator para explicar a dinâmica do *skew*. Contudo, evidencia-se que momentos de ordem superior apresentam pior ajuste aos dados, refletindo a deficiência do modelo em capturar as características da distribuição empírica. Os autores indicaram a especificação do ativo - EGARCH, que considera o efeito alavancagem e a heterocedasticidade da variância - como uma potencial fonte de erro, podendo incluir saltos e inovações não gaussianas.

Chen, Zhen e Shenghong (2016) propuseram uma generalização do modelo proposto por Bedendo e Hodges (2009), eliminando a restrição da equação de medida (dados observáveis - IV) ser linear, alegando que a parametrização não linear dispõe de maior *fitting*.

Define-se a função paramétrica g não linear que captura a superfície única da IV em cada instante de tempo - $\hat{\sigma}(t; K, T) = g(t, K, T, x^1_t, x^2_t, \dots, x^p_t)$, em que $\{x_i\}_{i=1, \dots, p}$ são os p fatores, denominados variáveis de estado, que direcionam a evolução da superfície, e cuja dinâmica pode ser descrita por um conjunto de SDEs. De forma intuitiva, o algoritmo de filtro de Kalman busca estimar os parâmetros destas SDEs e, com isso, derivar a dinâmica da IVS por meio do Lema de Itô. Neste trabalho, o modelo SABR foi utilizado como a função g . Visto que o filtro de Kalman requer que as equações de medida e de estado sejam lineares e as distribuições das variáveis sejam gaussianas, fez-se necessário a introdução do Filtro de Kalman Unscented (UKF). Ao contrário do Filtro de Kalman Estendido (EKF), que aproxima por meio da série de Taylor de primeira ordem as funções não linear e linear, ou seja, utilizando apenas um ponto, por exemplo o valor esperado, o UKF utiliza pontos estratégicos chamados de pontos sigma, incluindo a média, de uma distribuição gaussiana para criar uma outra distribuição gaussiana aproximada, após a transformação destes pontos a partir do mapeamento da função não linear.

Observando a performance da densidade preditiva do Filtro de Kalman linear, bem como a possibilidade de incluir uma variável exógena na equação discreta de estado, este trabalho baseia-se no artigo elaborado por Bedendo e Hodges (2009) para implementação do modelo no mercado brasileiro de opções de ações individuais. Cabe ressaltar que, baseado em pesquisas de artigos e dissertações locais, notou-se a aplicação do algoritmo de Filtro de Kalman apenas no mercado de juros brasileiro, deflagrando potencial contribuição para o estudo abordado neste trabalho. Ainda, outra contribuição é a exploração da relação dinâmica entre os coeficientes (ou variáveis de estado) dos *skews* de volatilidade das opções do Índice IBOVESPA e da ação individual. O objetivo é incrementar a informação utilizada pelo algoritmo de Filtro de Kalman e, deste modo, aprimorar o poder preditivo.

3 Fundamentação Teórica

Este capítulo descreve de forma detalhada o modelo de KF discreto e linear, bem como as características específicas do modelo proposto, além de uma breve descrição de 2 modelos alternativos.

3.1 O Modelo de Espaço de Estado

A abordagem do modelo de espaço de estado (MEE) visa a análise dos dados em séries temporais, possibilitando a extração de informações sobre a dinâmica do modelo. De modo geral, o MEE descreve a evolução de um sistema dinâmico com múltiplas entradas e saídas relacionadas por equações diferenciais parciais (EDPs). A flexibilidade para implementação do MEE advém da representação destas EDPs de ordem n por um conjunto de n EDPs de primeira ordem.

Diversos modelos podem ser colocados no formato de espaço de estado, sendo mais utilizado em modelos com parâmetros variantes no tempo, como ocorre no processo de filtragem de KF, em cuja abordagem os parâmetros são estimados de forma recursiva, ou seja, incorporando os dados disponíveis em cada instante de tempo.

O MEE permite decompor uma série de tempo em componentes relevantes, como ocorre nos modelos de análise de componentes principais - PCA, utilizado para análise de tendências, extração de ruídos, interpolação e suavização de dados incompletos.

Harvey (1989) define, de forma genérica, um sistema linear de espaço de estado a partir de uma equação de medida, que relaciona as variáveis observáveis (\mathbf{y}_t) ao vetor de estado (não observável $-\mathbf{x}_t$), que pode ser, por exemplo, o vetor de coeficientes da regressão:

$$\mathbf{y}_t = \mathbf{Z}_t \mathbf{x}_t + \mathbf{d}_t + \epsilon_t, t = 1, \dots, T \quad (3.1)$$

em que \mathbf{y}_t é um vetor $N \times 1$, \mathbf{x}_t é um vetor $m \times 1$, \mathbf{Z}_t é a matriz de transição $N \times m$, \mathbf{d}_t é um vetor $N \times 1$ e ϵ_t é um vetor $N \times 1$ de distúrbios serialmente não correlacionados, com média zero e matriz de covariância \mathbf{H}_t .

A equação de transição (ou estado) é gerada por um processo de Markov de primeira ordem:

$$\mathbf{x}_t = \mathbf{T}_t \mathbf{x}_{t-1} + \mathbf{c}_t + \eta_t, t = 1, \dots, T \quad (3.2)$$

em que \mathbf{T}_t é a matriz de transição $m \times m$, \mathbf{c}_t é um vetor $m \times 1$ e η_t é um vetor $g \times 1$ de distúrbios serialmente não correlacionados, com média zero e matriz de covariância \mathbf{Q}_t .

Segundo [Harvey \(1989\)](#), a condição para conhecimento do sistema completo em t é alcançada a partir do conhecimento da média \mathbf{x}_0 e matriz de covariância \mathbf{P}_0 , bem como da proposição de não correlação serial entre os ruídos $E[\epsilon_t \eta_t^T] = 0$. As matrizes de transição são não-estocásticas; mesmo que variem no tempo, o fazem de maneira pré-determinada

3.2 O Algoritmo Discreto e Linear do Filtro de Kalman

De forma comparativa ao método de estimação de Mínimos Quadrados Ordinários (MQO), cujo modelo de regressão linear é determinístico e as variáveis explicativas são utilizadas de forma conjunta na série temporal, o algoritmo do Filtro de Kalman é utilizado para estimar o estado completo de um processo dinâmico não estacionário, direcionado por um ruído, e com parâmetros variantes no tempo, denominados variáveis de estado (dados não observáveis), incorporando observações ruidosas com distribuição gaussiana ao sistema dinâmico já conhecido em t de modo a atualizar recursivamente a média condicional e covariância da densidade preditiva em $t + 1$ ([HARVEY, 1989](#)). Ressalta-se que tal algoritmo foi concebido para utilização em modelos lineares e gaussianos, sendo necessárias modificações no filtro para implementação em dados com distribuição que não seja normal, como o filtro de Kalman Estendido e *Unscented*, que não estão inseridos no escopo deste trabalho.

Pode-se dividir a implementação do algoritmo em 2 fases:

- 1ª Fase - Predição: Seja \mathbf{x}_k um vetor de estado $n \times 1$ normalmente distribuído, com média conhecida \mathbf{x}_{k-1} e matriz de covariância \mathbf{P}_{k-1} , a equação de propagação no formato espaço de estado permite obter a estimativa de estado em k , denominada estimativa *a priori* $\hat{\mathbf{x}}_{k|k-1} = E(\mathbf{x}_k | \mathcal{F}_{k-1})$, em que a σ -álgebra \mathcal{F}_{k-1} contempla toda informação disponível até o instante $k - 1$.

Sendo $k = 1$, assumimos conhecer o vetor inicial de estado \mathbf{x}_0 , com distribuição gaussiana, com média $\hat{\mathbf{x}}_0$ e matriz de covariância $\mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T]$,

$$\mathbf{x}_{k|k-1} = \mathbf{F}_k \mathbf{x}_{k-1} + \omega_k, \quad (3.3)$$

em que $\mathbf{x}_{k|k-1}$ é o vetor $n \times 1$, \mathbf{F}_k é a matriz de transição determinística $n \times n$ no instante k e ω_k é o vetor $n \times 1$ do distúrbio, assumindo correlação temporal nula e distribuição normal i.i.d. com média 0 e matriz de covariância \mathbf{Q} .

Aplicando o valor esperado na equação (3.3):

$$\begin{aligned} E(\mathbf{x}_{k|k-1}) &= \mathbf{F}_k E(\mathbf{x}_{k-1|k-1}) \\ \hat{\mathbf{x}}_{k|k-1} &= \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1}. \end{aligned}$$

Com isso, deduz-se o erro da estimativa *a priori* $\tilde{\mathbf{x}}_{k|k-1}$:

$$\mathbf{x}_{k|k-1} - \hat{\mathbf{x}}_{k|k-1} = \tilde{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \tilde{\mathbf{x}}_{k-1|k-1} + \omega_k.$$

Logo, obtém-se a matriz de covariância do erro *a priori*:

$$\mathbf{P}_{k|k-1} = E(\tilde{\mathbf{x}}_{k|k-1} \tilde{\mathbf{x}}_{k|k-1}^T) = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}. \quad (3.4)$$

- 2ª Fase - Atualização: a estimativa *a priori* é combinada com os dados observáveis disponíveis em cada instante k visando aprimorar a estimativa do vetor de estado, denominada estimativa *a posteriori*.

1. Processo de filtragem ou atualização por meio da equação de medida (ME)¹

$$\mathbf{y}_k = \mathbf{G}_k \mathbf{x}_k + \epsilon_k, \quad (3.5)$$

em que \mathbf{y}_k é o vetor $m \times 1$ de medição com m variáveis observáveis no instante k , \mathbf{G}_k é a matriz $m \times n$ que relaciona as variáveis de estado com as medidas observáveis e ϵ_k é o vetor $m \times 1$ do ruído branco gaussiano com média 0 e matriz de covariância \mathbf{R} .

Os resíduos de medição podem ser deduzidos a partir da aplicação do valor esperado na equação (3.5):

$$\begin{aligned} \mathbf{y}_k - \hat{\mathbf{y}}_{k|k} &= \tilde{\mathbf{y}}_k = \mathbf{G}_k \mathbf{x}_k + \epsilon_k - \mathbf{G}_k \hat{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{y}}_k &= \mathbf{G}_k \tilde{\mathbf{x}}_{k|k-1} + \epsilon_{k|k}. \end{aligned} \quad (3.6)$$

Mostra-se, portanto, que $E[\tilde{\mathbf{y}}_k] = 0$. Os resíduos são comumente denominados *inovações*, pois representam novas informações as quais o estado atual do filtro não possui.

A matriz de covariância dos resíduos é definida como:

$$\begin{aligned} \mathbf{C}_k &= E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^T] = E[(\mathbf{G}_k \tilde{\mathbf{x}}_{k|k-1} + \epsilon_{k|k})(\mathbf{G}_k \tilde{\mathbf{x}}_{k|k-1} + \epsilon_{k|k})^T] \\ \mathbf{C}_k &= \mathbf{G}_k \mathbf{P}_{k|k-1} \mathbf{G}_k^T + \mathbf{R}, \end{aligned} \quad (3.7)$$

assumindo que não há correlação temporal entre os ruídos do processo e medição - $E[\omega_k \epsilon_k^T] = 0$. Por definição, os resíduos são brancos (não serialmente correlacionados) - $E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] = 0$. A demonstração será apresentada na seção 3.3.2.

2. Atualização da estimativa *a posteriori* - $\hat{\mathbf{x}}_{k|k} = E(\mathbf{x}_k | \mathcal{F}_{k-1}, \mathbf{y}_k)$, caracterizado por um modelo Bayesiano (média condicional) ou, de forma equivalente, uma distribuição gaussiana conjunta de \mathbf{x}_k e \mathbf{y}_k :

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{G}_k \hat{\mathbf{x}}_{k|k-1}) = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k. \quad (3.8)$$

¹ Pode-se considerar, como exemplo, um sensor/GPS em um sistema dinâmico em que o vetor de estado contempla a posição e velocidade de uma partícula.

Nota-se que o algoritmo se baseia em uma combinação linear entre a estimativa *a priori* e o erro entre o valor real e valor esperado da medição. Para tanto, identifica-se uma matriz \mathbf{K}_k ($n \times m$), denominada ganho ótimo de Kalman, tal que a diferença entre o valor real e estimado *a posteriori* $\tilde{\mathbf{x}}_{k|k}$ seja mínima.

$$\mathbf{x}_k - \hat{\mathbf{x}}_{k|k} = \tilde{\mathbf{x}}_{k|k} = \mathbf{x}_k - (\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{G}_k\hat{\mathbf{x}}_{k|k-1})).$$

Substituindo \mathbf{y}_k por $\mathbf{G}_k\mathbf{x}_k + \epsilon_k$, obtém-se:

$$\tilde{\mathbf{x}}_{k|k} = \tilde{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{G}_k\tilde{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\epsilon_k \Rightarrow \tilde{\mathbf{x}}_{k|k} = \tilde{\mathbf{x}}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{G}_k) + \mathbf{K}_k\epsilon_k.$$

A matriz de covariância do erro da estimativa *a posteriori* será:

$$\mathbf{P}_{k|k} = E(\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T) = \mathbf{P}_{k|k-1} + \mathbf{K}_k\mathbf{G}_k\mathbf{P}_{k|k-1}\mathbf{K}_k^T\mathbf{G}_k^T + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T.$$

Logo:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} + \mathbf{K}_k(\mathbf{G}_k\mathbf{P}_{k|k-1}\mathbf{G}_k^T + \mathbf{R})\mathbf{K}_k^T. \quad (3.9)$$

Sabendo que o traço da matriz de covariância equivale à soma dos elementos da diagonal principal, neste caso a variância de cada fator $\{x_n\}_{n=1,\dots,4}$, o objetivo de minimizar a variância do erro *a posteriori* é obtido ao igualar a derivada do traço a zero para um dado \mathbf{K}_k , ou seja:

$$\frac{\partial \text{tr}(\mathbf{P}_{k|k})}{\partial \mathbf{K}_k} = 0,$$

obtendo, a partir da equação (3.9):

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{G}_k^T(\mathbf{G}_k\mathbf{P}_{k|k-1}\mathbf{G}_k^T + \mathbf{R})^{-1}. \quad (3.10)$$

Sem prejuízo na generalidade para a equação acima, pode-se incluir o termo $(\mathbf{K}_k\mathbf{G}_k\mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{G}_k\mathbf{P}_{k|k-1})$ e substituir o \mathbf{K}_k :

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{G}_k\mathbf{P}_{k|k-1} \Rightarrow \mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{G}_k)\mathbf{P}_{k|k-1}. \quad (3.11)$$

A linearidade das equações (3.3) e (3.5), em conjunto com a premissa de que \mathbf{x}_0 , ω_k e ϵ_k são variáveis gaussianas, implica que a função densidade de probabilidade condicional $p(\mathbf{x}_k|\mathbf{y}_k)$ também é gaussiana, caracterizada unicamente por uma média condicional $\hat{\mathbf{x}}_{k|k}$ e matriz de covariância $\mathbf{P}_{k|k}$.

O Filtro de Kalman é um algoritmo sequencial que pode ser utilizado para gerar $\hat{\mathbf{x}}_{k|k}$ e $\mathbf{P}_{k|k}$ para cada instante de tempo k , com uma densidade $\mathbf{x}_{k+1} \sim N(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1})$.

3.3 Os Modelos Propostos de Filtro de Kalman

3.3.1 Parametrização do Modelo Teórico de Filtro de Kalman

Este trabalho visa implementar o algoritmo de Filtro de Kalman, denominado modelo KF, com caráter preditivo, cujo processo de estimação baseia-se na incorporação

de novas medidas disponíveis em t , que são, neste trabalho, as curvas empíricas de volatilidade implícita para uma única maturidade, de modo a obter uma estimativa de estado aprimorada em $t + 1$. O objetivo é estimar de forma recursiva os coeficientes, denominados variáveis de estado não observáveis, que capturam a evolução do formato do *skew* de volatilidade. Estas variáveis de estado serão dadas por uma função polinomial ajustada às curvas observadas no mercado, cujas variáveis explicativas são quatro categorias da medida *moneyness* $\{m^i\}_{i=0,\dots,3}$, que buscam refletir os diferentes tipos de variação da curva. Define-se o *moneyness* como:

$$m = \frac{\log(\frac{E}{F})}{\sigma_{ATM}\sqrt{\tau}}, \quad (3.12)$$

em que E é o *strike*, F o preço futuro do ativo objeto e τ a maturidade da opção. Este conceito foi proposto por Tompkins (1999) com o objetivo de relacionar a distância entre o *strike* e o nível de preço do ativo objeto em função do preço atual do ativo S_t , τ e o σ_{ATM} . Desta forma, essa métrica normalizada pode ser interpretada como um desvio padrão, o que facilita a comparação direta entre diferentes ativos e modelos. Neste trabalho, considera-se, sem perda de generalidade, $\sigma_{ATM} = 1$.

De forma intuitiva, o polinômio cúbico visa capturar os diferentes tipos de choques no formato do *skew*, sendo interpretado de acordo com a Tabela 1.

Tabela 1 – Interpretação das variáveis explicativas da equação de medida

$\{m^i\}$	Descrição dos Fatores (ME)
$i = 0$	Nível do σ_{ATM} (<i>log moneyness</i> = 0)
$i = 1$	Inclinação (<i>slope</i>)
$i = 2$	Curvatura ou Convexidade (<i>smile</i>)
$i = 3$	Assimetria (<i>skewness</i>)

A parametrização da equação linear de medida baseia-se em estudos empíricos de modelos de regressão com diferentes formas funcionais, como cúbica, quadrática e linear, aplicados no período da amostra, indicando que a especificação do σ_t que melhor se ajusta às superfícies empíricas é a função polinomial cúbica do *moneyness* (m) e, portanto, considerar-se-á $\{\mathbf{X}_i\}_{i=1,\dots,4}$ como os coeficientes que conduzem o *skew* de volatilidade.

Logo, a equação de medida (ME) que representa o σ_t como um vetor $n \times 1$ com n *strikes* diferentes para cada instante de tempo t é definida como:

$$\sigma_t = \mathbf{M}_t \mathbf{X}_t + \epsilon_t, \quad (3.13)$$

em que \mathbf{M}_t é a matriz $n \times 4$ com os elementos $[\mathbf{M}]_{i,j} = m_i^{j-1}$, $\epsilon_t \sim (0, \mathbf{R})$ é o vetor dos ruídos serialmente não correlacionados, independentes de \mathbf{X}_t , sendo $\mathbf{R} = E[\epsilon_t \epsilon_t^T] = \sigma_\epsilon^2 I$ a matriz de covariância do ruído, assumindo que a distribuição gaussiana é independente e identicamente distribuída.

A previsão da estrutura a termo de volatilidade utilizando o modelo GARCH, proposto por Bollerslev e Kroner (1992), mostra que

$$\lim_{n \rightarrow \infty} E(\sigma_{t+n}^2 | t) = \frac{\alpha}{1 - (\beta - \alpha)},$$

indicando que há convergência para uma variância não condicional (de longo prazo ou assintótica), o que corrobora a premissa de que a volatilidade tende a reverter à média. Por conseguinte, a equação de estado é definida a partir da transformação de um processo contínuo de *Ornstein-Uhlenbeck* (OU) em um processo discreto, que pode representar um processo linear *Vector AutoRegressive* VAR(1)².

Seja uma SDE cuja solução representa um processo contínuo de OU:

$$d\mathbf{X}_t = \mathbf{A}(\mu - \mathbf{X}_{t-1})dt + \sigma_t d\mathbf{W}_t,$$

em que $d\mathbf{W}_t$ é um processo de Wiener, \mathbf{A} é a matriz diagonal 4×4 dos coeficientes de reversão à média de longo prazo³, μ é o vetor 4×1 das médias de longo prazo. Discretizando a SDE, em cuja equação o termo $d\mathbf{W}_t = \eta_t \sqrt{\Delta t}$, com $\eta_t \sim N(0, 1)$, tem-se:

$$\mathbf{X}_t - \mathbf{X}_{t-1} = \mathbf{A}\mu\Delta t - \mathbf{A}\mathbf{X}_{t-1}\Delta t + \sigma_t \eta_t \sqrt{\Delta t}.$$

Como neste trabalho considera-se os dados diários e, portanto, $\Delta t = 1$, a equação acima pode ser reescrita da forma:

$$\mathbf{X}_t = \mathbf{A}\mu - (\mathbf{I} - \mathbf{A})\mathbf{X}_{t-1} + \eta_t, \text{ com } \eta_t \sim N(0, \sigma_t^2).$$

Tomando $\mathbf{a} = \mathbf{A}\mu$, $\mathbf{U} = \mathbf{I} - \mathbf{A}$ e $\mathbf{u}_t = \eta_t$, obtém-se a equação de estado:

$$\mathbf{X}_t = \mathbf{a} + \mathbf{U}\mathbf{X}_{t-1} + \mathbf{u}_t, \quad (3.14)$$

que representa um processo VAR(1), sendo $\mathbf{u}_t \sim N(0, \mathbf{V})$ o vetor dos ruídos serialmente não correlacionados, independente de \mathbf{X}_{t-1} , com a matriz de covariância do erro igual a \mathbf{V} . Assume-se que a correlação cruzada entre os erros das equações de estado e de medição é nula - $E[\mathbf{u}_t \epsilon_t^T] = 0$.

Como já mencionado, uma das contribuições deste trabalho é a exploração do impacto do conteúdo informacional das superfícies diárias das opções do Índice IBOVESPA sobre a dinâmica dos coeficientes da superfície de opções de uma ação individual, neste caso, PETR4. Denominaremos este modelo como KF_IBOV.

O objetivo é construir uma estimação aprimorada de $\hat{\mathbf{X}}_{t|t}$ como a média condicional de \mathbf{X}_t , dados os valores passados dos vetores $\{\mathbf{X}_{t-1}^{IBOV}\}$ e $\{\sigma_t\}$, sendo

² Como já mencionado, o modelo dinâmico discreto pressupõe um processo VAR(1), ou de forma equivalente, um processo oculto de Markov.

³ Assumindo que a taxa de reversão à média de um coeficiente do *skew* não depende da taxa do outro coeficiente, o que parece ser bastante razoável.

$$\hat{\mathbf{X}}_{t|t} = E[\mathbf{X}_t | \sigma_t, \mathbf{X}_{t-1}^{IBOV}].$$

Portanto, a modelagem desta dinâmica conjunta advém da inclusão de uma variável exógena determinística \mathbf{X}_{t-1}^{IBOV} , representando uma série temporal do vetor de coeficientes $\{\mathbf{X}_{(0)}^{IBOV}, \mathbf{X}_{(1)}^{IBOV}, \dots, \mathbf{X}_{(t-1)}^{IBOV}\}$, previamente estimados pelo modelo KF. A equação de estado é definida como:

$$\mathbf{X}_t = \mathbf{a} + \mathbf{U}\mathbf{X}_{t-1}^{Eq} + \mathbf{B}(\mu_{IBOV} - \mathbf{X}_{t-1}^{IBOV}) + \mathbf{u}_t, \quad (3.15)$$

em que \mathbf{X}_t^{Eq} e \mathbf{X}_t^{IBOV} são os vetores dos coeficientes de cada ação individual e do Índice IBOVESPA, respectivamente; μ_{IBOV} é o vetor 4×1 das médias de longo prazo e \mathbf{B} é a matriz diagonal 4×4 das taxas de reversão à média dos coeficientes do IBOVESPA.

De acordo com as equações (3.3) e (3.8), considerando as estimativas iniciais dos fatores \mathbf{X}_0 e matriz de covariância do erro \mathbf{P}_0 e tomando $\mathbf{b} = \mathbf{B}\mu_{IBOV}$, $\mathbf{Z} = \mathbf{I} - \mathbf{B}$, obtém-se as estimativas *a priori* e *a posteriori*, como segue:

$$\hat{\mathbf{X}}_{t|t-1} = \mathbf{a} + \mathbf{U}\hat{\mathbf{X}}_{t-1|t-1} + \mathbf{b} + \mathbf{Z}\mathbf{X}_{t-1}^{IBOV}. \quad (3.16)$$

A partir das equações (3.4) e (3.11) e sabendo que a varável exógena é determinística e, portanto, sua matriz de covariância do erro $E[\tilde{\mathbf{X}}_{t-1}^{IBOV} \tilde{\mathbf{X}}_{t-1}^{IBOV^T}] = 0$, tem-se:

$$\begin{aligned} \mathbf{P}_{t|t-1} &= \mathbf{U}\mathbf{P}_{t-1|t-1}\mathbf{U}^T + \mathbf{V} \Rightarrow \\ \mathbf{P}_{t|t-1} &= \mathbf{U}[\mathbf{P}_{t-1|t-2}(\mathbf{I} - \mathbf{K}_{t-1}\mathbf{M}_{t-1})]\mathbf{U}^T + \mathbf{V}. \end{aligned}$$

Tomando o valor de \mathbf{K}_{t-1} e substituindo na equação acima, obtém-se:

$$\begin{aligned} \mathbf{K}_{t-1} &= \mathbf{P}_{t-1|t-2}\mathbf{M}_{t-1}^T(\mathbf{M}_{t-1}\mathbf{P}_{t-1|t-2}\mathbf{M}_{t-1}^T + \mathbf{R})^{-1} \text{ e} \\ \mathbf{P}_{t|t-1} &= \mathbf{U}[\mathbf{P}_{t-1|t-2} - \mathbf{P}_{t-1|t-2}\mathbf{M}_{t-1}\mathbf{K}_{t-1}]\mathbf{U}^T + \mathbf{V}. \end{aligned} \quad (3.17)$$

O cálculo da matriz de covariância do erro *a priori* $\mathbf{P}_{t|t-1}$ utiliza em sua fórmula o valor da matriz de covariância do erro *a priori* de lag anterior $\mathbf{P}_{t-1|t-2}$. Portanto, sabendo o valor de $\mathbf{P}_{t|t-1}$, tem-se o \mathbf{K}_t , equação (3.10), e a estimativa *a posteriori*:

$$\hat{\mathbf{X}}_{t|t} = \hat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t(\sigma_t - \mathbf{M}_t^T \hat{\mathbf{X}}_{t|t-1}). \quad (3.18)$$

Como resultado, obtém-se, de forma recursiva, a projeção do valor esperado $\hat{\mathbf{X}}_{t+1}$ e a matriz de covariância \mathbf{P}_{t+1} para o instante subsequente. Para cada instante de tempo t no período fora da amostra, tem-se a densidade preditiva em $t + 1$ dos fatores que direcionam o *skew* de volatilidade $\mathbf{X}_{t+1} \sim N(\hat{\mathbf{X}}_{t+1}, \mathbf{P}_{t+1})$.

3.3.2 Método de Estimação

De forma geral, o procedimento de calibração consiste em encontrar o conjunto de parâmetros que minimizam uma função perda entre a volatilidade prevista pelo modelo e os dados de mercado, enquanto satisfazem algumas restrições adicionais, como valores específicos para os parâmetros, ausência de arbitragem, etc. Trata-se, via de regra, de problemas de otimização não-linear com restrições de desigualdade também não-lineares, o que requer alta capacidade computacional.

Como mencionado no item 3.2, por definição, os resíduos de medição do algoritmo de KF são serialmente não correlacionados, conforme a demonstração a seguir, com base nos cálculos de Gibbs (2011).

Sabendo que o erro da estimativa *a priori* no período k é:

$$\begin{aligned}\tilde{\mathbf{x}}_{k|k-1} &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \\ \tilde{\mathbf{x}}_{k|k-1} &= \mathbf{F}_k \tilde{\mathbf{x}}_{k-1|k-1} + \omega_k,\end{aligned}$$

e o erro da estimativa *a posteriori*, aplicando a equação (3.6), será:

$$\begin{aligned}\tilde{\mathbf{x}}_{k|k} &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k} \\ \tilde{\mathbf{x}}_{k|k} &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k \tilde{\mathbf{y}}_k \\ \tilde{\mathbf{x}}_{k|k} &= \tilde{\mathbf{x}}_{k|k-1} - \mathbf{K}_k (\mathbf{G}_k \tilde{\mathbf{x}}_{k|k-1} + \epsilon_k) \\ \tilde{\mathbf{x}}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{G}_k) \tilde{\mathbf{x}}_{k|k-1} - \mathbf{K}_k \epsilon_k.\end{aligned}$$

Logo,

$$\begin{aligned}E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] &= E[(\mathbf{G}_k \tilde{\mathbf{x}}_{k|k-1} + \epsilon_k)(\mathbf{G}_{k-1} \tilde{\mathbf{x}}_{k-1|k-2} + \epsilon_{k-1})^T] \\ E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] &= E[(\mathbf{G}_k (\mathbf{F}_k \tilde{\mathbf{x}}_{k-1|k-1} + \omega_k) + \epsilon_k)((\mathbf{G}_{k-1} \tilde{\mathbf{x}}_{k-1|k-2} + \epsilon_{k-1})^T)]\end{aligned}$$

$$\begin{aligned}E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] &= E[(\mathbf{G}_k \mathbf{F}_k ((\mathbf{I} - \mathbf{K}_{k-1} \mathbf{G}_{k-1}) \tilde{\mathbf{x}}_{k-1|k-2} - \mathbf{K}_{k-1} \epsilon_{k-1}) \\ &\quad + \mathbf{G}_k \omega_k + \epsilon_k)(\mathbf{G}_{k-1} \tilde{\mathbf{x}}_{k-1|k-2} + \epsilon_{k-1}^T)].\end{aligned}$$

Seguindo as premissas de que $E[\epsilon_k] = 0$, $E[\omega_k \epsilon_k^T] = 0$, $E[\tilde{\mathbf{x}}_k \epsilon_k^T] = 0$ e tomando o valor esperado da equação acima, tem-se:

$$E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] = \mathbf{G}_k \mathbf{F}_k (-\mathbf{K}_{k-1} (\mathbf{G}_{k-1} \mathbf{P}_{k-1,k-2} \mathbf{G}_{k-1}^T + \mathbf{R}_{k-1}) + \mathbf{P}_{k-1,k-2} \mathbf{G}_{k-1}^T).$$

Substituindo o valor de \mathbf{K}_{k-1} por $\mathbf{P}_{k-1,k-2} \mathbf{G}_{k-1}^T (\mathbf{G}_{k-1} \mathbf{P}_{k-1,k-2} \mathbf{G}_{k-1}^T + \mathbf{R}_{k-1})^{-1}$, mostra-se que:

$$E[\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_{k-1}^T] = 0.$$

Portanto, o fato de os resíduos $\tilde{\mathbf{y}}_k$ serem independentes e identicamente distribuídos (ruído branco), assumindo uma distribuição gaussiana, com uma matriz de covariância \mathbf{C}_k conhecida, possibilita a implementação do algoritmo de máxima verossimilhança (*maximum likelihood estimation* - *MLE*) ao facilitar a tratabilidade matemática e computacional do produto das funções densidade de probabilidade (PDF) conjunta dos dados observados; além de permitir a validação do modelo por meio da aplicação de testes, como o teste do resíduo padronizado.

Por definição, o MLE identifica os parâmetros desconhecidos, tratados como não randômicos ou estáticos, que maximizam a probabilidade do processo descrito pelo modelo obter variáveis aleatórias ($\mathbf{y}_i, i = 1, \dots, m$) que sejam aderentes aos dados observados. Logo, a função verossimilhança (L) é definida como a distribuição de probabilidade conjunta de \mathbf{y}_i dado um conjunto de parâmetros $\mathbf{x}_i, i = 1, \dots, n$, como segue:

$$\mathbf{L}(\mathbf{y}|\mathbf{x}) = p(\mathbf{Y}|\mathbf{x})(\mathbf{y}|\mathbf{x}) = p(\mathbf{Y}|\mathbf{x})(y_1, \dots, y_n|x_1, \dots, x_n). \quad (3.19)$$

Assumindo uma distribuição gaussiana:

$$p(\mathbf{Y})\mathbf{y} = \frac{1}{(2\pi)^{m/2}|\mathbf{P}|^{1/2}} \exp(-\frac{1}{2}(\mathbf{y} - \bar{\mathbf{y}})^T \mathbf{P}_{yy}^{-1}(\mathbf{y} - \bar{\mathbf{y}})),$$

em que $|\mathbf{P}_{yy}|$ é o determinante de \mathbf{P}_{yy} , equivalente ao produto das variâncias na diagonal da matriz.

E aplicando a função logarítmica na função L , tem-se:

$$\ln \mathbf{L}(\mathbf{y}|\mathbf{x}) = -\frac{1}{2}[m \ln(2\pi) + \ln|\mathbf{P}_{yy}| + (\mathbf{y} - \bar{\mathbf{y}})^T \mathbf{P}_{yy}^{-1}(\mathbf{y} - \bar{\mathbf{y}})], \quad (3.20)$$

sendo \mathbf{P}_{yy} e $\bar{\mathbf{y}}$ computados a partir dos valores de \mathbf{x} . Logo, o ML estima o \mathbf{x} de modo que a função seja máxima, ou de forma equivalente, iguala-se o gradiente da função a zero $\frac{\partial \ln \mathbf{L}(\mathbf{y}|\mathbf{x})}{\partial \mathbf{x}} = 0$.

3.4 Modelo Benchmark

A performance do método do Filtro de Kalman para descrever a dinâmica do *skew* de volatilidade é comparada ao modelo proposto por Gonçalves e Guidolin (2006), denominado, neste trabalho, como Modelo II, cujo objetivo é explorar a possibilidade de predição da volatilidade implícita ao incluir no modelo as propriedades da série temporal dos fatores que conduzem o formato da IVS. A abordagem se divide em 2 estágios:

- *Fitting* das IVS diárias (*cross sectional*) por meio do modelo de mínimos quadrados ordinários (MQO) com variáveis explicativas - *moneyness* e maturidade - não lineares, como segue:

$$\log \sigma_t = \beta_0 + \beta_1 \mathbf{M}_t + \beta_2 \mathbf{M}_t^2 + \beta_3 \tau_t + \beta_4 \mathbf{M}_t \tau_t + \epsilon_t, \quad (3.21)$$

em que a variável *moneyness* $\mathbf{M} = \frac{\log[\frac{K_t}{\exp(r\tau_t)S}]}{\sqrt{(\tau_t)}}$, ϵ_t é o termo do erro e τ_t é o prazo até a maturidade como fração do ano (360 dias).

Como resultado, para cada dia da amostra, obtém-se o vetor estimado dos coeficientes $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4)'$.

- Implementação do modelo VAR(p) nos coeficientes estimados no item anterior, de modo a capturar a variação no tempo da IVS,

$$\hat{\beta}_t = \mu + \sum_{j=1}^p \phi \hat{\beta}_{t-j} + \mathbf{u}_t, \quad (3.22)$$

em que $\mathbf{u}_t \sim N(0, \mathbf{\Omega})$ sendo i.i.d.

A escolha da ordem p do modelo VAR baseou-se no critério de informação Bayesiano (BIC), considerando p máximo igual a 12, e variando-o para cada janela do período de estimação. Portanto, o modelo VAR produz projeções diárias dos coeficientes estimados $\hat{\beta}_{t+1}$, o que permite projetar a IVS para $t + 1$, além do preço das opções pela fórmula de BS.

O capítulo a seguir descreve a metodologia utilizada para implementação e calibração dos modelos, bem como a obtenção e tratamento dos dados.

4 Metodologia

Este capítulo descreve o processo de implementação do modelo de Filtro de Kalman na sua forma linear, bem como os ajustes aplicados no modelo II, proposto por [Gonçalves e Guidolin \(2006\)](#), utilizado aqui como *benchmark*. Além da forma de tratamento dos dados no mercado brasileiro de opções.

4.1 Tratamento dos Dados

O conjunto de dados consiste nos preços diários de fechamento das opções do Índice IBOVESPA e das opções de compra (*call*) de Petrobras PN (PETR4) negociadas na B3 no período de 02 de Janeiro de 2015 a 28 de Dezembro de 2018. A utilização apenas dos dados da *call* se deu em função do potencial não sincronismo entre os preços de fechamento, definidos na B3 como último negócio, das *calls* e *puts*, que poderia distorcer a projeção dos *skews*. Para implementação das estratégias com os preços projetados das opções, utilizamos o preço de uma *put* sintética a partir da paridade *call-put*, como segue:

$$Put_t = Call_t - S_t + Kexp(-r_t\tau),$$

sendo S_t o preço do ativo objeto, r_t a taxa pré¹ interpolada e τ a maturidade para cada período t .

Cabe ressaltar que a escolha pela utilização das *calls* fundamentou-se no maior numero e volume de negócios, ao incluir as opções americanas no conjunto de dados. Neste caso, no entanto, tratamos tais opções como se fossem do tipo Européia, já que não houve distribuição de dividendos no período total de estimação. De acordo com [Hull \(2005\)](#), nunca é melhor exercer uma opção de compra americana, sobre uma ação que não paga dividendos, antes da data de expiração, pois esta sempre vale, pelo menos, o valor que seria ganho com o exercício, e geralmente vale mais.

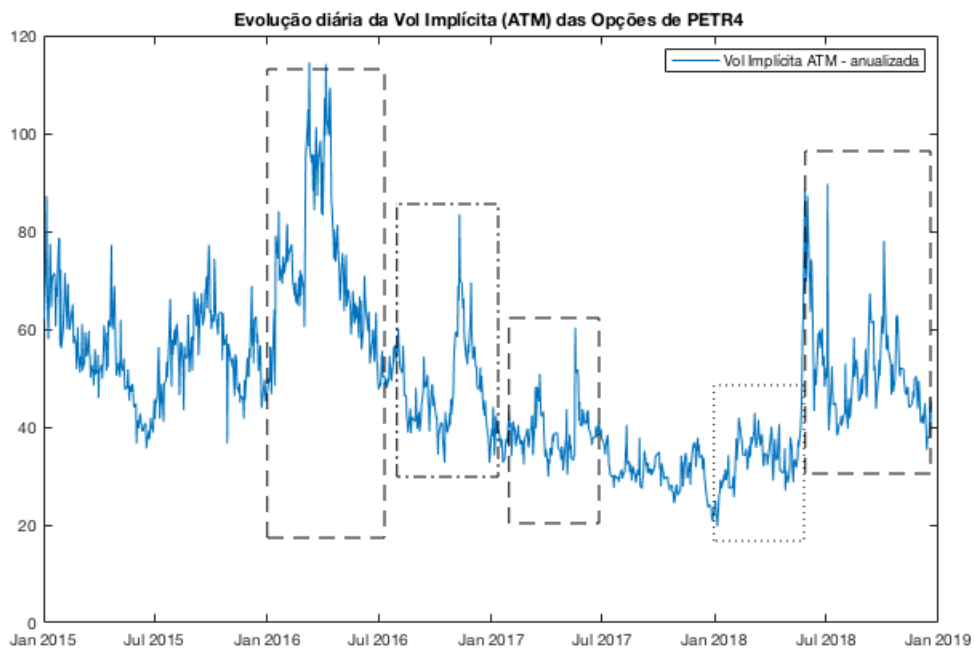
Como o objetivo de trabalho é modelar a dinâmica do *skew* de volatilidade, ou seja, com maturidade fixa e, considerando a maior liquidez dos vértices mais curtos, utilizou-se os contratos com vencimento do mês subsequente à data de negociação, com rolagem para o próximo contrato quando restarem 6 dias úteis para o vencimento, de modo a preservar a liquidez e conteúdo informacional dos preços das opções. O fator *moneyness* utilizado na equação de medida foi gerado a partir dos dados diários dos preços à vista de fechamento (PETR4 e Índice IBOVESPA) e a taxas de juros pré fixadas obtidas por meio dos dados

¹ Utilizamos o *swap* PróxDI, com vencimentos fixos, divulgados pelo terminal Bloomberg, como *proxy* do contrato de DI Futuro (Depósito Interfinanceiro).

diários dos *swaps* PróxDI (Depósito Interbancário) com prazos fixos de vencimento e interpoladas linearmente para as respectivas maturidades das séries de opções.

Utilizou-se um ano para estimação e o primeiro semestre do ano subsequente para avaliar a performance de predição do modelo, escolhendo as janelas de acordo com o nível de volatilidade do ativo subjacente, de modo a avaliar os modelos propostos em períodos mais voláteis, conforme Figura 1.

Figura 1 – Evolução diária da Volatilidade Implícita ATM como *proxy* da volatilidade de PETR4.



Os critérios de exclusão e filtragem das séries de opções são descritos a seguir:

- Exclusão dos contratos com menos de 6 dias uteis para o vencimento, considerando que o valor temporal da opção reduz em uma velocidade maior à medida que se aproxima do vencimento e, portanto, eleva a sensibilidade da volatilidade ao não sincronismo de preço das opções (DUMAS; FLEMING; WHALEY, 1998). Além de reduzir substancialmente a liquidez dos contratos, contendo menos informação relevante para a projeção da IVS.
- Exclusão das opções com *strikes* extremos (*deep in the money* e *out of the money*), cujo *moneyness* absoluto é maior do que $0,5 - |(\frac{Strike}{Forward}) - 1| > 0,5$ (CONT; FONSECA, 2002).

Hentschel (2003) explora em seu trabalho a sensibilidade do cálculo da IV ao erro de medição, como por exemplo o *bid/ask spread* e não sincronismo entre os preços das opções e os ativos subjacentes, por meio da análise da formula de BS. Assumindo o

erro de medição i.i.d. e nulo, o preço teórico da opção (C) com os dados de mercado será igual ao preço observado \bar{C} para uma determinada volatilidade $\bar{\sigma}$:

$$C(\bar{S}, \bar{K}, \bar{r}, \bar{t}, \bar{\sigma}) = \bar{C}. \quad (4.1)$$

Se os preços observados contêm erros de medição, tem-se $\bar{C} = C + dc$.

Logo: $C(S + dS, K + dK, r + dr, t + dt, +d\sigma) = C + dc$

Diferenciando (4.1) em relação à volatilidade implícita, obtém-se:

$$\frac{\partial C}{\partial S} \frac{dS}{d\sigma} + \dots + \frac{\partial C}{\partial \sigma} \frac{d\sigma}{d\sigma} = \frac{dC}{d\sigma} \Rightarrow d\sigma = \left(\frac{\partial C}{\partial \sigma} \right)^{-1} [dc - \frac{\partial C}{\partial S} dS - \dots - \frac{\partial C}{\partial t} dt]. \quad (4.2)$$

Sabendo que a função C (preço teórico da *call*) é monotônica em relação ao σ e, portanto, $\left(\frac{\partial C}{\partial \sigma} \right)^{-1} = \frac{\partial \sigma}{\partial C}$, deduz-se:

$$-\frac{\partial C}{\partial x_i} \left(\frac{\partial C}{\partial \sigma} \right)^{-1} = \frac{\partial \sigma}{\partial x_i} \Rightarrow d\sigma = \frac{\partial \sigma}{\partial x'} dx, \quad (4.3)$$

sendo $x = (C, S, K, r, t)'$.

Neste caso, considerando, como exemplo, o fator de risco vega originado a partir da fórmula BS: $\frac{\partial C}{\partial \sigma} = S\sqrt{t}\phi(z)$, sendo $\phi(\cdot)$ a função densidade de probabilidade normal e

$$z = \frac{\ln\left(\frac{S}{Ke^{-rt}}\right)}{\sigma\sqrt{t}} + \frac{1}{2}\sigma\sqrt{t}.$$

Evidencia-se, com isso, que opções *DITM* e *DOTM* apresentam $|z|$ grande e o $\phi(z)$ tende a zero, ou seja, $\left(\frac{\partial C}{\partial \sigma} \right)^{-1} = \frac{\partial \sigma}{\partial C}$ aumenta significativamente.

Portanto, pequenos erros de medida nos *inputs* de C produzem grandes erros na volatilidade implícita ($d\sigma$), tornando o processo de calibragem dos parâmetros e projeção menos robusto.

- Exclusão de opções que violam a restrição de arbitragem (CHEN, 1997), para a qual $C \geq \max[0, S - PV(K)]$.

Se $C < S - PV(K) = -C + S - PV(K) > 0$ equivale a um fluxo de caixa positivo em $t + 0$ oriundo da venda do *spot* e um fluxo zero ($S_T > K$) ou positivo ($S_T < K$) no vencimento. Com isso, configura-se uma arbitragem derivada do ganho financeiro sem possibilidade de perda.

Após a aplicação dos critérios de exclusão, obteve-se em média 16 séries (*strikes*) de opção de compra negociadas para cada período.

4.2 Implementação dos modelos KF e KF_IBOV

A serie temporal dos *skews* de volatilidade fornece informações sobre como a IV varia em função dos diferentes *strikes* para cada dia (*cross sectional*), e permite a modelagem da evolução estocástica dos fatores (variáveis não observáveis) que direcionam os *skews* no tempo. Para tanto, utiliza-se um modelo de espaço de estado a fim de incorporar toda a informação disponível para atualização do sistema dinâmico (HARVEY, 1989).

A despeito das inconsistências teóricas do modelo de BS, utilizou-se este método de apreçamento apenas para conversão dos preços das opções em volatilidade implícita para cada período t , sem prejuízo para a implementação dos modelos, cujo propósito é avaliar a performance de predição.

Para cada período t , tem-se a equação de medida (ME) como um vetor $n \times 1$ com n *strikes*, sendo σ_t uma função do *moneyness* $\{m^i\}_{i=0,\dots,3}$. Como *proxy* de um modelo linear e discreto, pode-se interpretar estes fatores como as variáveis explicativas de uma regressão linear, cujos coeficientes $\{\mathbf{X}_i\}_{i=1,\dots,4}$ direcionam o *skew* de volatilidade². Considerando as equações (3.13) e (3.15), obtém-se as equações de medição e de propagação, respectivamente:

$$\sigma_t = \mathbf{M}_t \mathbf{X}_t + \epsilon_t, \quad (4.4)$$

$$\mathbf{X}_t = \mathbf{a} + \mathbf{U} \mathbf{X}_{t-1} + \mathbf{u}_t, \quad (4.5)$$

sendo $\mathbf{a} = \mathbf{A}\mu$ um vetor 4×1 e $\mathbf{U} = \mathbf{I} - \mathbf{A}$ uma matriz diagonal 4×4 .

A característica sequencial do modelo KF pressupõe que, para $t = 0$, os valores iniciais $\hat{\mathbf{X}}_t$ e \mathbf{P}_t (antes da incorporação dos dados observados $[\sigma_t]$), sendo $E_{t-1}[\mathbf{X}_t]$ e $Var_{t-1}[\mathbf{X}_t]$, respectivamente, sejam conhecidos, assumindo uma distribuição gaussiana multivariada. A metodologia utilizada para estimar estes valores será demonstrada mais adiante.

Como primeiro passo, o processo de filtragem para gerar a estimativa *a posteriori* do $\hat{\mathbf{X}}_t$ para 1 dia, considerando os coeficientes \mathbf{X}_t e a matriz de covariância \mathbf{P}_t , incorpora os dados observados σ_t . De acordo com a equação (3.18), obtém-se a estimativa *a posteriori*:

$$\hat{\mathbf{X}}_{t[EstPosteriori]} = \hat{\mathbf{X}}_t + \mathbf{K}_t(\sigma_t - \mathbf{M}_t^T \hat{\mathbf{X}}_t), \quad (4.6)$$

em que $\mathbf{K}_t = \mathbf{P}_t \mathbf{M}_t^T (\mathbf{M}_t \mathbf{P}_t \mathbf{M}_t^T + \mathbf{R})^{-1}$.

Em seguida, a estimativa aprimorada (*a posteriori*) $\hat{\mathbf{X}}_{t[EstPosteriori]}$ propaga-se por meio do sistema dinâmico, equações (3.16) e (3.17) de modo a gerar uma projeção ótima $\hat{\mathbf{X}}_{t+1}$ e \mathbf{P}_{t+1} para o período subsequente:

$$\hat{\mathbf{X}}_{t+1} = \mathbf{a} + \mathbf{U} \hat{\mathbf{X}}_{t[EstPosteriori]}, \quad (4.7)$$

² Foram analisadas diferentes formas funcionais (mudando o número de variáveis explicativas) dos modelos de regressão, apontando que o polinômio cubico apresenta menor RMSE e maior R^2 ajustado.

$$\mathbf{P}_{t+1} = \mathbf{U}[\mathbf{P}_t - \mathbf{P}_t \mathbf{M}_t \mathbf{K}_t] \mathbf{U}^T + \mathbf{C} \mathbf{C}', \quad (4.8)$$

sendo $\mathbf{C} \mathbf{C}' = \mathbf{V}$ a diagonalização da matriz de covariância do ruído branco da equação de estado.

Para a modelagem conjunta do modelo KF_IBOV, considera-se a inclusão da série temporal do vetor de coeficientes tanto no período de estimação como no período fora da amostra $\{\mathbf{X}_{(0)}^{IBOV}, \mathbf{X}_{(1)}^{IBOV}, \dots, \mathbf{X}_{(t-1)}^{IBOV}\}$, como uma variável exógena determinística, estimados pelo algoritmo KF, cujo objetivo é aprimorar a performance do estimativa *a posteriori*:

$$\hat{\mathbf{X}}_{t|[\text{EstPosteriori}]} = E[\mathbf{X}_t | \sigma_t, \mathbf{X}_{(t-1)}^{IBOV}].$$

De acordo com a equação (4.7):

$$\hat{\mathbf{X}}_{t+1} = \mathbf{a} + \mathbf{U} \hat{\mathbf{X}}_{t|[\text{EstPosteriori}]} + \mathbf{b} + \mathbf{Z} \mathbf{X}_{(t-1)}^{IBOV}, \quad (4.9)$$

sendo $\mathbf{X}_{(t)}^{IBOV}$ os vetores dos coeficientes do Índice IBOVESPA; μ_{IBOV} é o vetor 4×1 das médias de longo prazo e \mathbf{B} é a matriz diagonal 4×4 das taxas de reversão à média. Toma-se $\mathbf{b} = \mathbf{B} * \mu_{IBOV}$ e $\mathbf{Z} = \mathbf{I} - \mathbf{B}$.

Em suma, para cada período t , aprimora-se a estimativa (*a priori*) $\hat{\mathbf{X}}_t$ ao incorporar os dados observáveis do *skew* de volatilidade, equação (4.6), obtendo, por meio da equação de propagação, a projeção ótima de 1 dia do valor esperado $\hat{\mathbf{X}}_{t+1}$ e a matriz de covariância \mathbf{P}_{t+1} . Assumindo uma distribuição gaussiana dos coeficiente, por definição do modelo, obtém-se para cada período a densidade preditiva $\mathbf{X}_{t+1} \sim N(\hat{\mathbf{X}}_{t+1}, \mathbf{P}_{t+1})$.

4.3 Estimação dos Modelos

Conforme descrito na subseção 3.3.2, por definição, o modelo de Filtro de Kalman pressupõe que as variáveis observáveis σ_t seguem uma distribuição gaussiana multivariada e os ruídos de medição são independentes e identicamente distribuídos. Com isso, realizou-se a estimação dos parâmetros por meio do algoritmo de máxima verossimilhança (MLE).

De acordo com equação (3.20), excluindo o termo constante $[-\frac{1}{2} \ln(2\pi)]$, sem prejuízo para o calculo final, e considerando a matriz de covariância dos resíduos de medição igual a $\mathbf{M}_t \mathbf{P}_t \mathbf{M}_t^T + \mathbf{R}_t$, equação (3.7), obtém-se para cada período t :

$$\ln \mathbf{L}(\sigma_t | [\mathbf{X}_i]_{i=1, \dots, 4}) = \ln |\mathbf{M}_t \mathbf{P}_t \mathbf{M}_t^T + \sigma_\epsilon^2 \mathbf{I}| + (\sigma_t - \bar{\sigma}_t)^T (\mathbf{M}_t \mathbf{P}_t \mathbf{M}_t^T + \sigma_\epsilon^2 \mathbf{I})^{-1} (\sigma_t - \bar{\sigma}_t), \quad (4.10)$$

em que $\bar{\sigma}_t = \mathbf{M}_t^T \hat{\mathbf{X}}_{t|t-1}$, ou seja, $\bar{\sigma}_t$ em função da medida *moneyness* e da estimativa *a priori*, e $\sigma_\epsilon^2 \mathbf{I} = \mathbf{R}$, pois assumiu-se que o ruído da equação de medida é *i.i.d.*

Denominando $\ln \mathbf{L}(\sigma_t | [\mathbf{X}_i]_{i=1, \dots, 4})$ como *output*, que é o valor escalar da função para cada período t , o algoritmo de otimização busca minimizar a somatória do *output* para todo o período de estimação por meio do algoritmo do Filtro de Kalman, variando o conjunto de parâmetros encapsulado nas variáveis de estado em $\{\mathbf{X}_i\}_{i=1, \dots, 4}$:

$$\text{output} = \sum_{t=1}^N \ln \mathbf{L}(\sigma_t | \Theta),$$

com $\Theta = [\mathbf{a}, \mathbf{U}, \mathbf{C}, \sigma_\epsilon^2]$, N =número de períodos e $\Theta_{IBOVESPA} = [\mathbf{b}, \mathbf{Z}, \mathbf{C}, \sigma_\epsilon^2]$, sendo os coeficientes exógenos do IBOVESPA.

Definimos a matriz dos coeficientes de reversão à média \mathbf{A} como diagonal, entendendo ser razoável supor que não há dependência entre as taxas de reversão dos diferentes coeficientes do *skew*. Logo, há redução do número total de parâmetros a serem estimados para 19, sendo 4 de $\mathbf{U} = \mathbf{I} - \mathbf{A}$ (diagonal), 4 de $\mathbf{a} = \mathbf{A}\mu$ (vetor), 10 de \mathbf{C} (matriz triangular de $\mathbf{V} = \mathbf{C}\mathbf{C}'$ 4×4) e 1 da volatilidade do erro de medida σ_ϵ^2 .

Seguindo o trabalho de Bedendo e Hodges (2009), os valores de $\hat{\mathbf{X}}_t$ e \mathbf{P}_t , com $t = 1$, foram estimados, respectivamente, como os coeficientes de uma regressão linear com o mesmo polinômio cúbico aplicado aos dados observados do primeiro *skew*; e seguindo a dinâmica do modelo II (*benchmark*) (GONÇALVES; GUIDOLIN, 2006), aplicamos um modelo de regressão nos *skews* diários, produzindo uma série temporal destes coeficientes. Em seguida, estimamos um modelo VAR(1) visando analisar a evolução no tempo destes fatores estimados³.

Como resultado, conforme a equação (3.22) para $p = 1$, estima-se VAR(1):

$$\hat{\beta}_t = \mathbf{a} + \mathbf{U}\hat{\beta}_{t-1} + \mathbf{u}_t, \quad (4.11)$$

com $\mathbf{u}_t \sim N(0, \mathbf{\Omega})$ sendo i.i.d.

Obtendo $\mathbf{A} = \mathbf{I} - \mathbf{U}$ e $\mathbf{V} = \mathbf{C}\mathbf{C}' = \mathbf{\Omega}$ (matriz de covariância do ruído do processo).

Para simplificação, definimos \mathbf{P}_1 como a matriz de covariância assintótica do processo OU, equação (4.5), que por definição é $\mathbf{P}_1 = \mathbf{C}(\mathbf{A} + \mathbf{A}')\mathbf{C}'$ (ver APÊNDICE A para maiores detalhes).

Como *proxy* do ruído da equação de medida $\sigma_\epsilon^2 \mathbf{I}$, obteve-se a variância $\sigma_{\epsilon_1}^2$ (escalar) da série temporal dos resíduos $[\sigma_t - \bar{\sigma}_t]$ do y da regressão.

Neste caso, visto que a execução do algoritmo MLE do Matlab requer o *input* de valores iniciais dos parâmetros $[\Theta_0]$, utilizamos os valores estimados acima $[\mathbf{a}, \mathbf{U}, \mathbf{C}, \sigma_{\epsilon_1}^2 \mathbf{I}]$ e, por conseguinte, houve significativa redução de tempo de execução no processo de otimização do $[\Theta]$.

³ Conforme demonstração na seção 3.3.1, o processo OU discreto da equação de estado pode ser representado pelo VAR(1), estimado, neste caso, com uma matriz diagonal dos coeficientes de modo a replicar as características da equação de estado do modelo de Filtro de Kalman.

Para tanto, deve-se aplicar algumas condições de contorno para execução do processo de otimização numérica, como segue:

- Seja \mathbf{U} a matriz dos coeficientes do processo, equação (4.5), e seja γ os autovalores de \mathbf{U} , faz-se necessário $|\gamma| < 1$ para que o processo seja estacionário. Logo, dado que \mathbf{U} é diagonal, impõe-se a restrição de que o módulo da diagonal de \mathbf{U} seja menor do que 1.
- Seja $\mu = (\mu_{X_1}, \mu_{X_2}, \mu_{X_3}, \mu_{X_4})'$ o vetor 4×1 da média de longo prazo dos coeficientes, tem-se as seguintes restrições:

Coefficiente $[X]$	Condição
X_1 [Nível]	$\mu_{X_1} \geq 0$
X_2 [Inclinação]	$-\infty < \mu_{X_2} < \infty$
X_3 [Curvatura]	$\mu_{X_3} > 0$
X_4 [Assimetria]	$-\infty < \mu_{X_4} < \infty$

Ressalta-se que consideramos o padrão empírico de que as volatilidades de *strikes* extremos são maiores para impor a a restrição em μ_{X_3} . Sabendo que $\mathbf{a} = \mu\mathbf{A}$ e $\mathbf{A} \geq 0$, pois é a taxa (velocidade) de reversão, é necessário que \mathbf{a} satisfaça as condições de μ .

- Faz-se necessário que a matriz simétrica de covariância \mathbf{V} (4×4) seja positiva definida (ou semi definida). Para tanto, os autovalores de \mathbf{V} devem ser maiores do que -1 e menores do que 1. Nestas condições, pode-se decompor ou diagonalizar a matriz como $\mathbf{V} = \mathbf{C}\mathbf{C}'$, estimando apenas a matriz \mathbf{C} . A fim de reforçar que a condição dos auto-valores seja satisfeita, inserimos uma função no algoritmo MLE de modo a assegurar que os elementos da diagonal de \mathbf{C} sejam sempre não negativos.
- A variância do erro de medida (escalar) deve ser não negativa - $\sigma_{\epsilon_1}^2 \geq 0$.

4.4 Implementação do Modelo II

De acordo com a equação (3.21), e como resultado da análise do R^2 e RMSE, obteve-se o polinômio cúbico aplicado às IVS diárias:

$$\log \sigma_t = \beta_0 \mathbf{M}_t^0 + \beta_1 \mathbf{M}_t + \beta_2 \mathbf{M}_t^2 + \beta_3 \mathbf{M}_t^3 + \epsilon_t, \quad (4.12)$$

obtendo o vetor de estado por meio do modelo de mínimos quadrados ordinários (MQO)⁴ $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)'$ para cada período t .

⁴ O modelo de mínimos quadrados generalizados (MQG), indicado quando há heteroscedasticidade na variância do erro, foi aplicado, porém as estatísticas descritivas (RMSE e R^2) apresentaram pequena diferença. Neste caso, por simplificação, optou-se pelo MQO.

Como ajuste ao modelo desenvolvido por Gonçalves e Guidolin (2006), consideramos apenas a medida *moneyness* como variável explicativa, uma vez que este trabalho utiliza maturidade fixa para cada série de opções.

Os estimadores dos parâmetros do modelo VAR(p), seguindo a equação (3.22), são obtidos pelo método de máxima verossimilhança condicional, utilizando o algoritmo MLE de maximização numérica do Matlab, supondo normalidade de \mathbf{u}_t . A escolha da ordem p do processo baseou-se no critério de informação Bayesiano (BIC), considerando p máximo igual a 10, e variando-o para cada janela do período de estimação. Neste caso, a pré amostra igual a 10 permaneceu constante a fim de manter a coerência na comparação dos modelos. O modelo VAR(p) é descrito a seguir:

$$\hat{\beta}_t = \mu + \sum_{j=1}^p \phi \hat{\beta}_{t-j} + \mathbf{u}_t, \quad (4.13)$$

em que $\mathbf{u}_t \sim N(0, \mathbf{\Omega})$

Portanto, o modelo VAR produz projeções diárias dos coeficientes estimados $\hat{\beta}_{t+1}$, o que permite projetar a IVS para $t + 1$, e os preços das opções pela fórmula de BS.

4.5 Método de análise da performance de predição

4.5.1 Previsão do *skew* pelos modelos KF e KF_IBOV

Para $t = 1$, no período fora da amostra, os valores iniciais necessários para executar o algoritmo de Filtro de Kalman ($\hat{\mathbf{X}}_1$ e \mathbf{P}_1) são oriundos do último dia N da janela de estimação, sendo $\hat{\mathbf{X}}_1 = \hat{\mathbf{X}}_{N|[\text{EstPosterior}]}$.

De forma sequencial, propaga-se a estimativa a posteriori, equações (4.7) e (4.9), obtendo a projeção ótima dos coeficientes para $(t + 1)$.

Obtém-se, desta forma, as volatilidades implícitas por meio da equação de medida (4.4), cujos *inputs* são os coeficientes estimados e medida *moneyness* projetada:

$$\hat{\sigma}_{t+1} = \mathbf{M}_{t+1}^T \hat{\mathbf{X}}_{t+1}. \quad (4.14)$$

4.5.2 Previsão do *skew* pelo modelo II

Considerando a equação (4.13), com VAR(p), após a estimação dos parâmetros, e sendo $\mathcal{F}_t = [\hat{\beta}_s : s \leq t]$ o σ -álgebra com as informações disponíveis até o instante t , obtém-se, de forma recursiva, a projeção dos coeficientes para o horizonte h :

$$E[\hat{\beta}_{t+h} | \mathcal{F}_t] = \mu + \phi_1 E[\hat{\beta}_{t+h-1} | \mathcal{F}_t] + \dots + \phi_p E[\hat{\beta}_{t+h-p} | \mathcal{F}_t],$$

sendo $E[\mathbf{u}_{t+h} | \mathcal{F}_t] = 0$, para todo $h > 0$.

Para $h = 1$, temos:

$$\hat{\beta}_t(1) = \mu + \phi_1 \hat{\beta}_t + \dots + \phi_p \hat{\beta}_{t-p+1}. \quad (4.15)$$

Desta forma, as IVs correspondentes são obtidas por meio do polinômio da regressão da equação (4.12):

$$\hat{\sigma}_{t+1} = \exp(\mathbf{M}_{t+1}^T \hat{\beta}_{t+1}). \quad (4.16)$$

Para os 3 modelos, a medida *moneyness* \mathbf{M}_{t+1} é calculada com base nos *strikes* da curva t imediatamente anterior. Para tanto, conforme abordado por Gonçalves e Guidolin (2006), assume-se que a melhor previsão para o dia subsequente é o preço de hoje ($S_{t+1} = S_t$ e $r_{t+1} = r_t$), sendo S e r o preço *spot* do ativo subjacente e a taxa pré interpolada para a maturidade de cada opção, respectivamente.

Os preços das *calls* e *puts* são derivados da função de BS, utilizando o vetor projetado de IVs ($\hat{\sigma}_{t+1}$), e considerando a maturidade como fração do ano base 360 dias a partir dos vencimentos das opções/séries de t . De forma similar, calcula-se a grega Delta das *calls* e *puts*.

Optamos pela imposição desta proposição *martingale* a fim de focar primordialmente na análise do poder preditivo dos 3 modelos propostos, pois a inclusão de outro modelo para o ativo objeto, como propuseram Gonçalves e Guidolin (2006) ao implementar o modelo EGARCH(1,1), poderia dificultar a distinção entre os fatores principais que conduziram os resultados do período fora da amostra.

Ainda que possa haver algum impacto na qualidade da projeção da IV, que é uma função do *moneyness*, ao assumir variação nula do S , entendemos que o horizonte de projeção de 1 dia tende a mitigar um potencial efeito de distorção. Além disso, dado que aplicamos esta proposição para os 3 modelos, conseguimos atingir um dos objetivos deste trabalho, sob a ótica de comparação de performance preditiva dos coeficientes dos modelos. Vale ressaltar que o delta projetado, utilizando a IV estimada, pode divergir para cada modelo. Logo, deve-se considerar este impacto nas estratégias com portfólios delta-neutro, que incluem a variação de preço do ativo objeto no resultado.

Por exemplo, o portfólio com *straddle* ATM, com delta 50, reflete exclusivamente o poder de predição dos modelos, visto que a somatória dos deltas da *call* e *put* é nula e, portanto, não há impacto da oscilação nos preços do ativo objeto no resultado do *trading*. Em contrapartida, o portfólio posicionado na inclinação da curva, denominado *risk reversal*, embora seja delta-neutro, pode sofrer diferenças significativas em seu valor de mercado em função da variação realizada de preço do ativo subjacente.

Como consequência da irregularidade do *grid* diário das séries/*strikes*, ou seja, o caso em que um determinado *strike* em t não é negociado em $t + 1$, refletindo a escassez do número de séries negociadas no mercado local, utilizamos os preços das opções para

deltas específicos a fim de normalizar a comparação das medidas estatísticas entre o *skew* projetado e observado em $t + 1$. Para tanto, adotamos o método *spline* cúbico de modo a interpolar as IVs para os deltas 25, 50 e 75%. Para mais detalhes acerca deste método de interpolação, ver APÊNDICE B.

Os métodos para validação da técnica do Filtro de Kalman linear para modelagem da dinâmica do *skew* de volatilidade nos períodos fora da amostra no que tange à significância estatística e potencial geração de valor econômico são descritos nas seções a seguir:

4.5.3 Indicadores Estatísticos de Desempenho

Computamos as medidas estatísticas de acurácia aplicadas sobre os dados projetados e observados dos preços das *calls* e *puts*, bem como suas respectivas IVs $[\hat{\sigma}_{ATM,t+1} - \sigma_{ATM,t+1}]$, com os deltas 25, 50 e 75, ou OTM, ATM e ITM, para cada modelo:

1. Raiz do Erro Quadrático Médio (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{\sigma}_{ATM,t+1} - \sigma_{ATM,t+1})^2}{N}}, \quad (4.17)$$

sendo N o número de observações.

2. Erro Médio Absoluto (MAE)

$$MAE = \frac{\sum_{i=1}^N |\hat{\sigma}_{ATM,t+1} - \sigma_{ATM,t+1}|}{N}, \quad (4.18)$$

3. Média da Acurácia Direcional (MDA)

$$MDA = \frac{\sum_{i=1}^N \mathbf{1}_{\text{SINAL}(\sigma_{ATM,t+1} - \sigma_{ATM,t}) = \text{SINAL}((\hat{\sigma}_{ATM,t+1} - \sigma_{ATM,t}))}}{N}, \quad (4.19)$$

sendo a função indicador igual a **1** quando há igualdade de sinais na equação. Portanto, o MDA é a frequência média em que a variação da IV projetada tem o mesmo sinal do que a IV realizada.

4.5.4 Estratégias de *Trading*

Apuramos o desempenho e aderência aos dados empíricos dos modelos por meio da projeção da variação diária do valor de mercado de portfólios de opções com estratégias que refletem a sensibilidade às variações de diferentes naturezas na curva de volatilidade.

Os portfólios são delta-neutro ao assumir posição comprada ou vendida no ativo subjacente, cuja quantidade é o $\Delta_{(t+1)}$ derivado da fórmula de BS utilizando a IV projetada pelos modelos. Neste caso, considerando o theta de apenas 1 dia, cujo impacto no valor do portfólio é limitado, observa-se que os riscos inerentes aos portfólios são o

vega e gamma, sendo o primeiro objeto de estudo deste trabalho. Como já mencionado, as opções OTM/ITM são baseadas nos deltas -0,25/-0,75 e 0,25/0,75 para as *puts* e *calls*, respectivamente.

Para cada período t , montamos os portfólios delta-neutro, cujo $\Delta_{(t)}$ utiliza a IV observada, e calculamos o valor de mercado para t e $t + 1$ com base nos dados observados, de modo a computar a variação diária do valor de mercado na amostra.

Em paralelo, a partir da IV projetada para $t + 1$ $[\hat{\sigma}_{t+1}]$, e utilizando novamente a fórmula de BS, tem-se a variação diária projetada do valor de mercado dos portfólios. Neste caso, utilizamos o delta projetado $\Delta_{(t+1)}$.

A seguir, temos as posições e os respectivos fluxos de caixa utilizados na elaboração das estratégias de *trading*:

Tabela 2 – Fluxos de caixa de posições *delta hedged* em *Call/Put* OTM.

Posição/Sinal	$Opcao_{Delta}$	$Delta\ Hedge$	Caixa
Compra (-)	$Call_{25}$	$\Delta_{Call_{25}}Spot$	$-(Call_{25} - \Delta_{Call_{25}}Spot)$
Venda (+)	$Call_{25}$	$\Delta_{Call_{25}}Spot$	$+(Call_{25} - \Delta_{Call_{25}}Spot)$
Compra (-)	Put_{25}	$\Delta_{Put_{25}}Spot$	$-(Put_{25} + \Delta_{Call_{25}} Spot)$
Venda (+)	Put_{25}	$\Delta_{Put_{25}}Spot$	$+(Put_{25} + \Delta_{Put_{25}} Spot)$

1. Venda de *straddle*: posição vendida em uma *call* ATM e uma *put* ATM, refletindo a sensibilidade do portfólio às mudanças paralelas no nível da volatilidade implícita ATM. O portfólio perde (ganha) valor quando o nível de volatilidade sobe (cai).

Seja V_t o valor de mercado do portfólio em t :

$$V_{Straddle_t} = (Call_{50D,t} - \Delta_{Call_{50D,t}}S_t) + (Put_{50D,t} + |\Delta_{Put_{50D,t}}|S_t).$$

Sabendo que $|\Delta_{Put_{50D,t}}| - \Delta_{Call_{50D,t}} = 0$, obtém-se a variação de V_t para $t + 1$:

$$\begin{aligned} dV_{Straddle_Sample} &= (Call_{50D,t} - Call_{50D,t+1}) + (Put_{50D,t} - Put_{50D,t+1}), \\ dV_{Straddle_Forecast} &= (Call_{50D,t} - \hat{Call}_{50D,t+1}) + (Put_{50D,t} - \hat{Put}_{50D,t+1}), \end{aligned}$$

onde $\hat{Call}_{50D,t+1}$ e $\hat{Put}_{50D,t+1}$ são os valores projetados para 1 dia da *call* e *put*.

Repetimos o processo para cada período t e, de forma recursiva, projetamos a trajetória futura de variação diária do valor de mercado dos portfólios. O objetivo é analisar os primeiros momentos da distribuição desta série temporal de variações diárias e comparar com a amostra.

2. Compra de *risk reversal*: posição vendida em uma *put* OTM e comprada em uma *call* OTM, refletindo a sensibilidade do portfólio às variações na inclinação da curva

de volatilidade. Se houver aumento da inclinação, configurando um *skew* positivo ($\sigma_{\text{CALL}_{25D}} > \sigma_{\text{PUT}_{25D}}$), há ganho para o portfólio. De forma análoga, o *skew* negativo resulta em perda para a posição.

$$\begin{aligned} V_{RR_t} &= (Put_{25D,t} + |\Delta_{Put_{25D,t}}|S_t) - (Call_{25D,t} - \Delta_{Call_{25D,t}}S_t), \\ dV_{RR_Sample} &= (Put_{25D,t} - Put_{25D,t+1}) - (Call_{25D,t} - Call_{25D,t+1}) \\ &\quad - (\Delta_{Call_{25D,t}} + |\Delta_{Put_{25D,t}}|)(S_{t+1} - S_t), \\ dV_{RR_Forecast} &= (Put_{25D,t} - \hat{Put}_{25D,t+1}) - (Call_{25D,t} - \hat{Call}_{25D,t+1}) \\ &\quad - (\hat{\Delta}_{Call_{25D,t+1}} + |\hat{\Delta}_{Put_{25D,t+1}}|)(S_{t+1} - S_t), \end{aligned}$$

sendo $\hat{\Delta}_{Call_{25D,t+1}}$ e $\hat{\Delta}_{Put_{25D,t+1}}$ os deltas projetados para 1 dia da *call* e *put*, respectivamente, considerando as IVs projetadas.

3. Venda de *butterfly spread*: posição vendida em uma *call* OTM e uma *put* OTM, e comprada em uma *call* ATM e uma *put* ATM, cuja sensibilidade é em relação à curvatura (convexidade) da curva. Se houver redução da convexidade, ou seja, elevação do nível das IVs próximas à IV_{ATM} e decréscimo das IVs de *strikes* mais extremos, o resultado do portfólio é positivo e vice-versa.

$$\begin{aligned} V_{BF_t} &= (Put_{25D,t} + |\Delta_{Put_{25D,t}}|S_t) + (Call_{25D,t} - \Delta_{Call_{25D,t}}S_t) \\ &\quad - (Call_{50D,t} + Put_{50D,t}), \\ dV_{BF_Sample} &= -[(Call_{50D,t} - Call_{50D,t+1}) + (Put_{50D,t} - Put_{50D,t+1})] \\ &\quad + [(Call_{25D,t} - Call_{25D,t+1}) + (Put_{25D,t} - Put_{25D,t+1})] \\ &\quad + [(|\Delta_{Put_{25D,t}}| - \Delta_{Call_{25D,t}})(S_{t+1} - S_t)], \\ dV_{BF_Forecast} &= -[(Call_{50D,t} - \hat{Call}_{50D,t+1}) + (Put_{50D,t} - \hat{Put}_{50D,t+1})] \\ &\quad + [(Call_{25D,t} - \hat{Call}_{25D,t+1}) + (Put_{25D,t} - \hat{Put}_{25D,t+1})] \\ &\quad + [(|\hat{\Delta}_{Put_{25D,t+1}}| - \hat{\Delta}_{Call_{25D,t+1}})(S_{t+1} - S_t)]. \end{aligned}$$

4.5.5 Valor Econômico das Estratégias de *Trading*

Adicionalmente, incluímos a análise da evolução do valor de mercado de uma estratégia de *trading* a partir de sinais de compra e venda emitidos por cada modelo implementado anteriormente.

De forma geral, se $\hat{\sigma}_{ATM,t+1} > \sigma_{ATM,t}$ ($\hat{\sigma}_{ATM,t+1} < \sigma_{ATM,t}$), monta-se uma posição comprada (vendida) em volatilidade, ou seja, compra (vende) uma unidade de *call* e *put* de mesmo *strike* (*straddle*), considerando os deltas especificados acima, com os preços próximos ao fechamento de mercado em t , com liquidação da posição também perto do

fechamento⁵ em $t + 1$. Por exemplo, considerando um portfólio com injeção de recursos ou posição comprada nas opções:

$$V_t = -(Put_{ATM,t} + Call_{ATM,t}).$$

Para cada período t , há um investimento fixo de R\$10.000,00 e, portanto, uma quantidade $Q_t = 10.000/V_t$. O resultado do portfólio em $t + 1$ é obtido como:

$$Res_{tOUT} = Q_t[(Call_{ATM,t} - Call_{ATM,t+1}) + (Put_{ATM,t} - Put_{ATM,t+1})].$$

Se houver geração de fluxo de caixa com a venda das opções (venda de *straddle*), com $V_t > 0$, tem-se $Q_t = 10.000/|V_t|$, assumindo que este valor gerado e o valor diário disponível são investidos no CDI *over*. Logo, o resultado do portfólio será:

$$Res_{tIN} = Res_{tOUT} + 20.000 \exp \frac{r_t}{252},$$

em que r_t é a taxa pré interpolada em base anual.

Se $Res_{tOUT} < 0$, adotamos a premissa de que o portfólio é auto financiável, ou seja, há o incremento do carregio do CDI *over* no resultado negativo, sendo $Res_{tOUT} \exp \frac{r_t}{252}$.

Visto que o intuito desta etapa é analisar a qualidade do poder preditivo do modelo a partir da geração de valor econômico das estratégias de *trading*, optamos por implementar apenas a estratégia *straddle*, cujo delta *hedge* é igual a zero, de modo a eliminar o possível impacto da variação do preço do ativo objeto.

Para tanto, definimos algumas regras de negociação como:

- A execução do *trading* ocorre apenas se houver uma diferença mínima de 5 *basis points* entre a volatilidade projetada em $t + 1$ e a realizada em t , de modo a mitigar o viés do fator de risco *theta*, que, por definição, reduz o valor da opção em função do tempo.
- Não há negociação no dia em que há rolagem do vencimento da série, pois se assume que a liquidação da posição em $t + 1$ deve seguir os mesmos parâmetros das opções negociadas em t . Neste caso, $Res_{tOUT} = 10.000 \exp \frac{r_t}{252}$.

Para efeito de comparação de performance dos modelos, utilizamos o índice de *Sharpe*:

$$IS = \frac{(\mu_{ResOUT} - r_f)}{\sigma_{ResOUT}}, \quad (4.20)$$

em que μ_{ResOUT} e σ_{ResOUT} é a média e o desvio padrão dos resultados diários, respectivamente; e r_f a média da taxa pré interpolada para cada período.

⁵ Assumimos que cada *trade* é executado perto do fechamento de mercado, pois os dados utilizados para atualizar os 3 modelos são preços de fechamento.

No capítulo seguinte serão apresentados os resultados da análise comparativa de performance dos modelos abordados, aplicados sobre os dados da ação PETR4 e do índice IBOVESPA, bem como a descrição detalhada dos possíveis diagnósticos.

5 Resultados

Este capítulo descreve os detalhes da aplicação de ferramentas estatísticas a fim de avaliar a performance dos modelos nos períodos de estimação e fora da amostra. Além disso, há detalhamento dos resultados das estratégias de *trading*, abordando os possíveis diagnósticos e limitações.

5.1 Validação de performance dos modelos (*in sample*)

A Tabela 3 mostra os parâmetros estimados dos modelos KF e KF_IBOV. Denota-se, como exemplo, a janela 2015-1S16 como o período de estimação de Janeiro_15 a Dezembro_15 e o período fora da amostra sendo o primeiro semestre de 2016 (1S16). Portanto, os períodos de estimação são: Janeiro_15 a Dezembro_15, Junho_15 a Junho_16, Janeiro_16 a Dezembro_16, Janeiro_17 a Dezembro_17 e Junho_17, Junho_18.

Tabela 3 – Parâmetros estimados dos modelos KF e KF_IBOV

	KF				KF_IBOV			
2015-1S16	Diag_U	a	Diag_A	μ	Diag_U	a	Diag_A	μ
X(1)-Nível	0,861	0,077	0,139	0,551	0,515	0,013	0,485	0,027
X(2)-Inclinacao	0,564	-0,014	0,436	-0,032	0,151	0,076	0,849	0,090
X(3)-Curvatura	0,450	0,060	0,550	0,109	-0,335	0,000	1,335	0,000
X(4)-Assimetria	0,273	-0,004	0,727	-0,005	-0,207	-0,130	1,207	-0,108
Jun_15_16 - 2S16	Diag_U	a	Diag_A	μ	Diag_U	a	Diag_A	μ
X(1)-Nível	0,973	0,018	0,027	0,663	0,563	0,000	0,437	0,000
X(2)-Inclinacao	0,308	-0,013	0,692	-0,019	-0,270	0,063	1,270	0,050
X(3)-Curvatura	0,286	0,068	0,714	0,095	-0,523	0,000	1,523	0,000
X(4)-Assimetria	0,094	-0,007	0,906	-0,007	-0,340	-0,136	1,340	-0,101
2016 - 1S17	Diag_U	a	Diag_A	μ	Diag_U	a	Diag_A	μ
X(1)-Nível	0,953	0,028	0,047	0,600	0,577	0,000	0,423	0,000
X(2)-Inclinacao	0,264	0,002	0,736	0,002	-0,141	0,090	1,141	0,079
X(3)-Curvatura	0,266	0,079	0,734	0,108	-0,643	0,000	1,643	0,000
X(4)-Assimetria	0,082	-0,019	0,918	-0,020	-0,108	-0,086	1,108	-0,078
2017 - 1S18	Diag_U	a	Diag_A	μ	Diag_U	a	Diag_A	μ
X(1)-Nível	0,702	0,103	0,298	0,346	0,382	0,012	0,618	0,019
X(2)-Inclinacao	0,404	-0,011	0,596	-0,018	0,129	0,091	0,871	0,104
X(3)-Curvatura	0,428	0,109	0,572	0,191	-0,002	0,000	1,002	0,000
X(4)-Assimetria	0,348	-0,018	0,652	-0,028	0,054	-0,132	0,946	-0,140
Jun_17_18 - 2S18	Diag_U	a	Diag_A	μ	Diag_U	a	Diag_A	μ
X(1)-Nível	0,841	0,056	0,159	0,355	0,462	0,000	0,538	0,000
X(2)-Inclinacao	0,585	-0,009	0,415	-0,023	0,039	0,078	0,961	0,082
X(3)-Curvatura	0,529	0,089	0,471	0,189	-0,173	0,000	1,173	0,000
X(4)-Assimetria	0,539	-0,013	0,461	-0,027	0,340	-0,175	0,660	-0,266

Para o modelo KF, o coeficiente X_1 (nível ATM) apresenta menor taxa de reversão à média de longo prazo em períodos de maior volatilidade, por exemplo, nas janelas de

2015 e 2016, conforme a Figura 1, indicando a IV ATM como *proxy* da volatilidade do ativo subjacente. De forma análoga, os coeficientes X_2 (inclinação), X_3 (curvatura) e X_4 (assimetria) apresentam valores significativamente maiores nestes períodos mais voláteis.

Nota-se que a média de longo prazo de X_1 é diretamente proporcional à volatilidade do ativo. Vale ressaltar que a média da curvatura é maior em janelas de 2017 e Junho_17, Junho_18.

Em relação ao modelo KF_IBOV, a despeito de as taxas de reversão serem maiores em relação às taxas do KF, mantém-se o racional descrito acima, sendo X_1 menor e os outros fatores maiores em períodos voláteis.

Tabela 4 – Matrizes de correlação dos modelos KF e KF_IBOV.

KF					KF_IBOV			
2015-1516	X_(1)	X_(2)	X_(3)	X_(4)	X_(1)	X_(2)	X_(3)	X_(4)
X(1)-Nível	1,000				1,000			
X(2)-Inclinacao	-0,407	1,000			-0,389	1,000		
X(3)-Curvatura	-0,205	-0,323	1,000		-0,242	-0,224	1,000	
X(4)-Assimetria	0,226	-0,153	-0,839	1,000	0,300	-0,359	-0,694	1,000
Desvio_Padrão	0,052	0,064	0,062	0,041	0,059	0,064	0,064	0,045

Jun_15_16 - 2516	X_(1)	X_(2)	X_(3)	X_(4)	X_(1)	X_(2)	X_(3)	X_(4)
X(1)-Nível	1,000				1,000			
X(2)-Inclinacao	0,077	1,000			0,098	1,000		
X(3)-Curvatura	-0,243	-0,625	1,000		-0,256	-0,588	1,000	
X(4)-Assimetria	0,143	-0,438	-0,407	1,000	0,168	-0,445	-0,351	1,000
Desvio_Padrão	0,056	0,067	0,054	0,039	0,066	0,069	0,059	0,042

2016-1517	X_(1)	X_(2)	X_(3)	X_(4)	X_(1)	X_(2)	X_(3)	X_(4)
X(1)-Nível	1,000				1,000			
X(2)-Inclinacao	-0,099	1,000			-0,078	1,000		
X(3)-Curvatura	-0,296	-0,583	1,000		-0,253	-0,533	1,000	
X(4)-Assimetria	0,308	-0,542	-0,341	1,000	0,194	-0,541	-0,095	1,000
Desvio_Padrão	0,052	0,067	0,051	0,041	0,062	0,071	0,059	0,074

2017-1518	X_(1)	X_(2)	X_(3)	X_(4)	X_(1)	X_(2)	X_(3)	X_(4)
X(1)-Nível	1,000				1,000			
X(2)-Inclinacao	-0,442	1,000			-0,398	1,000		
X(3)-Curvatura	-0,227	-0,407	1,000		-0,213	-0,446	1,000	
X(4)-Assimetria	0,276	-0,622	-0,347	1,000	0,240	-0,618	-0,259	1,000
Desvio_Padrão	0,037	0,072	0,050	0,053	0,040	0,071	0,052	0,061

Jun_17_18 - 2518	X_(1)	X_(2)	X_(3)	X_(4)	X_(1)	X_(2)	X_(3)	X_(4)
X(1)-Nível	1,000				1,000			
X(2)-Inclinacao	-0,609	1,000			-0,606	1,000		
X(3)-Curvatura	-0,012	-0,656	1,000		-0,072	-0,515	1,000	
X(4)-Assimetria	0,403	-0,716	0,155	1,000	0,302	-0,619	0,067	1,000
Desvio_Padrão	0,055	0,101	0,064	0,050	0,066	0,096	0,116	0,083

A Tabela 4 apresenta as matrizes de correlação dos fatores estimados pelos modelos KF e KF_IBOV. Para ambos os modelos, evidencia-se que a correlação negativa entre o X_1 (nível) e X_2 (inclinação) é maior em períodos menos voláteis, com valores médio e máximo em módulo (janela de Junho_17 a Junho_18) de -0,29 e -0,60 para KF e -0,27 e -0,60 para KF_IBOV.

Nota-se que há considerável correlação negativa entre X_3 (curvatura) e X_2 (inclinação), com valor médio das janelas de -0,45 para KF e -0,40 para KF_IBOV, e X_4 e X_2 , com média de -0,38 para KF e -0,46 para KF_IBOV.

A Tabela 5 reporta a correlação cruzada entre as inovações (ou resíduos $[\hat{X}_{t+1|EstPost}] - \hat{X}_{t|EstPost}$) dos coeficientes estimados pelo modelo KF com os dados de PETR4 e os dados dos *skews* do índice IBOVESPA. Observa-se que não há significativa correlação entre os fatores de PETR4 e IBOV, sendo 0,40 o maior valor médio para $X_{1,IBOV}$ e $X_{1,PETR4}$.

Tabela 5 – Correlação cruzada entre os coeficientes KF_PETR4 e KF_IBOVESPA.

	X(1)_PETR4	X(2)_PETR4	X(3)_PETR4	X(4)_PETR4	X(1)_IBOV	X(2)_IBOV	X(3)_IBOV	X(4)_IBOV
X(1)_PETR4	1,000							
X(2)_PETR4	-0,564	1,000						
X(3)_PETR4	0,273	-0,527	1,000					
X(4)_PETR4	0,387	-0,575	-0,119	1,000				
X(1)_IBOV	0,396	-0,150	0,035	0,087	1,000			
X(2)_IBOV	0,077	0,032	-0,045	-0,001	0,000	1,000		
X(3)_IBOV	0,039	-0,072	0,004	0,120	-0,024	0,108	1,000	
X(4)_IBOV	0,077	-0,148	0,104	0,122	0,172	-0,241	0,492	1,000
Desvio_Padrão	0,047	0,070	0,049	0,041	0,018	0,046	0,130	0,178

A Figura 2 demonstra a série temporal dos coeficientes diários que conduzem os *skews* em cada período t , estimados pela equação (4.7) com os dados das opções de PETR4 e do índice IBOVESPA. Observa-se maior descolamento dos dados entre $X_{3,IBOV}$ versus $X_{3,PETR4}$ e $X_{4,IBOV}$ versus $X_{4,PETR4}$.

Neste caso, o objetivo é avaliar a relação conjunta para cada *skew* (*cross-sectional*). Mais adiante, testaremos o desempenho do modelo KF_IBOV no período fora da amostra de modo a explorar a relação dinâmica entre os coeficientes.

Em relação à parametrização do modelo II, aplicamos o método MQO para cada período t , obtendo os vetores diários estimados ($\hat{\beta}$). A Tabela 6 reporta os parâmetros estimados pela equação (4.12), avaliando a aderência aos dados empíricos nos períodos de estimação. Por simplificação, utilizamos a média dos valores das janelas de estimação, sem prejuízo para a análise, visto que se observou forte linearidade das medidas estatísticas ao longo dos períodos, com estreita oscilação em relação aos valores médios.

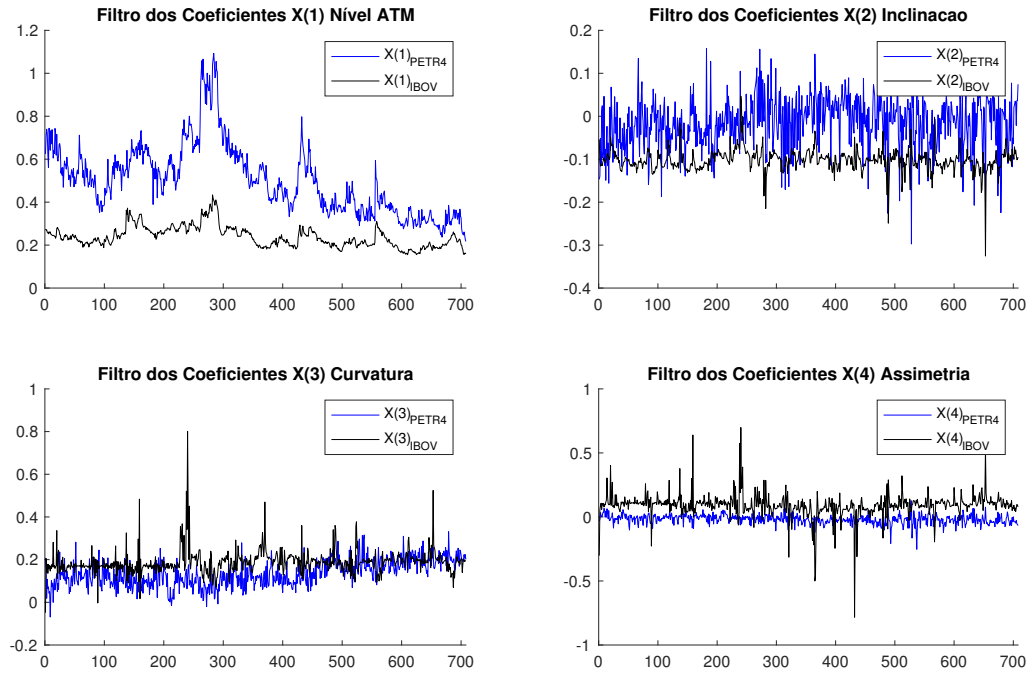
Tabela 6 – Resumo das estatísticas para os parâmetros estimados pelo modelo II.

	Média	DesPad	Min	Max	Skew	Kurtosis	LB_5	LB_10
X(1)-Nível	-0,741	0,216	-1,249	-0,103	0,392	3,642	776,553	1.260,925
X(2)-Inclinacao	-0,013	0,163	-0,478	0,904	0,869	8,701	5,738	11,921
X(3)-Curvatura	0,231	0,223	-1,741	0,979	-3,612	33,989	22,451	36,153
X(4)-Assimetria	-0,002	0,216	-0,851	1,539	1,950	20,492	12,204	19,687
R_quad	0,713	0,208	0,073	0,988	-1,034	3,949	1,000	1,000
R_quad_aju	0,666	0,241	-0,073	0,985	-1,031	3,930	1,000	1,000
RMSE	0,074	0,039	0,015	0,222	0,975	4,057	1,000	1,000

Denota-se R_quad_aju como R^2 ajustado (coeficiente de determinação), utilizado

para comparar a forma funcional dos modelos lineares, variando o número de fatores ou variáveis explicativas; e RMSE a raiz do erro quadrático médio das IVS.

Figura 2 – Evolução dos coeficientes estimados pelo modelo KF_PETR4 e KF_IBOVESPA para o período de Jan_2015 a Dez_2017.



Ainda na Tabela 6, na média, $R_{\text{quad_aju}}$ é igual a 66%, com mínimo de -7% e máximo de 98%. Constatou-se, por meio da série de tempo dos valores diários de $R_{\text{quad_aju}}$ e RMSE (não reportado nesta Tabela), que há considerável variação do poder explicativo da equação (4.12).

Os valores das estatísticas do teste de Ljung-Box para as *lags* 5 e 10 indicam que há significativa autocorrelação temporal para todos os coeficientes, corroborando a tese de que há instabilidade na série de tempo dos fatores estimados, e que a implementação do modelo VAR a fim de explorar as propriedades dinâmicas destes fatores se mostra adequada.

Em relação ao teste, de forma geral, assume-se que a estatística Q tem uma distribuição χ^2 com h graus de liberdade, sendo h o número de defasagens. O teste rejeita a hipótese nula de que não há autocorrelação, ou seja, as variáveis são i.i.d, se $Q > \chi^2_{1-\alpha, h}$, em que α é o nível de significância, aqui adotado como sendo 1%. Portanto, o valor crítico é igual a 0,40 e 2,15 para 5 e 10 *lags*, respectivamente.

5.2 Análise dos Indicadores Estatísticos de Previsão (*out of sample*)

A seguir, implementamos um método recursivo de *back-testing* para os deltas 25, 50 e 75, cujo objetivo é mensurar as propriedades preditivas dos modelos por meio da análise das medidas estatísticas do erro de previsão da IV e preço da opção, considerando a média destas medidas para as 5 janelas fora da amostra, que são: Janeiro_16 a Junho_16, Julho_16 a Dezembro_16, Janeiro_17 a Junho_17, Janeiro_18 a Junho_18 e Julho_18 a Dezembro_18. A escolha das janelas reflete a necessidade de avaliar o desempenho dos modelos em diferentes cenários, combinando períodos com diferentes níveis de volatilidade do ativo subjacente, conforme Figura 1.

Vale salientar que a utilização dos deltas acima, além de permitir a análise econômica dos modelos por meio das estratégias de *trading*, descritas adiante, permite a análise de desempenho dos modelos condicionada às restrições de disponibilidade dos dados observados utilizados como *input* na calibragem dos modelos. Por definição, conforme a seção 4.1, opções OTM apresentam menor liquidez à medida que se aproximam do vencimento, como reflexo, dentre outros fatores, do seu *gamma* reduzido, que é taxa de variação do delta em relação à variação do preço do ativo objeto. De forma intuitiva, o *gamma* reduzido indica que há menor probabilidade de exercício da opção, ou seja, tornar-se ITM e, por isso, há impacto no número de negócios.

A Tabela 7 reporta os indicadores calculados com base nas volatilidades implícitas, e respectivos preços das opções calculados pela fórmula de BS. Na tabela, RMSE_IV é a raiz do erro quadrático médio aplicado sobre as IVs projetadas e observadas, ao passo que RMSE_Px utiliza os preços das opções; MAE é o erro médio absoluto e MDA é a média da acurácia direcional.

Tabela 7 – Valores médios das medidas estatísticas do erro de previsão.

	RMSE_IV			RMSE_Px			MAE_IV			MAE_Px			MDA		
	25D	ATM	75D	25D	ATM	75D	25D	ATM	75D	25D	ATM	75D	25D	ATM	75D
KF	0,055	0,061	0,100	0,034	0,069	0,227	0,037	0,042	0,064	0,027	0,051	0,152	0,488	0,543	0,599
Mdl II	0,158	0,167	0,179	0,169	0,240	0,274	0,126	0,132	0,138	0,137	0,188	0,204	0,512	0,512	0,525
KF_IBOV	0,069	0,074	0,103	0,064	0,093	0,229	0,051	0,054	0,072	0,051	0,074	0,161	0,497	0,509	0,562

Definimos, a seguir, os principais pontos evidenciados na Tabela 7:

- Evidencia-se a possibilidade de predição da dinâmica do *skew* de volatilidade das opções de PETR4, fundamentando-se nos valores médios acima de 0,50 da medida MDA. A leitura predominante é de que o modelo KF apresenta os maiores valores de MDA para as opções ATM e ITM, mostrando um valor marginalmente menor para as opções OTM em relação aos valores dos modelos II e KF_IBOV.

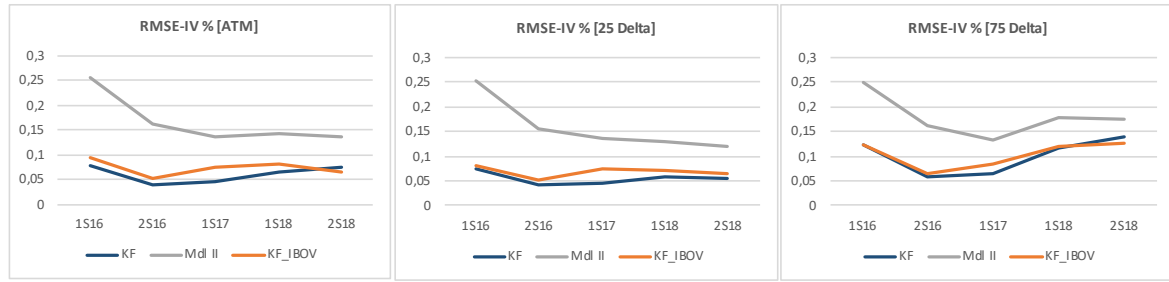
- Os movimentos dos *skews* de volatilidade das opções do índice IBOVESPA não provém potencial conteúdo informacional ao modelo KF de PETR4 de modo a aprimorar a projeção dos coeficientes e, por conseguinte, elevar a robustez do modelo, denominado, neste trabalho, como KF_IBOV. Nota-se que os erros de previsão dos modelos KF e KF_IBOV são similares para as opções ITM, como será demonstrado no teste de acurácia adotado. Embora haja uma correlação cruzada entre os coeficientes que direcionam o formato dos *skews* diários de PETR4 e do índice IBOVESPA, conforme Tabela 5, apontando que parte da variação da IVS da opções de ação pode ser explicada por movimentos passados da IVS do Índice, concluímos que a relação dinâmica não aumenta o poder de predição do modelo KF_IBOV. Pode-se elencar como possível diagnóstico a liquidez reduzida das opções do Índice vis-à-vis as opções de PETR4.

De forma intuitiva, observa-se empiricamente que o *skew* das opções de ações individuais apresenta diferentes *strikes* e, com isso, é razoável assumir que as oscilações no *skew* não são concomitantes às execuções dos negócios das opções com determinados deltas. Logo, se as opções do índice apresentassem maior liquidez, com maior número de negócios em diferentes deltas e, assumindo que há correlação entre os preços dos ativos objetos (ou sensibilidade do retorno da ação em relação ao retorno do índice, definido como beta), caberia investigar se o conteúdo informacional dos *skews* passados do índice, adicionado recursivamente ao processo de filtragem do algoritmo KF, poderia impactar a projeção dos coeficientes da IVS da ação, o que violaria a hipótese eficiência de mercado.

- Em suma, o modelo KF, que inclui apenas informações dos movimentos passados da IVS das opções de PETR4, apresentou uma performance preditiva de 1 dia superior aos modelos II e KF_IBOV, como evidenciado pelos indicadores de RMSE_IV, MAE_IV e MDA.

De forma adicional, pode-se mapear na Figura 3 um padrão de inclinação da curva do erro de previsão (RMSE-IV) para o modelo II, sendo maior nas janelas 1S16 e 2S16, o que corrobora a inabilidade deste modelo em períodos mais voláteis, de acordo com a Figura 1. Em contrapartida, os modelos KF e KF_IBOV mostram maior linearidade do erro ao longo dos períodos para as opções ATM e OTM, e leve inclinação da curva do erro nas janelas 1S18 e 2S18 para as opções ITM (75 delta), o que reflete a consistência do método recursivo de filtragem de KF, desvinculando a qualidade de predição do modelo do nível de volatilidade do ativo objeto.

Figura 3 – Evolução do RMSE-IV para o 3 modelos nos períodos fora da amostra.



A Tabela 8 reporta o teste de acurácia de predição, cuja metodologia foi desenvolvida por Diebold (1995). O objetivo é testar a performance dos modelos KF_IBOV e II versus o modelo KF. O teste é computado utilizando uma função perda $f(e_{it})$, sendo i o modelo testado. Definimos a função perda como o erro quadrático de previsão, logo: $f(e_{it}) = e_{it}^2$.

Com isso, compara-se os erros de projeção dos modelos $d_t = e_{1t}^2 - e_{2t}^2$, e define-se a hipótese nula $E(d_t) = 0 \forall t$ versus hipótese alternativa $E(d_t) \neq 0$.

Seja DM a estatística teste:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\sigma^2}{T}}}, \quad (5.1)$$

em que \bar{d} e σ^2 é a média e variância da série histórica dos períodos fora da amostra consolidados, e T o número de observações da série. Por definição, a estatística teste DM é assintoticamente distribuída como uma normal padrão (ver APÊNDICE C para maiores detalhes). Portanto, o valor crítico da estatística teste é igual ao valor crítico da normal padrão (z); logo, rejeita-se H_0 se $|DM| > z_{\alpha/2}$.

Tabela 8 – Teste de acurácia de previsão (contra o Modelo KF).

	Mdl II				KF_IBOV			
	DM_IV		DM_Px		DM_IV		DM_Px	
ATM	-17,00	*	-15,95	*	-7,11	*	-10,37	*
25D	-17,77	*	-19,46	*	-10,71	*	-11,88	*
75D	-10,57	*	-5,18	*	-0,91		-0,50	

O (*) indica que os valores são estatisticamente significantes com um nível de 5%, com valor crítico igual a $z_{\alpha/2} = 1.96$. Deste modo, verifica-se que a hipótese nula de igualdade da acurácia das previsões entre os modelo KF e II é rejeitada para os diferentes deltas, indicando que KF apresenta maior robustez. Em relação ao KF_IBOV, pode-se afirmar que há igualdade de acurácia de predição apenas para as opções ITM.

5.3 Resultados das estratégias de *Trading*

Conforme exposto anteriormente, a análise das medidas de acurácia, reportada de forma detalhada na Tabela 9, cujos valores superam, majoritariamente, a média de

50%, sugere que os *skews* diários das opções de PETR4, e os preços correspondentes, são estatisticamente previsíveis como resultado da implementação do método do Filtro de Kalman linear. A seguir, analisaremos a significância econômica do poder preditivo do modelos propostos, cuja finalidade engloba gestão de risco e alavancagem de resultados financeiros por meio de estratégias de *trading* com opções.

Tabela 9 – Evolução dos valores da medida MDA para os 3 modelos.

	KF - MDA %			Mdl II - MDA %			KF_IBOV - MDA %		
	ATM	25	75	ATM	25	75	ATM	25	75
1516	0,612	0,545	0,645	0,533	0,558	0,517	0,563	0,600	0,639
2516	0,554	0,471	0,620	0,484	0,517	0,512	0,512	0,504	0,537
1517	0,561	0,455	0,616	0,484	0,475	0,529	0,421	0,413	0,500
1518	0,529	0,488	0,576	0,590	0,492	0,579	0,537	0,463	0,562
2518	0,458	0,479	0,537	0,467	0,517	0,488	0,512	0,504	0,570

5.3.1 Análise das densidades preditivas

Para fins de gestão de risco, validamos a aderência com que modelo KF projeta a trajetória futura de variação diária do valor de mercado de alguns portfólios, que contemplam as estratégias *straddle*, *risk reversal* e *butterfly*, sensíveis à choques de diferentes naturezas na curva, como nível, inclinação e curvatura, respectivamente.

Seguindo o seção 4.5.4, para cada período t das janelas fora da amostra, montamos as 3 estratégias¹, calculando o valor de mercado em t e $t + 1$ com base nos dados observados, bem como o valor em $t + 1$ com base nos preços das opções projetados por cada modelo. Com isso, tem-se a variação diária observada e projetada do valor de mercado dos 3 portfólios. Em seguida, computamos as estatísticas descritivas dos resultados diários para ambas as séries temporais, como média, desvio-padrão, assimetria e excesso de curtose, conforme Tabela 10 que mostra as estatísticas descritivas para as densidades realizada e projetada dos resultados diários dos portfólios de opções com as 3 estratégias descritas para cada período fora da amostra.

Vale enfatizar que os modelos abordados utilizam os mesmos fatores, que conduzem o formato dos *skews*, oriundos da equação de estado linear e, portanto, teoricamente, configuram-se como igualmente aptos a modelar a evolução dinâmica da curva de volatilidade. Neste contexto, o objetivo deste trabalho é verificar se o método de filtragem de KF, que possui método de estimação sequencial (recursivo), adicionando informações disponíveis de mercado em cada período t , apresenta performance superior em relação ao modelo II, e consistente em diferentes períodos, baseado na análise dos momentos das densidades preditivas produzidas pelo modelo. Além disso, mapeamos o impacto do conteúdo informacional das IVS das opções de IBOVESPA na modelagem das IVS das opções de PETR4.

¹ Considera-se a compra/venda de uma unidade da *call/put*.

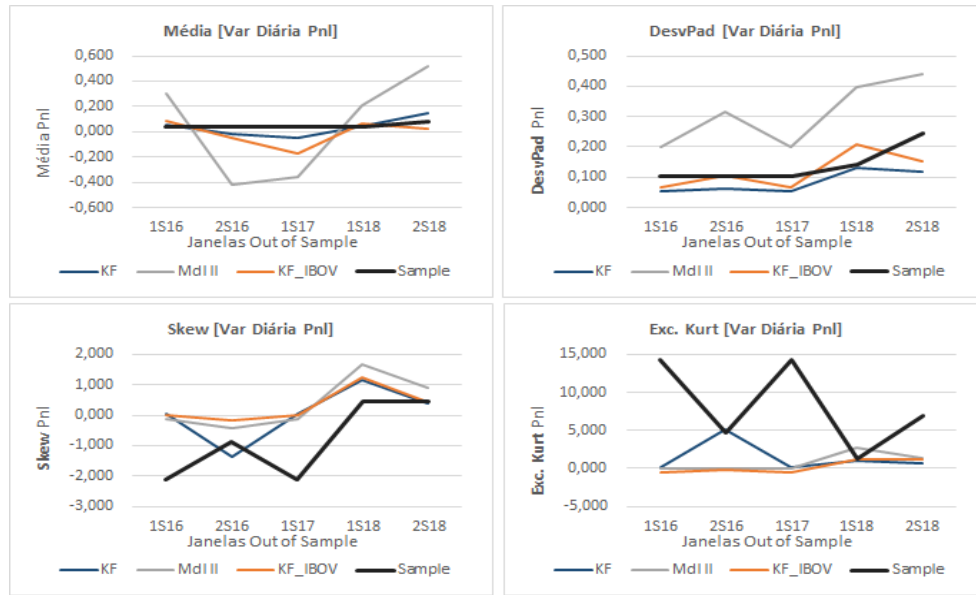
Tabela 10 – Resumo das estatísticas dos resultados diários dos portfólios.

15-1516					Jun_15_16 - 2516				16-1517					
	Média	DesPad	Skew	Exc Kurt		Média	DesPad	Skew	Exc Kurt		Média	DesPad	Skew	Exc Kurt
Short Straddle ATM														
Sample	0,0348	0,1013	-2,1132	14,2354	0,0385	0,1044	-0,8763	4,7183	0,0414	0,1013	-2,1132	14,2354		
KF	0,0516	0,0521	0,0271	0,1418	-0,0152	0,0638	-1,3754	5,1680	-0,0451	0,0521	0,0271	0,1418		
Mdl II	0,3060	0,2009	-0,1444	0,0449	-0,4202	0,3135	-0,4352	-0,0577	-0,3523	0,2009	-0,1444	0,0449		
KF_IBOV	0,0843	0,0684	-0,0139	-0,6120	-0,0507	0,1050	-0,1530	-0,1495	-0,1725	0,0684	-0,0139	-0,6120		
Long Risk Reversal 25D														
Sample	-0,0120	0,1791	-0,1919	0,9078	-0,0222	0,2135	-0,0301	1,1050	0,0121	0,1950	2,1735	11,9882		
KF	-0,0230	0,1773	-0,0446	0,1340	-0,0313	0,2164	0,1113	0,4372	0,0205	0,2107	1,7948	9,4740		
Mdl II	-0,0001	0,1412	-0,0072	0,2304	-0,0337	0,2495	0,2030	-0,0270	0,0353	0,2532	2,4733	14,8733		
KF_IBOV	-0,0152	0,1771	-0,0549	0,1184	-0,0567	0,2234	0,0360	0,0859	0,0197	0,2285	1,9748	11,0153		
Short Butterfly Spread (ATM / 25D)														
Sample	-0,0210	0,1061	0,1082	1,3633	-0,0265	0,0925	0,8868	4,5419	-0,0264	0,0847	1,7624	9,5869		
KF	-0,0215	0,0712	0,0253	0,9252	0,0023	0,0567	0,4216	2,5065	0,0146	0,0552	0,9219	1,9354		
Mdl II	-0,0940	0,1232	-1,0731	1,1734	0,0722	0,0733	0,7798	3,0089	0,0652	0,0588	0,5364	0,3658		
KF_IBOV	-0,0230	0,0668	-0,0467	1,0907	0,0016	0,0648	0,0063	3,4950	0,0385	0,0562	1,0330	2,2136		

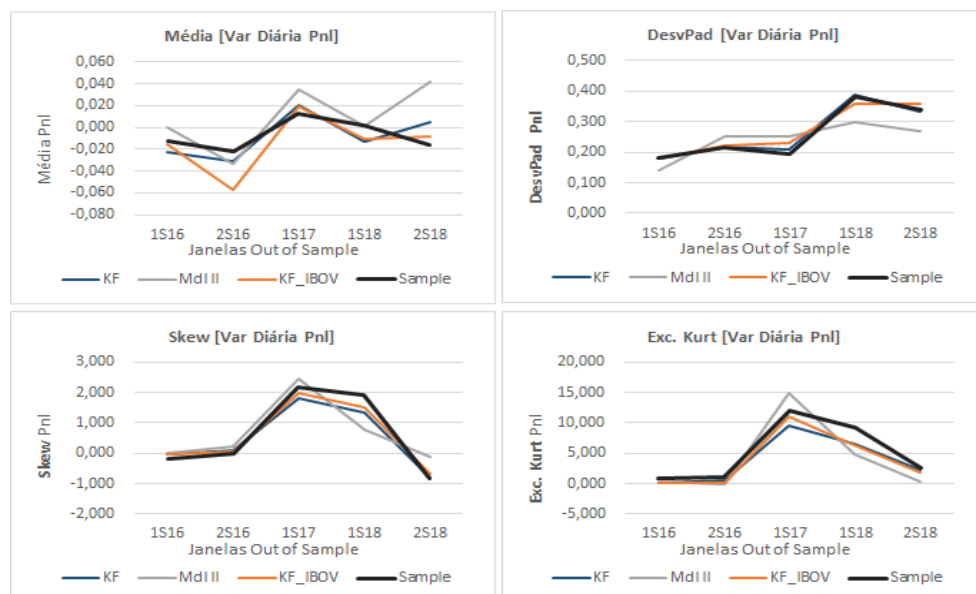
17-1518					Jun_17_18 - 2518				
	Média	DesPad	Skew	Exc Kurt		Média	DesPad	Skew	Exc Kurt
Short Straddle ATM									
Sample	0,0392	0,1424	0,4665	1,2323	0,0761	0,2462	0,4706	6,9200	
KF	0,0419	0,1294	1,1697	0,9818	0,1468	0,1181	0,3734	0,6396	
Mdl II	0,2135	0,3967	1,6763	2,7006	0,5183	0,4418	0,9008	1,3973	
KF_IBOV	0,0679	0,2079	1,2424	1,1851	0,0262	0,1541	0,4177	1,2016	
Long Risk Reversal 25D									
Sample	0,0016	0,3802	1,9252	9,2636	-0,0156	0,3377	-0,8372	2,6522	
KF	-0,0125	0,3888	1,3527	6,4525	0,0049	0,3326	-0,7942	2,3552	
Mdl II	0,0008	0,2982	0,7868	4,7921	0,0419	0,2672	-0,0983	0,3399	
KF_IBOV	-0,0112	0,3586	1,4987	6,4221	-0,0077	0,3597	-0,6811	1,9016	
Short Butterfly Spread (ATM / 25D)									
Sample	-0,0236	0,1245	-0,2233	4,0141	-0,0075	0,3058	1,0189	8,0469	
KF	-0,0124	0,0900	-0,0308	5,3068	-0,0538	0,1222	0,6367	2,2481	
Mdl II	-0,0566	0,1430	-1,8025	4,0016	-0,1470	0,1971	-0,5720	1,6051	
KF_IBOV	-0,0190	0,0920	-0,2135	3,5424	-0,0712	0,1286	0,5571	1,8486	

A seguir, mapeamos os resultados diários (PnL) das densidades para cada portfólio de opções:

- Venda de *straddle* ATM: Em geral, de acordo com a Figura 4, observa-se que os primeiros momentos das densidades preditivas obtidas pelo modelo KF apresentam maior aderência aos dados empíricos em comparação aos valores dos modelo II e KF_IBOV, embora o desvio padrão dos resultados diários obtido pelo KF seja marginalmente menor para todas as janelas, sendo os valores do modelo KF_IBOV mais próximos aos dados observados. Os momentos superiores das densidades geradas por KF apresentam menor assimetria e excesso de curtose vis-à-vis os dados empíricos. Neste caso, o descolamento destes momentos do modelo KF_IBOV mostra-se mais estreito para alguns períodos. Cabe ressaltar que as densidades obtidas pelo modelo II mostram significativo viés na média das variações diárias, refletindo a deficiência do modelo em capturar as características da distribuição empírica.

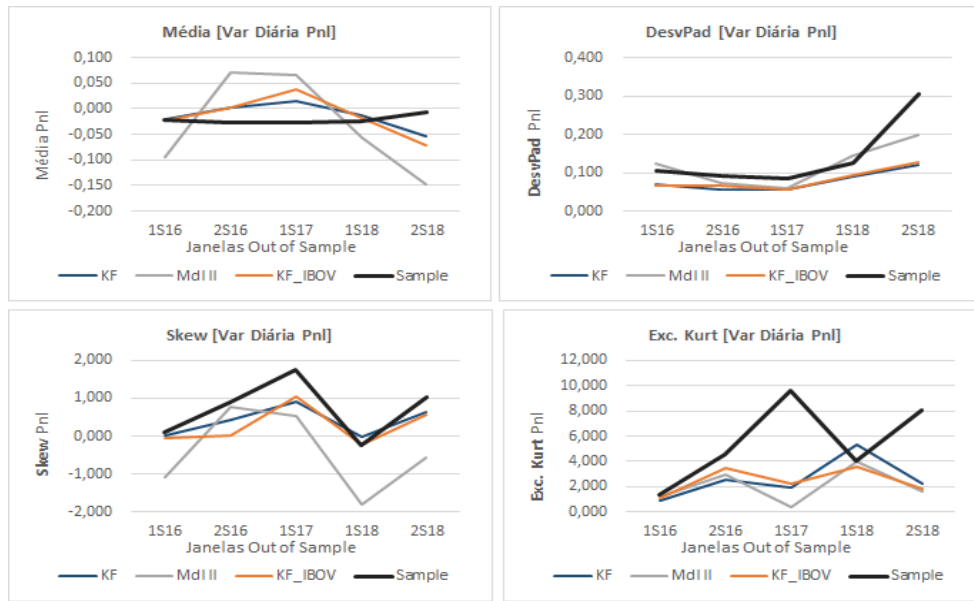
Figura 4 – Momentos das distribuições do PnL diário para *straddle* ATM.

- Compra de *risk reversal* com opções 25 Delta: Observa-se, na Figura 5, que as densidades do modelo KF mostram majoritariamente maior robustez, para os diferentes períodos, principalmente nos primeiros momentos (média e desvio), visto que os valores do excesso de curtose ficaram levemente subestimados. Ainda, verifica-se maior alinhamento da variância e assimetria ao longos dos períodos. Em seguida, destaca-se o modelo KF_IBOV com aderência significativa dos desvios e momentos superiores ao dados observados. O modelo II, embora tenha mostrado menor deslocamento nos primeiros momentos, apresentou discrepância acentuada nas caudas das distribuições vis-à-vis as densidades da amostra.

Figura 5 – Momentos das distribuições do PnL diário para *risk reversal* com 25 Delta.

- Venda de *butterfly spread* com opções 50 e 25 Delta: A leitura predominante da Tabela 10 e Figura 6 é de menor ajuste das distribuições dos 3 modelos em relação aos dados da amostra. O modelo KF mostra um viés de baixa nos valores da variância, o que impacta os valores das caudas, apresentado menor assimetria positiva para algumas janelas, e menor excesso de curtose. O modelo II apresenta valores da média e desvio mais consistentes, com distorções relevantes nas caudas.

Figura 6 – Momentos das distribuições do PnL diário para *butterfly spread* com 50 e 25 Delta.



5.3.2 Análise do valor econômico das estratégias

A estratégia implementada a seguir, denominada estratégia de sinais, visa corroborar o potencial de predição de cada modelo a partir da possibilidade de alavancar o resultado financeiro com *trading* de opções fundamentado nos sinais emitidos pelos modelos.

Para fins de comparação entre os modelos e, considerando a utilização dos preços de fechamento na modelagem, a execução e liquidação da posição vigente do *trade* deve ocorrer, de forma sistemática, perto do fechamento de mercado em t e $t + 1$, respectivamente.

Utilizamos a estratégia *straddle*, que considera o movimento puro de predição da IV, enquanto as outras estratégias, embora sejam delta-neutro, ficam à mercê das variações de preço do ativo subjacente, para o qual assumimos a proposição *martingale*.

Como limitação desta análise, cabe salientar que a estratégia de sinais não captura a magnitude do movimento do preço projetado em $t + 1$ contra o preço observado em t , que refletiria o erro de previsão, sendo substancialmente menor no modelo KF, conforme demonstrado anteriormente. Uma alternativa seria variar a quantidade alocada no portfólio em função da magnitude do movimento esperado da IV projetada pelo modelo. Desta

forma, supondo menor erro de previsão do modelo KF, entende-se que o desempenho seria superior em relação aos modelos II e KF_IBOV.

Tabela 11 – Resumo estatístico das estratégias de sinais com *straddle* [25, 50 e 75 Delta].

15-1516				Jun_15_16 - 2516			16-1517		
	Média	DesPad	IS	Média	DesPad	IS	Média	DesPad	IS
Straddle ATM									
KF	0,0421	0,1056	0,3936	0,0091	0,0718	0,1201	-0,0202	0,0824	-0,2506
Mdl II	0,0310	0,1093	0,2789	-0,0194	0,0696	-0,2869	-0,0251	0,0869	-0,2938
KF_IBOV	0,0228	0,1116	0,1995	-0,0039	0,0731	-0,0608	-0,0346	0,0839	-0,4171
Buy & Hold	0,0023	0,0485	0,0356	0,0034	0,0307	0,0948	-0,0016	0,0272	-0,0763
Straddle 25D									
KF	0,0280	0,2139	0,1284	-0,0145	0,0852	-0,1768	-0,0186	0,1144	-0,1661
Mdl II	0,0319	0,2192	0,1431	-0,0193	0,0928	-0,2136	-0,0224	0,1221	-0,1874
KF_IBOV	0,0248	0,2181	0,1114	0,0018	0,0912	0,0137	-0,0245	0,1230	-0,2027
Buy & Hold	0,0023	0,0485	0,0356	0,0034	0,0307	0,0948	-0,0016	0,0272	-0,0763
Straddle 75D									
KF	0,0488	0,0982	0,4919	0,0103	0,0681	0,1431	0,0196	0,1004	0,1911
Mdl II	0,0261	0,1081	0,2369	-0,0176	0,0663	-0,2741	-0,0058	0,1099	-0,0571
KF_IBOV	0,0484	0,1000	0,4789	-0,0008	0,0720	-0,0182	-0,0119	0,1054	-0,1169
Buy & Hold	0,0027	0,0489	0,0356	0,0038	0,0312	0,0948	-0,0020	0,0274	-0,0763

17-1518				Jun_17_18 - 2518		
	Média	DesPad	IS	Média	DesPad	IS
Straddle ATM						
KF	0,0167	0,0830	0,1985	0,0229	0,1158	0,1953
Mdl II	0,0270	0,0797	0,3351	0,0218	0,1162	0,1856
KF_IBOV	0,0186	0,0824	0,2228	0,0083	0,1181	0,0679
Buy & Hold	0,0000	0,0386	-0,0066	0,0022	0,0279	0,0698
Straddle 25D						
KF	0,0117	0,1078	0,1058	0,0115	0,1678	0,0671
Mdl II	0,0219	0,1119	0,1931	0,0151	0,1683	0,0881
KF_IBOV	0,0059	0,1113	0,0510	-0,0109	0,1217	-0,0915
Buy & Hold	0,0000	0,0386	-0,0066	0,0022	0,0279	0,0698
Straddle 75D						
KF	0,0108	0,0949	0,1109	0,0205	0,1333	0,1523
Mdl II	0,0266	0,0917	0,2876	0,0249	0,1325	0,1857
KF_IBOV	0,0127	0,0937	0,1328	0,0086	0,1341	0,0620
Buy & Hold	-0,0003	0,0384	-0,0066	0,0028	0,0279	0,0698

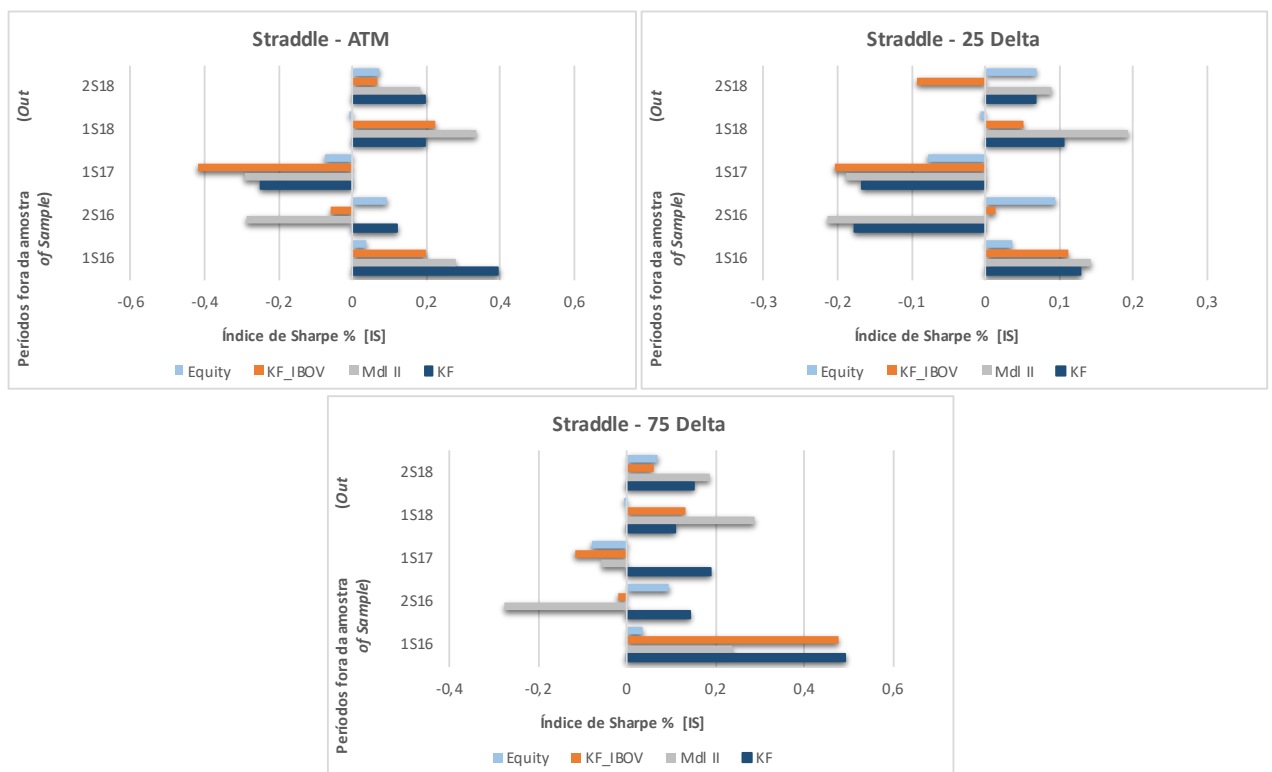
Adicionalmente, por simplificação, utilizamos o índice de *Sharpe* com a finalidade de comparar a performance dos modelos, todavia, deve-se considerar que este indicador utiliza apenas os primeiros momentos (média e desvio) e, portanto, pode limitar a análise comparativa das distribuições de retorno diário.

Como informação complementar, aplicamos as mesmas métricas descritivas sobre os dados do retorno diário dos preços de PETR4, denominando tal estratégia como *buy-and-hold equity*, utilizada como *benchmark*.

A Tabela 11 reporta os resultados das estratégias de sinais com *straddle* [25, 50 e 75 Delta] e *buy-and-hold* de PETR4 para todos os períodos fora da amostra, apresentando a média e desvio padrão dos retornos diários, bem como o índice de *Sharpe* (IS). De forma

geral, os resultados da Tabela 11 estão fortemente ancorados aos fundamentos da medida MDA previamente explicitados, mostrando que o modelo KF obtém maior retorno diário em relação ao modelo II em 80% dos períodos com o *straddle* ATM, 40% com *straddle* 25 Delta e 60% com *straddle* 75 Delta. Analisando o modelo KF de forma isolada, identifica-se que o modelo gerou IS superior ao *benchmark* (*equity*) em 80% dos períodos com ATM, 60% com 25 Delta e 100% com 75 Delta. Considerando a diferença acentuada entre os retornos do KF e *equity*, parece razoável assumir que os custo diários de transação não penalizariam o resultado consolidado. Contra o modelo KF_IBOV, KF apresenta maior retorno em 80% dos períodos com o *straddle* ATM, 25 e 75 Delta.

Figura 7 – Evolução do índice de Sharpe para *straddle* [25, 50 e 75 Delta].



De acordo com a Figura 7, nota-se que a maior discrepância do índice de *Sharpe* do modelo KF em relação ao modelo II, incluindo os 3 Deltas, encontra-se no período 1S18, que apresenta menor patamar de volatilidade do ativo objeto, com menores picos, conforme Figura 1. Em contrapartida, o modelo KF apresentou melhor performance para ATM e 75 Delta, e retorno ligeiramente menor com 25 Delta, para o período mais volátil - 1S16, o que reforça o argumento de que a capacidade preditiva do modelo KF é mais robusta e consistente em períodos mais voláteis, ver Figura 3.

Destaca-se que os resultados das diferentes formas de validação de performance de predição dos modelos se mostraram consistentes. Os erros de previsão do modelo KF foram menores em relação aos modelos II e KF_IBOV nas 5 janelas fora da amostra e

para todos os deltas, mostrando maior linearidade ao longo dos períodos, conforme as curvas da Figura 3.

A despeito da evidência de que houve menor aderência das densidades projetadas dos 3 modelos, principalmente nos momentos superiores, sendo mais acentuada nas estratégias que refletem a inclinação e curvatura, a performance do modelo KF se mostrou superior considerando a média, desvio e cauda das distribuições vis-à-vis a dos outros modelos.

De forma consolidada, os resultados provenientes das estratégias de sinais ratificam os indicadores de acurácia da Tabela 9, mostrando maior potencial para geração de valor econômico do modelo KF para as opções ATM e ITM, principalmente em períodos com maiores picos de volatilidade. Para as opções OTM (25 Delta), a capacidade preditiva do modelo KF é levemente inferior em comparação com o modelo II, refletindo possivelmente a menor liquidez destas opções, o que tende a limitar o aprimoramento da estimação obtido pelo processo de filtragem do *framework* KF.

6 Conclusão

O objetivo deste trabalho foi implementar o algoritmo de Filtro de Kalman linear, que modela a evolução estocástica dos coeficientes que conduzem o formato da curva de volatilidade implícita das opções de PETR4, e comparar sua performance com um modelo de fatores implementado em 2 estágios: regressão linear para os coeficientes das IVS diárias; e VAR(p) para modelar a evolução dos coeficientes, considerando as propriedades das séries de tempo. Ainda, exploramos o impacto dos movimentos passados dos fatores dos *skews* do índice IBOVESPA de modo a aprimorar o poder preditivo do modelo. O critério para escolha das opções de PETR4 baseou-se na elevada liquidez histórica e significativo peso da ação na composição do Índice IBOVESPA.

Identificamos, para o período *out of sample* consolidado, que a performance do modelo KF proposto é majoritariamente superior em comparação aos modelos alternativos II e KF_IBOV, considerando a análise preditiva nas vertentes estatística e econômica, esta última englobando gestão de risco e geração de resultado com estratégias de *trading*, sem adicionar custos de transação. A despeito de haver leve correlação cruzada entre alguns coeficientes da IVS do Índice e de PETR4, evidenciou-se que não houve incremento do conteúdo informacional de modo a elevar o potencial de predição do modelo KF_IBOV.

Importante destacar o embasamento detalhado no trabalho no que tange à existência de um padrão preditivo da curva de volatilidade implícita das opções de PETR4. Neste contexto, além de empregar essas características de predição da IVS na administração de risco e portfólio, pode-se utilizá-las no apreçamento e *hedge* de opções exóticas.

Como limitação para implementação dos modelos, pode-se citar a metodologia empregada pela B3 para divulgação dos dados, indicando o preço de fechamento como o último negócio. Neste contexto, visando limitar o risco de assincronismo entre os preços de fechamento das opções, com impacto relevante na projeção dos coeficientes e, por conseguinte, das IVs, optamos pela utilização dos preços apenas das *calls* em detrimento dos *skews* completos, que considerando as (*calls* e *puts*). Neste caso, é factível assumir que as curvas completas refletiriam um número maior de negócios com diferentes *strikes*, com potencial para aprimorar o processo de filtragem do modelo KF.

Uma possível extensão deste trabalho seria a implementação de modelos alternativos com um número reduzido de fatores, como, por exemplo, o *sticky-delta* e *vega-gamma*, cuja modelagem envolve apenas a volatilidade ATM. Neste caso, o objetivo seria ratificar a performance mais robusta do modelo KF, para projeção da inclinação ou curvatura do *skew*, vis-à-vis a dos modelos alternativos. Além disso, seria interessante replicar o processo para outras maturidades e gerar projeções dos coeficientes, considerando a correlação

dinâmica entre os fatores das diferentes maturidades. De forma complementar, o modelo KF poderia ser implementado em outras opções de ações menos líquidas, de modo a verificar o impacto do fator liquidez na performance do modelo, além de explorar a relação dinâmica dos movimentos passados dos *skews* de IBOVESPA com outras ações.

Como abordagem alternativa, pode-se exercitar a implementação de uma equação de estado não linear, o que requer a utilização do algoritmo *Unscented* KF, contemplando os pontos sigma. O objetivo seria verificar se há acentuado descolamento entre a performance dos modelos linear e não linear no período fora da amostra.

Referências

- 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.
- BEDENDO, M.; HODGES, S. The dynamics of the volatility skew: a kalman filter approach. *J. Bank. Financ*, v. 33, n. 6, p. 1156–1165, 2009.
- BOLLERSLEV, T.; KRONER, K. Arch modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, Elsevier, v. 52, n. 1-2, p. 5–59, 1992.
- CARR, P.; WU, L. Analyzing volatility risk and risk premium in option contracts: A new theory. *Journal of Financial Economics*, v. 120, n. 1, p. 1–20, 2016.
- CHEN, B. C. Empirical performance of alternative option pricing models. *The Journal of Finance*, v. 52, n. 5, p. 2003–2049, 1997.
- CHEN, S. a.; ZHEN, Z. b.; SHENGHONG, L. An efficient estimate and forecast of the implied volatility surface: A nonlinear kalman filter approach. *Economic Modelling*, n. 58, p. 655–664, 2016.
- CONT, R.; FONSECA, J. D. Dynamics of implied volatility surfaces. *INSTITUTE OF PHYSICS PUBLISHING*, v. 2, p. 45–60, 2002.
- DERMAN, E.; KANI, I. The volatility smile and its implied tree. *Goldman Sachs Quantitative Strategies Research Notes*, 1994.
- DIEBOLD, M. Comparing predictive accuracy. *Journal of Business Economic Statistics*, v. 13, n. 3, p. 253–63, 1995.
- DUMAS, B.; FLEMING, J.; WHALEY, R. Implied volatility functions: Empirical. *Journal of Finance*, v. 57, p. 2059–2106, 1998.
- FLEMING, J. The quality of market volatility forecasts implied by sp 100 index option. *Journal of Empirical Finance*, v. 2, p. 17–45, 2002.
- GIBBS, B. P. *Advanced Kalman Filter, Least-Squares and Modeling*. [S.l.: s.n.], 2011.
- GONÇALVES, S.; GUIDOLIN, M. Predictable dynamics in the sp 500 index options implied volatility surface. *Journal of Finance*, v. 79, n. 3, p. 1531–1635, 2006.
- HARVEY, A. *Forecasting, Structural Time Series Models and the Kalman Filter*. [S.l.]: Cambridge University Press, Cambridge., 1989.
- HENTSCHEL, L. Errors in implied volatility estimation. *The Journal of Financial and Quantitative Analysis*, v. 38, n. 4, p. 779–810, 2003.
- HESTON, S.; NANDI, S. A closed-form garch option valuation model. *Review of Financial Studies*, n. 13, p. 585–625, 2000.

- HESTON, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, v. 6, n. 2, p. 327–343, 1993.
- HULL, J. C. *Options, Futures and Other Derivatives*. [S.l.: s.n.], 2005.
- KOU, S. A jump-diffusion model for option pricing. *Management Science*, v. 48, n. 8, p. 1086–1101, 2002.
- MERTON, R. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, v. 3, n. 1-2, p. 125–144, 1976.
- SCHONBUCHER, P. A market model for stochastic implied volatility. *Soc. London*,, v. 357, p. 2071–2092, 1999.
- TOMPKINS, R. G. Implied volatility surfaces: Uncovering regularities for options. *Vienna University of Technology Working Paper*, n. 49, 1999.
- VILLA, C.; FENGLER, M.; HARDLE, W. *The dynamics of implied volatilities: a common principal components approach*. [S.l.], 2003.

Apêndices

APÊNDICE A – Variância assintótica do processo de Ornstein-Uhlenbeck

Seja uma equação diferencial estocástica (SDE), que representa um processo de Ornstein-Uhlenbeck (OU) com distribuição gaussiana:

$$dx_t = \theta(\mu - x_t)dt + \sigma_t dw_t, \quad (\text{A.1})$$

onde dw_t é um processo de Wiener. Tomando $\mu = 0$, podemos reescrever (A.1) como:

$$dx_t = -\theta x_t dt + \sigma_t dw_t. \quad (\text{A.2})$$

Seja uma função conhecida $y = f(x, t) = x e^{\theta t}$, aplica-se o Lema de Itô a fim de solucionar a SDE (para simplificação, denotaremos $f(x, t)$ simplesmente por f):

$$df = f + f' dx_t + \frac{1}{2} f'' dx_t^2.$$

Sabendo que, dadas as características da variação quadrática do movimento browniano, $dx_t^2 = 0$ e $dw_t^2 = dt$, e neste caso, tem-se:

$$dy = [-\theta x_t \frac{df}{dx} + \frac{df}{dt} + \frac{1}{2} \sigma_t^2 \frac{d^2 f}{dx^2}] dt + \sigma_t \frac{df}{dx} dw_t \Rightarrow dy = [-\theta x_t e^{\theta t} + \theta x_t e^{\theta t}] dt + \sigma_t e^{\theta t} dw_t,$$

e

$$dy = \sigma_t e^{\theta t} dw_t. \quad (\text{A.3})$$

Integrando os dois lados da equação (A.3):

$$\begin{aligned} y_t - y_0 &= \sigma \int_0^t e^{\theta s} dw_s \\ x_t e^{\theta t} &= x_0 + \sigma \int_0^t e^{\theta s} dw_s \\ x_t &= x_0 e^{-\theta t} + \sigma \int_0^t e^{\theta(s-t)} dw_s \end{aligned}$$

Portanto, determina-se o processo OU com distribuição normal com média $x_0 e^{-\theta t}$ e variância $\frac{\sigma^2}{2\theta} (1 - e^{-2\theta t})$.

Portanto, a variância assintótica ou de longo prazo do processo OU pode ser definida como:

$$\lim_{t \rightarrow \infty} Var(x_t) = \frac{\sigma^2}{2\theta}.$$

Nota-se que quanto maior a velocidade de convergência à média, θ , menor a variância assintótica.

APÊNDICE B – *Spline* cúbico

Dados $n + 1$ pontos (nós) do plano, pode-se encontrar um polinômio de grau n que interpola esses pontos. Tal método é denominado aproximação por polinômios seccionados. Por definição, busca-se o menor grau n em cada intervalo entre os nós para o qual o polinômio é diferenciável nos nós mais extremos. No caso de $n = 3$, tem-se o polinômio cúbico ou *spline* cúbico.

Seja $f : [a, b] \rightarrow \mathbb{R}$ uma função e $a = x_0 < x_1 < \dots < x_n = b$ um conjunto de nós, o *spline* cúbico é uma função $S : [a, b] \rightarrow \mathbb{R}$ satisfazendo:

- $S(x)$ é um polinômio cúbico, indicado por $S_j(x)$ no intervalo $[x_j, x_{j+1}]$ para cada $0 \leq j \leq n - 1$;
- $S(x_j) = f(x_j)$, para cada $0 \leq j \leq n$;
- $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$, para cada $0 \leq j \leq n - 2$;
- $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$, para cada $0 \leq j \leq n - 2$;
- $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$, para cada $0 \leq j \leq n - 2$;

Seja $S_j(x)$ cada secção do *spline* para o intervalo $0 \leq j \leq n - 1$, cuja forma pode ser escrita:

$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$; sabendo que $S_j(x_j) = a_j = f(x_j)$, $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ e tomando $x_{j+1} - x_j = h_j$, tem-se:

$$a_{j+1} = a_j + b_j h_j + c_j h_j^2 + d_j h_j^3, 0 \leq j \leq n - 1; \quad (\text{B.1})$$

tomando $b_n = S'(x_n)$ e lembrando que $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$, tem-se:

$$b_{j+1} = b_j + 2c_j h_j + 3d_j h_j^2, 0 \leq j \leq n - 1; \quad (\text{B.2})$$

tomando $c_n = \frac{S''(x_n)}{2}$, tem-se:

$$c_{j+1} = c_j + 3d_j h_j, 0 \leq j \leq n - 1; \quad (\text{B.3})$$

isolando d_j em B.3 e substituindo em B.1 e B.2, tem-se:

$$a_{j+1} = a_j + b_j h_j + \frac{h_j^2}{3}(2c_j + c_{j+1}), 0 \leq j \leq n - 1; \quad (\text{B.4})$$

$$b_{j+1} = b_j + h_j(c_j + c_{j+1}), 0 \leq j \leq n - 1; \quad (\text{B.5})$$

isolando b_j em B.4 e tomando $j = j - 1$, tem-se:

$$h_{j-1}c_{j-1} + 2(h_{j-1} + h_j)c_j + h_jc_{j+1} = \frac{3}{h_j}(a_{j+1} - a_j) - \frac{3}{h_{j-1}}(a_j - a_{j-1}). \quad (\text{B.6})$$

Portanto, tem-se um sistema linear, cuja incógnita é o c_j .

APÊNDICE C – Teste de acurácia de previsão de Diebold e Mariano (DM)

Seja $f(e_{it})$ uma função perda definida como o erro quadrático de previsão do modelo i , logo: $f(e_{it}) = e_{it}^2$.

Seja $d_t = e_{1t}^2 - e_{2t}^2$ a diferença do erro de projeção dos modelos 1 e 2 testados, defini-se o teste de hipótese, sendo:

- $H_0 : E[d_t] = 0 \forall t$,
- $H_1 : E[d_t] \neq 0$,

tomando $\sqrt{T}(\bar{d} - \mu)$, sendo $\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$ a média amostral e $\mu = E[d_t]$ a média da população.

Denominado $f_d(0) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_d(k)$ como a densidade espectral da potência com frequência zero e $\gamma_d(k)$ a autocovariância com lag k , assumindo que o processo $\{d_t\}$ é estacionário.

Tem-se: $\sqrt{T}(\bar{d} - \mu) \sim N(0, 2\pi f_d(0))$, e pode-se definir a estatística teste como:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}}, \quad (\text{C.1})$$

onde $\hat{f}_d(0) = \frac{1}{2\pi} \hat{\gamma}_d(0)$ o estimador consistente de $f_d(0)$.

Logo, reescrevendo (C.1), obtém-se:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\gamma}_d(0)}{T}}}. \quad (\text{C.2})$$

Neste caso, pode-se considerar $\hat{\gamma}_d(0)$ como a variância da série história do d_t .

Por definição, a estatística teste DM é assintoticamente distribuída como uma normal padrão. Logo, rejeita-se H_0 se $|DM| > z_{\alpha/2}$, indicando que não há igualdade de acurácia nas projeções dos modelos.