



INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA

---

# Modelo de 2 fatores para a curva de contratos futuros do VIX: derivação, estimação e aplicações

Autor: **Felipe do Nascimento Fernandes**

Orientador: **Roberto Imbuzeiro Moraes Felinto de Oliveira**

Coorientador: **Fernando Antônio de Lucena Aiube**

Rio de Janeiro  
Janeiro de 2020



Para minha família.



## **Agradecimentos**

Agradeço, em primeiro lugar, a Deus e à minha família; em especial aos meus pais pelo carinho e paciência nos períodos mais turbulentos ao longo deste curso. Agradeço aos meus amigos pela compreensão nas inúmeras ausências por conta dos estudos. Agradeço aos professores Marco Antônio Cunha de Oliveira e Marcelo Camarão Ganem que me incentivaram e me ajudaram antes mesmo do início deste curso. Agradeço, em especial, a Fernando Antônio de Lucena Aiube, por sua paciência e didática ao me ajudar em todas as etapas na confecção deste trabalho. Agradeço, ainda, aos professores Jorge Zubelli, Roberto Imbuzeiro, Yuri Saporito e Rodrigo Targino por todo conhecimento compartilhado ao longo do mestrado. Deixo também minha gratidão aos grandes amigos que fiz nessa jornada: Lucas Paiva de Carvalho, Luiz Ricardo Martins, Clarissa Pontual Braga, Rafael Bessa, Yuri Resende Fonseca, Maurício Darós Andrade e, em especial, a Ciro Barbosa Paolucci, pelas incontáveis vezes que se disponibilizaram a tirar minhas dúvidas e me ajudar nos estudos. Sem vocês eu certamente não teria completado esta fase. Aos monitores Daniel Tordecilla e Dyego Soares que conseguiam sempre deixar claro o conteúdo apresentado. E, não poderia deixar de agradecer também aos meus colegas de trabalho da Icatu Seguros e Icatu Vanguarda, particularmente, aos meus gestores que souberam ser pacientes e generosos, permitindo que eu dividisse o meu tempo entre entregas profissionais e acadêmicas, tornando possível a conclusão desta etapa tão importante na minha vida.



## Resumo

Neste estudo, estimamos os futuros do VIX usando um modelo gaussiano, seguindo o artigo de Avellaneda e Papanicolau (2018). Utilizamos o modelo de dois fatores com ambas as variáveis seguindo um processo de Ornstein-Uhlenbeck que considera a estacionariedade, conforme observado a partir de dados empíricos. Adotamos duas estratégias diferentes: na primeira, seguimos rigorosamente a referência acima e incluímos o VIX à vista e todos os futuros do VIX como variáveis de observação (ou CMF - futuros de vencimento constante); na segunda abordagem, estimamos o modelo usando apenas os futuros do VIX. Nos dois casos, foi utilizada a metodologia do filtro de Kalman. Na segunda abordagem, encontramos erros menores na amostra. A estimação do modelo sem o VIX à vista envolveu um menor tempo computacional. A contribuição da dissertação com a estimação alternativa permite a simulação de instrumentos atrelados a contratos mais longos com menores erros. Uma vez estimados os hiperparâmetros, realizamos a simulação de alguns produtos, como ETFs e ETNs. Da perspectiva prática, essa modelagem é útil para entender o comportamento dos contratos futuros do VIX, pois todos os instrumentos derivados do VIX podem ser usados por agentes para estabelecer estratégias para proteger suas posições no mercado.

**Palavras-chave:** VIX Futuro, ETNs, ETFs, Modelos de fator gaussiano, Filtro de Kalman .



## **Abstract**

In this study we estimate the VIX futures using Gaussian models following the paper by Avellaneda and Papanicolau (2018). We use the two-factor model with both variables following an Ornstein-Uhlenbeck processes that accounts for the stationarity as observed from empirical data. We adopt two different strategies: in the first, we strictly followed the reference above and include the VIX spot and all VIX futures as observation variables, (or CMF - constant maturity futures); in the second approach we estimate the model using only VIX futures. In both cases the Kalman filter methodology was used. In the second approach we found smaller errors in sample. The estimation of the model without the VIX spot involved a shorter computational time. The dissertation's contribution with the alternative estimation allows the simulation of instruments linked to longer contracts with less errors. Once the hyperparameters were estimated we proceeded the simulation of some products such as ETFs and ETNs. From the practical perspective this modeling is useful to understand the behavior of futures contracts on VIX, because all the VIX instruments can be used by agents to stablish strategies to hedge their positions in the market.

**Key-words:** VIX Future, ETNs, ETFs, Gaussian factor models, Kalman Filter.



# Sumário

|  |           |
|--|-----------|
| Sumário  | ix        |
| Introdução   | 1         |
| <b>1 Revisão Bibliográfica</b>                                       | <b>3</b>  |
| 1.1 O índice VIX e a volatilidade como uma classe de ativo . . . . . | 3         |
| 1.2 Artigo objeto . . . . .  | 6         |
| <b>2 Modelo de 2 Fatores</b>   | <b>9</b>  |
| 2.1 Modelo na medida física (real) $\mathbb{P}$ . . . . .            | 9         |
| 2.2 Modelo na medida neutra ao risco $\mathbb{Q}$ . . . . .          | 13        |
| 2.3 Modelo na forma espaço-estado . . . . .                          | 15        |
| <b>3 Resultados</b>  | <b>19</b> |
| 3.1 Característica dos dados e obtenção das CMFs . . . . .           | 19        |
| 3.2 Estimação do modelo . . . . .                                    | 22        |
| 3.3 Análise dos resultados . . . . .                                 | 23        |
| <b>4 Aplicações aos ETPs</b>   | <b>29</b> |
| <b>5 Conclusão</b>   | <b>39</b> |
| Referências Bibliográficas   | 40        |
| <b>A Fórmula do VIX</b>  | <b>43</b> |
| <b>B Revisão Teórica</b>   | <b>51</b> |
| B.1 Probabilidade e Cálculo Estocástico . . . . .                    | 51        |
| B.2 Conceitos de finanças e apreçamento neutro ao risco . . . . .    | 55        |
| B.3 Filtro de Kalman . . . . .                                       | 59        |
| B.4 Método de Euler-Maruyama . . . . .                               | 64        |
| <b>C Gráficos: estimação</b>   | <b>67</b> |



# Introdução

Desde o seu advento, em 2004, os contratos futuros do VIX são colocados como importantes produtos para a diversificação e a proteção dos portfólios de investimento. Neste sentido, os profissionais de mercado podem negociá-los pelos mais diversos motivos, sejam eles: especulação, exposição direcional, arbitragem, diversificação ou até mesmo para *hedge* de volatilidade.

Mais recentemente, qualquer tipo de investidor — por exemplo, um fundo de pensão ou um indivíduo — passou a também ter acesso às negociações de volatilidade, como um ativo que compõe a sua carteira de investimento. Nas bolsas de valores americanas (e em algumas outras bolsas espalhadas pelo mundo), observou-se a ascensão de produtos denominados *Exchange Traded Notes* (ETNs) e *Exchange Traded Funds* (ETFs) que, juntos, compõem a classe dos *Exchange Traded Products* (ETPs) e podem ser encontrados para as mais diversas categorias de ativos (como ações, *commodities*, etc.). No caso da volatilidade, esses produtos são negociados nas bolsas e buscam replicar a rentabilidade de contratos futuros do VIX com maturidades (vencimentos) constantes.

Diante desse cenário, os pesquisadores Marco Avellaneda e Andrew Papanicolaou escreveram um artigo (Avellaneda and Papanicolaou [2018]) na qual estudam os principais aspectos associados aos contratos futuros do VIX e as implicações atreladas aos ETPs. Os autores propõem um modelo multifatorial para a dinâmica da estrutura a termo do VIX e derivam as fórmulas inerentes aos índices de rentabilidade que esses produtos tentam replicar. Essas contribuições mostram-se como um ferramental muito útil capaz de trazer um bom entendimento sobre a dinâmica evolutiva dos contratos futuros do VIX, além de permitir a simulação de cenários para analisar a rentabilidade esperada dos respectivos produtos.

Por conta disso, este trabalho possui como objetivo aprofundar o entendimento e a derivação do modelo proposto em Avellaneda and Papanicolaou [2018], além de obter uma estimativa para os parâmetros. Nesta etapa, vale destacar que realizamos dois procedimentos de estimação. No primeiro, seguimos a metodologia do artigo citado anteriormente (e denominamos estimação original) e, como uma forma de complementar a análise, sugerimos uma estimação (que chamamos de estimação alternativa) que nos remete ao caso clássico em que o preço à vista é uma variável latente (não observável). Tal proposição não afeta a robustez do modelo e nos parece mais adequada uma vez que o VIX à vista não é diretamente operado pelos *traders* de mercado. Além disso, na estimação alternativa observamos uma maior rapidez na convergência do filtro. De posse das estimativas dos parâmetros, somos capazes de realizar simulações e análises mais robustas para os ETPs. Para isso, será necessário fazer uma revisão das ferramentas matemáticas utilizadas para derivação do modelo e dos aspectos práticos inerentes aos respectivos produtos.

Sendo assim, este trabalho será dividido em 5 capítulos. O capítulo 1 explicará melhor o

que é o VIX e o que ele representa e falará também das características dos produtos (ETNs e ETFs). Em seguida, vamos revisar os principais pontos do estudo feito no artigo e explicitar a contribuição deste.

No capítulo 2, realizamos os cálculos para derivar o modelo proposto no artigo e o escrevemos na forma espaço-estado, de modo a implementar a otimização e obter estimativas para os parâmetros (com o Filtro de Kalman).

Em seguida, no capítulo 3, apresentamos o resultado do procedimento de estimação e propomos uma discussão para uma otimização alternativa à realizada no artigo original.

De posse dos valores dos parâmetros, no capítulo 4 podemos abordar com maior ênfase as estratégias que os ETPs realizam para replicar a rentabilidade dos contratos futuros de maturidade constante do VIX e derivar os respectivos índices, de modo a simular diversos cenários para os respectivos produtos.

Por último, o capítulo 5 traz as considerações finais sobre o trabalho e algumas sugestões de pesquisas futuras que podem ser interessantes diante de tudo o que aqui foi colocado.

Vale destacar também que algumas derivações complementares foram colocadas nos apêndices com o intuito de dar maior clareza a todos os aspectos associados ao processo de obtenção dos parâmetros para esse modelo. Em especial, o apêndice B vai apresentar algumas definições matemáticas inerentes às disciplinas de cálculo estocástico, estatística e outros tópicos em finanças que foram utilizados no decorrer do trabalho, de modo a dar maior entendimento ao leitor sobre as ferramentas necessárias para a derivação do modelo proposto.

# Capítulo 1

## Revisão Bibliográfica

### 1.1 O índice VIX e a volatilidade como uma classe de ativo

De acordo com CBOE [2018], o índice VIX foi criado em 1993 com o intuito de medir a volatilidade implícita esperada para os próximos 30 dias, nas opções *at-the-money* do índice S&P 100, e logo se tornou a principal referência para o mercado de volatilidade de ações nos EUA.

Em 2003, a metodologia inicial foi revista e o índice VIX passou a refletir uma nova maneira de medir a volatilidade esperada. Essa nova metodologia se baseia no índice S&P 500, o principal índice de ações dos EUA, e estima a volatilidade esperada, agregando os preços ponderados das opções de compra e de venda do SPX em uma ampla variedade de preços de exercício. Como foi definido um procedimento para replicar a exposição à volatilidade com um portfólio de opções SPX, essa nova metodologia transformou o índice VIX, de um conceito abstrato para um padrão prático a fim de utilizar a volatilidade como uma nova classe de ativo, facilitando sua negociação e a busca por *hedge*. Para se ter uma ideia mais clara de como o índice é calculado, colocamos a sua respectiva derivação no Apêndice A.

Na figura 1.1, apresentamos a evolução diária histórica dos preços de fechamento do VIX e do S&P 500. Nela, podemos observar as principais características presentes nos diversos modelos de volatilidade utilizados na literatura, por exemplo: *i) Clusters de Volatilidade* - Momentos de pequenas variações (estabilidade/baixa volatilidade) tendem a se suceder por um determinado período de tempo, assim como períodos de grandes variações (instabilidade/alta volatilidade); *ii) Efeito de Alavancagem* - Correlação negativa entre o preço do ativo e a volatilidade (para alguns mercados, principalmente ações e índice), ou seja, quando o preço cai a volatilidade aumenta, e vice-versa; *iii) Reversão à média* - Quando a volatilidade aumenta ou diminui, ela tende a reverter para algum nível de longo prazo.

Essa correlação negativa sugere um benefício de diversificação ao incluir a volatilidade em uma carteira de investimentos. Os contratos futuros e as opções do VIX foram projetados para oferecer pura exposição à volatilidade em um único pacote eficiente. Esses produtos foram lançados respectivamente em Março/2004 e Fevereiro/2006.

Vale ainda destacar que o VIX à vista não é diretamente negociado de modo que, para acessar o mercado de volatilidade, o *trader* tem que recorrer aos mercados futuros, de opções e, mais recentemente de ETFs/ETNs (todos eles dependentes do VIX futuro). Uma vez que os preços futuros do VIX convergem para o preço do VIX à vista no vencimento, eles fornecem

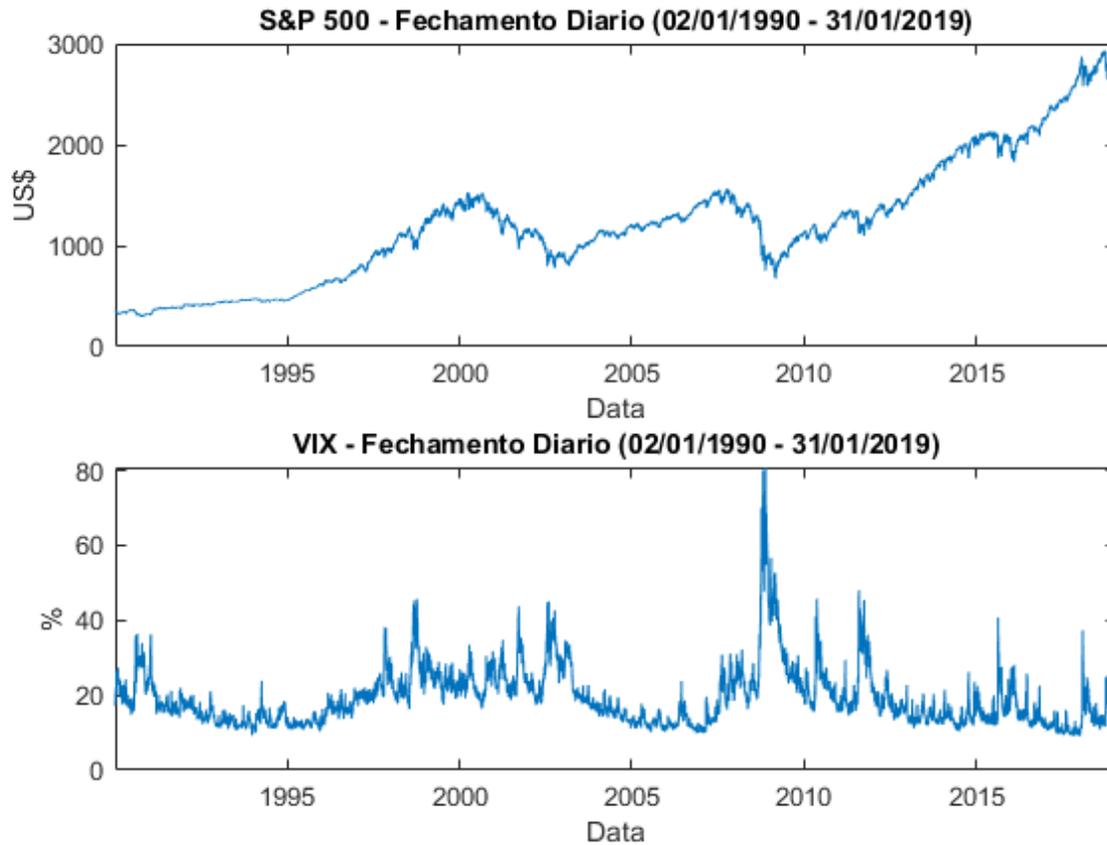


Figura 1.1: VIX e S&P 500 - preços de fechamento diário

uma expectativa de mercado negociável para onde o VIX estará na data de vencimento de cada futuro.

Por conta dessa característica, nesse trabalho, vamos voltar a atenção para os *Exchange Traded Products* - ETPs (*Exchange Traded Notes* - ETNs e *Exchange Traded Funds* - ETFs) do VIX. Conforme descrito em Hill et al. [2015] ETFs e ETNs são produtos negociados em bolsa (ETPs) vinculados ao valor de um índice subjacente e negociados em bolsas de valores globais. ETFs e ETNs diferem nos seguintes aspectos: *i) ETFs* - compartilham os recursos de investimento dos fundos mútuos. Um investidor compra ações de um ETF para possuir uma participação proporcional nos ativos agrupados. Como os fundos mútuos, os ETFs geralmente são gerenciados por um consultor de investimentos mediante uma taxa e regulados pela Lei das Sociedades de Investimento de 1940 (EUA). Os ETFs são respaldados por garantias e, portanto, normalmente têm um risco de contraparte menor do que um ETN; *ii) ETNs* - são obrigações de dívida não subordinadas não garantidas, estruturadas como uma promessa de pagamento de um padrão de retorno com base no retorno do índice declarado menos taxas de administração. Os ETNs são registrados de acordo com a Lei de Valores Mobiliários de 1933 (EUA). Se o emissor de um ETN declarar falência, o investidor está na mesma posição que todos os outros credores quirografários do emissor e pode perder parte ou todo o investimento.

É importante notar que os ETPs, em geral, fornecem aos investidores: liquidez, velocidade de execução, transparência e baixos custos de transação, difíceis de imitar com outros instrumentos; por isso, tais características têm atraído cada vez mais recursos para esse tipo de

produto. Para uma explicação mais detalhada da dinâmica operacional de funcionamento dos ETPs, sugerimos: Boroujerdi and Fogertey [2015].

No que diz respeito aos ETPs de volatilidade, podemos dizer que são classificados principalmente em termos de: *i) Maturidade* - esses produtos são comparados aos índices construídos com maturidades constantes, geralmente de 1 mês (*short-term*) ou 5 meses (*mid-term*), mas, a manutenção da respectiva estratégia requer um rebalanceamento diário de uma carteira de ativos que contém de 2 a 4 contratos futuros (como será visto ao longo deste trabalho); *ii) Alavancagem* - outro aspecto que classifica os ETPs do VIX está relacionado à sua alavancagem no retorno do benchmark subjacente. A maior parte dos ETPs é comparada a  $(+/- 1\times)$  ou  $(+/- 2\times)$  o retorno diário dos índices futuros de curto (*short-term*) e médio prazo (*mid-term*) do S&P 500 VIX. Um ETP que compra contratos futuros do VIX, com alavancagem  $+1$  se propõe a replicar o retorno diário do benchmark subjacente. Um ETP que compra contratos futuros do VIX com alavancagem  $+2$  tenta oferecer um retorno duas vezes maior que o retorno diário de seu índice de referência, enquanto retornos ETP inversos ou que vendem contratos futuros do VIX são múltiplos dos retornos negativos do índice de referência  $(-1\times, -2\times)$ .

Ao leitor que tiver interesse em explicações mais detalhadas das estratégias e análises mais profundas das respectivas performances históricas, indicamos Alexander and Korovilas [2012] e Gregory and Droumaguet [2017].

Ainda neste aspecto das classificações dos ETPs, podemos mostrar uma visão mais ampla de como esse mercado está organizado atualmente. Por meio de uma pesquisa feita ao site ETF.db [2019], observamos os principais produtos do mercado e quais as estratégias mais oferecidas aos investidores atualmente. Vimos que este mercado é muito concentrado, pois os 5 principais produtos possuem 77% do volume total sob gestão. Além disso, a estratégia de compra de volatilidade é a mais comum, com 85% dos ativos totais alocados nela e, por último, podemos notar que a busca por produtos que se propõem a replicar as variações de curto prazo respondem por 78% do total do mercado.

Podemos colocar também que a ascensão do mercado de ETPs para o VIX trouxe maior liquidez para os contratos futuros deste ativo; pois, apesar dos contratos futuros do VIX serem listados desde 2004, não observamos um volume significativo de negociações até o lançamento dos primeiros ETPs de VIX em 2009.

Diante dessa necessidade de um maior entendimento sobre a dinâmica dos contratos futuros do VIX (para conseguir aplicá-los nos produtos existentes no mercado), diversos estudos vêm sendo realizados. Um estudo de caráter mais inicial e menos formal em termos matemáticos (porém bem explicado) pode ser encontrado em Gregory and Timcenko [2016]. No presente trabalho, vamos utilizar o artigo escrito por Avellaneda e Papanicolaou (2018) como base para nos aprofundar no modelo proposto e nos resultados encontrados.

As contribuições dos autores em Avellaneda and Papanicolaou [2018] são: o modelo proposto e a derivação de fórmulas associadas aos índices de retorno dos ETPs. A partir das equações do modelo, podemos realizar simulações para o VIX à vista e para os contratos futuros, respeitando a dinâmica observada ao longo da série histórica. Ou seja, o modelo consegue replicar as variações dos contratos futuros do VIX, respeitando a dinâmica existente entre os vértices. Isso possibilita a simulação de cenários mais robustos (condizentes com a realidade). E, a partir das

equações propostas para os índices de retorno dos ETPs, somos capazes de simular também a rentabilidade esperada dos respectivos produtos.

Como será visto ao longo do trabalho, da perspectiva prática, essa modelagem é útil para entender o comportamento dos contratos futuros do VIX e mostra-se como um trabalho de grande importância, pois todos os instrumentos do VIX (opções, futuros e ETPs) podem ser usados pelos investidores para estabelecer estratégias a fim de proteger suas posições no mercado.

Dito isto, podemos passar a uma revisão mais detalhada do artigo estudado e à sua abordagem para escrever o VIX como um modelo de fatores.

## 1.2 Artigo objeto

Em Avellaneda and Papanicolaou [2018], os autores estudaram a dinâmica do VIX Futuro e, argumentam que a estrutura a termo dos contratos futuros com maturidade constante pode ser modelada como um processo estocástico estacionário; ou seja, a principal premissa da modelagem é que a série temporal dos CMFs é estacionária.

Como características associadas ao VIX e aos contratos futuros, os autores colocam que se espera que os contratos futuros com vencimentos mais longos (maturidades mais distantes) sejam superiores à média estatística de longo prazo do VIX, pois tal característica remete à teoria do prêmio de risco de volatilidade em Carr and Wu [2008].

Observaram, ainda, que o nível mais provável da distribuição empírica do VIX é significativamente menor do que a média e isso se mostra consistente com o fato de que a estrutura a termo dos contratos futuros, normalmente, possui uma inclinação positiva ou, como é mais comumente chamada, está em um estado de *contango*.

Mas, a presença de caudas pesadas para a distribuição também indica que, esporadicamente, o VIX à vista será maior que a sua média de longo prazo. Portanto, também existem períodos no qual essa inclinação é negativa ou está no estado de *backwardation*, seja parcial (somente para os primeiros vencimentos) ou totalmente, (em toda a estrutura a termo dos contratos de vencimento futuro). Nesses períodos de *backwardation*, podemos ver que o VIX aumenta rapidamente, impulsionado por notícias exógenas e pelo aumento nos prêmios das opções de curto prazo.

A justificativa para a utilização de 2 fatores é feita a partir da Análise de Componentes Principais (ACP) realizada no artigo. Os autores consideraram as flutuações da curva de futuros do VIX em torno de sua média de longo prazo. Isto é consistente com a suposição de reversão à média/estacionariedade. Em outras palavras, essa abordagem para a ACP destina-se a determinar as várias curvas necessárias para reconstruir a curva CMF em qualquer data específica, podendo ser vista como uma superposição de formas.

Tal técnica foi denominada pelos autores de *Equilibrium PCA* e proporciona uma identificação da curva do VIX sob a perspectiva do reconhecimento de imagem, na qual cada componente é visto como fornecendo uma forma ou característica que é sobreposta à média, com um coeficiente de carga correspondente.

Além disso, estudaram o comportamento dos produtos ETNs e ETFs que se baseiam em estratégias de rolagem sobre os CMFs e calibraram os dados históricos para um modelo multifatorial (mais precisamente, 2 fatores). Esses fatores compõem um modelo lognormal com propriedades de reversão à média.

Os autores argumentam também que, na análise das estratégias, perceberam que a operação mantida até o vencimento (*buy and hold*), na qual: *i*) Vende-se um ETN/ETF que rola posições compradas nos contratos futuros; ou *ii*) Compra-se um ETN/ETF que rola posições vendidas nos contratos futuros, produzem lucros (teoricamente) quase certos, assume-se que os CMFs possuem propriedades de estacionariedade e ergodicidade. Mas complementam dizendo que sabe-se, também, que o VIX e os CMFs apresentam períodos em que ficam em *backwardation*, e isso traz grandes variações aos resultados das estratégias.

Para um aprofundamento maior na literatura sobre a modelagem da volatilidade e a sua aplicação aos derivativos, recomendamos a leitura de Bergomi [2015].

Dito tudo isso, vamos agora detalhar a derivação do modelo proposto, de modo que seremos capazes de encontrar estimativas dos valores dos parâmetros e simular a rentabilidade esperada para os produtos de mercado (ETNs/ETFs). Para o leitor que quiser mais detalhes sobre a teoria matemática associada à essa derivação, sugerimos olhar o apêndice [B](#).



# Capítulo 2

## Modelo de 2 Fatores

Neste capítulo, vamos realizar a derivação dos cálculos inerentes ao modelo proposto por Avellaneda and Papanicolaou [2018]. Primeiro demonstraremos o processo estocástico que satisfaz a EDE (na medida real  $\mathbb{P}$ ) assumida para cada fator. Em seguida, vamos explicar a modelagem feita para o Preço de Risco do Mercado (MPR) e realizar a respectiva mudança de medida, a partir dos resultados encontrados no Teorema de Girsanov. Com isso, seremos capazes de aplicar a fórmula do preço de um contrato futuro para chegarmos à equação final. Por último, apresentamos o modelo na forma espaço-estado, que será a base para a implementação do Filtro de Kalman no processo de calibragem.

### 2.1 Modelo na medida física (real) $\mathbb{P}$

Nesta seção vamos derivar os cálculos por trás do modelo proposto em Avellaneda and Papanicolaou [2018] para replicar a dinâmica observada nas curvas dos contratos futuros do VIX com maturidades constantes (CMFs).

O modelo proposto é uma difusão multifatorial gaussiana, tal que:

$$\ln(VIX_t) = \sum_{i=1}^d X_{it} \Leftrightarrow VIX_t = \exp\left(\sum_{i=1}^d X_{it}\right) \quad (2.1.1)$$

Por hipótese, cada fator  $X_i$  segue uma dinâmica unidimensional dada pelo processo estocástico de Ornstein-Uhlenbeck (O-U):

$$dX_{it} = \kappa_i(\mu_i - X_{it})dt + \sigma_i dW_{it}^{\mathbb{P}}, \quad \kappa_i > 0, \quad i = 1, \dots, d \quad (2.1.2)$$

No processo estocástico correspondente a  $X_{it}$ :

$\kappa_i > 0$  representa a velocidade de reversão à média;

$\mu_i > 0$  representa o nível da média de longo prazo do processo;

$\sigma_i > 0$  representa a variância do processo;

$dW_{it}^{\mathbb{P}}$  representa o movimento browniano associado à variância, na medida  $\mathbb{P}$  (Física).

Aqui vale observar ainda que  $(W_{1t}^{\mathbb{P}}, \dots, W_{dt}^{\mathbb{P}})$  é um vetor  $d$  – dimensional de movimentos brownianos, com matriz de correlação  $\rho_{ij}$ .

Além disso, de acordo com este modelo, o  $VIX$  é estacionário com uma distribuição de equilíbrio *lognormal* com:

$$\mathbb{E}^{\mathbb{P}}[\ln(VIX)] = \sum_{i=1}^d \mu_i, \quad Var[\ln(VIX)] = \sum_{ij=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j}$$

E, por conta dessa *lognormalidade*, podemos escrever:

$$\mathbb{E}^{\mathbb{P}}[VIX] = \exp\left(\sum_{i=1}^d \mu_i + \frac{1}{2} \sum_{ij=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j}\right) \quad (2.1.3)$$

Diante desses fatos, vamos agora detalhar as contas relacionadas à resolução das equações diferenciais estocásticas e das principais estatísticas associadas (provando os resultados descritos acima).

Pela justificativa feita na Análise de Componentes Principais (PCA) do artigo, assumimos  $i = 1, \dots, d$  com  $d = 2$ . E, para encontrar o processo que satisfaz a dinâmica proposta anteriormente, primeiro identificamos o seguinte fator integrante:

$$y = Ce^{at} \Rightarrow f(t, X_{it}) = (\mu_i - X_{it})e^{\kappa_i t}$$

E, pelo lema de Itô (em B.1.7), precisamos calcular:

$$\begin{aligned} f_t(t, X_{it}) &= \frac{\partial f(t, X_{it})}{\partial t} = \kappa_i(\mu_i - X_{it})e^{\kappa_i t} \\ f_x(t, X_{it}) &= \frac{\partial f(t, X_{it})}{\partial X_{it}} = -e^{\kappa_i t} \\ f_{xx}(t, X_{it}) &= \frac{\partial^2 f(t, X_{it})}{\partial X_{it}^2} = 0 \end{aligned}$$

Com isso, escrevemos:

$$d[f(t, X_{it})] = \kappa_i(\mu_i - X_{it})e^{\kappa_i t} dt - e^{\kappa_i t} dX_{it}$$

E, substituindo a dinâmica proposta para  $dX_{it}$ , ficamos com:

$$\begin{aligned} d[f(t, X_{it})] &= \kappa_i(\mu_i - X_{it})e^{\kappa_i t} dt - e^{\kappa_i t} \left[ \kappa_i(\mu_i - X_{it})dt + \sigma_i dW_{it}^{\mathbb{P}} \right] \\ d[f(t, X_{it})] &= -e^{\kappa_i t} \sigma_i dW_{it}^{\mathbb{P}} \end{aligned}$$

Agora, integrando ambos os lados, obtemos:

$$f(t, X_{it}) - f(0, X_{i0}) = - \int_0^t e^{\kappa_i s} \sigma_i dW_{is}^{\mathbb{P}}$$

E, substituindo  $f(t, X_{it})$  e  $f(0, X_{i0})$ , teremos:

$$X_{it} = \mu_i + e^{-\kappa_i t} (X_{i0} - \mu_i) + \sigma_i \int_0^t e^{-\kappa_i(t-s)} dW_{is}^{\mathbb{P}} \quad (2.1.4)$$

Na qual  $X_{it}$  é o processo estocástico que satisfaz à dinâmica proposta pela EDE de O-U. A partir disso, podemos calcular a média e a variância do respectivo processo.

Sabemos ainda que a esperança da integral estocástica é igual a zero. Portanto:

$$\mathbb{E}^{\mathbb{P}}[X_{it}|X_{i0} = x_{i0}] = \mathbb{E}^{\mathbb{P}}\left[\mu_i + e^{-\kappa_i t}(X_{i0} - \mu_i) + \sigma_i \int_0^t e^{-\kappa_i(t-s)} dW_{is}^{\mathbb{P}} \middle| X_{i0} = x_{i0}\right]$$

Agora, pela linearidade da esperança, vendo que  $\mu_i$  não depende de  $t$  e, com o argumento do valor da integral estocástica, podemos reescrever:

$$\mathbb{E}^{\mathbb{P}}[X_{it}|X_{i0} = x_{i0}] = \mu_i + e^{-\kappa_i t}(x_{i0} - \mu_i) \quad (2.1.5)$$

e, analisando quando  $t \rightarrow \infty$ :

$$\lim_{t \rightarrow \infty} \mathbb{E}^{\mathbb{P}}[X_{it}|X_{i0} = x_{i0}] = \mu_i \quad (2.1.6)$$

Para o cálculo da matriz de covariância, escrevemos:

$$Cov^{\mathbb{P}}[X_{it}, X_{jt}] = \mathbb{E}^{\mathbb{P}}[(X_{it} - \mathbb{E}^{\mathbb{P}}[X_{it}])(X_{jt} - \mathbb{E}^{\mathbb{P}}[X_{jt}])]$$

De modo que, quando  $i = j$ , ficamos com:

$$Var^{\mathbb{P}}[X_{it}] = \mathbb{E}^{\mathbb{P}}[(X_{it} - \mathbb{E}^{\mathbb{P}}[X_{it}])^2]$$

Mas, pelo resultado encontrado em 2.1.5, e pela isometria de Itô, podemos verificar que:

$$\begin{aligned} Cov^{\mathbb{P}}[X_{it}, X_{jt}] &= \mathbb{E}^{\mathbb{P}}\left[\left(\sigma_i \int_0^t e^{-\kappa_i(t-s)} dW_{is}^{\mathbb{P}}\right)\left(\sigma_j \int_0^t e^{-\kappa_j(t-s)} dW_{js}^{\mathbb{P}}\right)\right] \\ &= \frac{\sigma_i \sigma_j \rho_{ij} e^{-(\kappa_i + \kappa_j)t}}{\kappa_i + \kappa_j} [e^{(\kappa_i + \kappa_j)t} - \underbrace{e^{(\kappa_i + \kappa_j)0}}_{=1}] \end{aligned}$$

De modo que ficamos com:

$$Cov^{\mathbb{P}}[X_{it}, X_{jt}] = \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j} (1 - e^{-2(\kappa_i + \kappa_j)t}) \quad (2.1.7)$$

E, quando  $t \rightarrow \infty$ :

$$\lim_{t \rightarrow \infty} Cov^{\mathbb{P}}[X_{it}, X_{jt}] = \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j} \quad (2.1.8)$$

Por último, quando  $i = j$ , temos que:

$$Var^{\mathbb{P}}[X_{it}] = \frac{\sigma_i^2}{2\kappa_i} (1 - e^{-2\kappa_i t}) \quad (2.1.9)$$

E, mais uma vez verificando quando  $t \rightarrow \infty$ :

$$\lim_{t \rightarrow \infty} Var^{\mathbb{P}}[X_{it}] = \frac{\sigma_i^2}{2\kappa_i} \quad (2.1.10)$$

Agora que já sabemos a média e a variância de  $X_t$ , podemos analisar a média e a variância do modelo proposto.

Uma vez que  $VIX_t = \exp\left(\sum_{i=1}^d X_{it}\right)$  e, por hipótese,  $\ln(VIX_t)$  segue uma difusão gaussiana multifatorial, podemos escrever:

$$VIX_t = \exp\left(\sum_{i=1}^d X_{it}\right) \Leftrightarrow \ln(VIX_t) = \ln\left(\exp\left(\sum_{i=1}^d X_{it}\right)\right) \Leftrightarrow \ln(VIX_t) = \sum_{i=1}^d X_{it}$$

Na qual:  $X_i \sim \mathcal{N}(\mu, \sigma^2)$ , já calculados anteriormente.

Além disso, sabemos que, se  $X \sim \mathcal{N}(\mu, \sigma^2)$  e, temos  $Y = e^X$ . Então,  $Y \sim \text{lognormal}$ , na qual  $\mathbb{E}[Y] = e^{(\mu + \frac{\sigma^2}{2})}$ . Portanto, escrevemos:

$$VIX_t \sim \text{lognormal} : \mathbb{E}^{\mathbb{P}}[VIX_t] = \exp\left(\mathbb{E}^{\mathbb{P}}[\ln(VIX_t)] + \frac{\text{Var}^{\mathbb{P}}[\ln(VIX_t)]}{2}\right) \quad (2.1.11)$$

E, calculando:

$$\mathbb{E}^{\mathbb{P}}[\ln(VIX_t)] = \mathbb{E}^{\mathbb{P}}\left[\sum_{i=1}^d X_{it}\right] = \sum_{i=1}^d \mathbb{E}^{\mathbb{P}}[X_{it}]$$

Mas, pelos cálculos realizados em 2.1.5 e 2.1.6, podemos escrever que "no equilíbrio de longo prazo", a esperança do modelo com  $d$  fatores é dada por:

$$\lim_{t \rightarrow \infty} \mathbb{E}^{\mathbb{P}}[\ln(VIX_t)] = \sum_{i=1}^d \mu_i \quad (2.1.12)$$

Já para a matriz de covariâncias, teríamos:

$$\text{Var}^{\mathbb{P}}[\ln(VIX_t)] = \text{Var}^{\mathbb{P}}\left[\sum_{i=1}^d X_{it}\right] = \sum_{i=1}^d \text{Cov}^{\mathbb{P}}[X_{it}, X_{jt}]$$

E, novamente, pelo que já foi calculado em 2.1.7 e 2.1.8, quando  $t \rightarrow \infty$ , a variância do modelo será:

$$\lim_{t \rightarrow \infty} \text{Var}^{\mathbb{P}}[\ln(VIX_t)] = \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j} \quad (2.1.13)$$

Como já havíamos colocado em 2.1.11 e, acrescentando o que foi calculado em 2.1.12 e 2.1.13, no "equilíbrio de longo prazo", chegamos ao mesmo resultado enunciado em 2.1.3:

$$\lim_{t \rightarrow \infty} \mathbb{E}^{\mathbb{P}}[VIX] = \exp\left(\sum_{i=1}^d \mu_i + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j}\right)$$

## 2.2 Modelo na medida neutra ao risco $\mathbb{Q}$

Nesta seção, apresentamos a escolha do Preço de Risco do Mercado (*Market Price of Risk* - MPR) e realizamos a mudança de medida, conforme explicado em B.2.6.

Para estender esse modelo para os CMFs, assume-se que existe uma medida de probabilidade  $\mathbb{Q}$  (Neutra ao Risco) e um processo  $d$ -dimensional e  $W^{\mathbb{P}}$ -adaptado  $\lambda_t$ , que é o MPR, na qual:

Os processos  $dW_{it}^{\mathbb{Q}} := dW_{it}^{\mathbb{P}} + \lambda_{it}dt$  com  $i = 1, \dots, d$ , são Movimentos Brownianos Padrão sob a medida  $\mathbb{Q}$  (com correlação  $\rho_{ij}$ );

O preço de cada contrato futuro (assim como em B.2.11) é o valor esperado na medida  $\mathbb{Q}$  (Neutra ao Risco) do  $VIX$ , na data de vencimento correspondente, ou seja:

$$V_t^\tau = \mathbb{E}_t^{\mathbb{Q}}[VIX_{t+\tau}] = \mathbb{E}_t^{\mathbb{Q}} \left[ \exp \left( \sum_{i=1}^d X_{it+\tau} \right) \right]$$

Em Avellaneda and Papanicolaou [2018], os autores escolheram os MPRs como funções lineares dos fatores, conforme abaixo:

$$\lambda_{it} = p_i + q_i X_{it} \quad (2.2.1)$$

Onde  $p_i$  e  $q_i$  são constantes. Mas, vale destacar que, com essas hipóteses, o MPR não será constante ao longo do tempo. A partir dessas considerações, podemos realizar o cálculo da mudança de medida e verificar o respectivo processo que irá satisfazer essa nova dinâmica.

Temos a seguinte dinâmica proposta na medida  $\mathbb{P}$ :

$$dX_{it} = \kappa_i(\mu_i - X_{it})dt + \sigma_i dW_{it}^{\mathbb{P}}$$

Com o respectivo prêmio de risco de mercado assumido e, pelo Teorema de Girsanov (em B.2.4), sabemos que:

$$dW_{it}^{\mathbb{Q}} := dW_{it}^{\mathbb{P}} + \lambda_{it}dt \Rightarrow dW_{it}^{\mathbb{P}} = dW_{it}^{\mathbb{Q}} - \lambda_{it}dt$$

Portanto, podemos escrever (substituindo  $dW_{it}^{\mathbb{P}}$  na nossa equação inicial):

$$\begin{aligned} dX_{it} &= \kappa_i(\mu_i - X_{it})dt + \sigma_i(dW_{it}^{\mathbb{Q}} - \lambda_{it}dt) \\ &= (\kappa_i + \sigma_i q_i) \left[ \frac{\kappa_i \mu_i - \sigma_i p_i}{\kappa_i + \sigma_i q_i} - X_{it} \right] dt + \sigma_i dW_{it}^{\mathbb{Q}} \end{aligned}$$

Agora, tomando:

$$\bar{\kappa}_i = \kappa_i + \sigma_i q_i \quad e \quad \bar{\mu}_i = \frac{\kappa_i \mu_i - \sigma_i p_i}{\kappa_i + \sigma_i q_i}$$

Vemos que, na Medida Martingal Equivalente (MME), a dinâmica para o processo  $X_{it}$  é dada por:

$$dX_{it} = \bar{\kappa}_i(\bar{\mu}_i - X_{it})dt + \sigma_i dW_{it}^{\mathbb{Q}} \quad (2.2.2)$$

E, seguindo os mesmos cálculos realizados na seção anterior, ficaremos com:

$$X_{it} = \bar{\mu}_i + e^{-\bar{\kappa}_i t}(X_{i0} - \bar{\mu}_i) + \sigma_i \int_0^t e^{-\bar{\kappa}_i(t-s)} dW_{is}^{\mathbb{Q}} \quad (2.2.3)$$

Na qual  $X_{it}$  é o processo estocástico que satisfaz à dinâmica proposta pela EDE de O-U, na medida  $\mathbb{Q}$  (neutra ao risco). De maneira mais genérica, podemos ainda escrever:

$$X_{it+\tau} = e^{-\bar{\kappa}_i \tau} X_{it} + (1 - e^{-\bar{\kappa}_i \tau}) \bar{\mu}_i + \sigma_i \int_t^{t+\tau} e^{-\bar{\kappa}_i(t+\tau-s)} dW_{is}^{\mathbb{Q}} \quad (2.2.4)$$

E, de maneira análoga, temos a média do respectivo processo:

$$\mathbb{E}^{\mathbb{Q}}[X_{it}|X_{i0} = x_{i0}] = \bar{\mu}_i + e^{-\bar{\kappa}_i t}(x_{i0} - \bar{\mu}_i) \quad (2.2.5)$$

e, analisando quando  $t \rightarrow \infty$ :

$$\lim_{t \rightarrow \infty} \mathbb{E}^{\mathbb{Q}}[X_{it}|X_{i0} = x_{i0}] = \bar{\mu}_i$$

E, como sabemos, a matriz de covariância é a mesma (nas duas medidas).

Agora que já calculamos os processos correspondentes na medida  $\mathbb{Q}$  (neutra ao risco) e os 2 primeiros momentos dessa variável aleatória, podemos colocar que, de acordo com B.2.11, calculamos o preço do contrato futuro do VIX como o valor esperado (na medida  $\mathbb{Q}$ ) do VIX em uma determinada data de expiração (vencimento), tal que:

$$V_t^\tau = \mathbb{E}_t^{\mathbb{Q}}[VIX_{t+\tau} | \mathcal{F}_t] = \mathbb{E}_t^{\mathbb{Q}} \left[ \exp \left( \sum_{i=1}^d X_{it+\tau} \right) \middle| \mathcal{F}_t \right] \quad (2.2.6)$$

Novamente, pela definição da esperança de uma variável aleatória *lognormal*, teremos:

$$\begin{aligned} V_t^\tau &= \mathbb{E}_t^{\mathbb{Q}}[VIX_{t+\tau} | \mathcal{F}_t] \\ &= \exp \left( \mathbb{E}_t^{\mathbb{Q}}[\ln(VIX_{t+\tau}) | \mathcal{F}_t] + \frac{1}{2} \text{Var}[\ln(VIX_{t+\tau}) | \mathcal{F}_t] \right) \\ &= \exp \left( \sum_{i=1}^d \bar{\mu}_i + e^{-\bar{\kappa}_i \tau}(X_{it} - \bar{\mu}_i) + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} (1 - e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau}) \right) \\ \ln(V_t^\tau) &= \sum_{i=1}^d \bar{\mu}_i + e^{-\bar{\kappa}_i \tau}(X_{it} - \bar{\mu}_i) + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} (1 - e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau}) \\ &= \sum_{i=1}^d \bar{\mu}_i + e^{-\bar{\kappa}_i \tau}(X_{it} - \bar{\mu}_i) + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \\ &= \underbrace{\sum_{i=1}^d \bar{\mu}_i + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j}}_{\ln(V^\infty)} + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau}(X_{it} - \bar{\mu}_i) - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \end{aligned}$$

Chegando no processo proposto em Avellaneda and Papanicolaou [2018] para a curva de contratos futuros com maturidade constante do VIX (CMF), na medida neutra ao risco:

$$\ln(V_t^\tau) = \ln(V^\infty) + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (X_{it} - \bar{\mu}_i) - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \quad (2.2.7)$$

Na qual,  $V^\infty = \sum_{i=1}^d \bar{\mu}_i + \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j}$  representa o valor assintótico da curva CMF quando  $\tau \rightarrow \infty$ .

Neste ponto do artigo original, os autores fazem ainda uma simplificação adicional. Como  $X_{it} - \bar{\mu}_i$  possui média zero sob a medida  $\mathbb{Q}$ , definimos:  $Y_{it} := X_{it} - \bar{\mu}_i$ . Usando a relação entre as medidas  $\mathbb{P}$  e  $\mathbb{Q}$ , pode-se derivar uma equação diferencial estocástica para  $Y_i$  na medida  $\mathbb{P}$ , de modo que:

$$\begin{aligned} dY_{it} &= dX_{it} = \kappa_i(\mu_i - X_{it})dt + \sigma_i dW_{it}^{\mathbb{P}} \\ &= \kappa_i(\mu_i^Y - Y_{it})dt + \sigma_i dW_{it}^{\mathbb{P}} \end{aligned}$$

onde:  $\mu_i^Y = \mu_i - \bar{\mu}_i$

Dessa forma, ficamos com:

$$dY_{it} = \kappa_i(\mu_i^Y - Y_{it})dt + \sigma_i dW_{it}^{\mathbb{P}} \quad (2.2.8)$$

E, como consequência, nosso modelo terá a seguinte formulação na medida real ( $\mathbb{P}$ ):

$$V_t^\tau = V^\infty \exp \left( \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} Y_{it} - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \right) \quad (2.2.9)$$

Portanto, de acordo com Avellaneda and Papanicolaou [2018], a posição de equilíbrio da curva CMF é dada pela média de longo prazo de  $Y_{it}$ , que é  $\mu_i^Y$ . Logo, teremos:

$$V_{eq}^\tau = V^\infty \exp \left( \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} \mu_i^Y - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \right) \quad (2.2.10)$$

## 2.3 Modelo na forma espaço-estado

Nesta seção, vamos escrever o modelo proposto na forma espaço-estado (formulação geral do problema para implementação do Filtro de Kalman), conforme descrito em B.3.1.

Da equação para o CMF em 2.2.7, podemos reorganizar os termos e escrever a seguinte forma para o modelo proposto:

$$V_{it}^T = \exp \left[ \sum_{i=1}^d \bar{\mu}_i + (X_{it} - \bar{\mu}_i) e^{-\bar{\kappa}_i (T-t)} + \frac{1}{2} \sum_{i,j=1}^d \frac{\rho_{ij} \sigma_i \sigma_j}{\bar{\kappa}_i + \bar{\kappa}_j} \left( 1 - e^{-(\bar{\kappa}_i + \bar{\kappa}_j)(T-t)} \right) \right] \quad (2.3.1)$$

Portanto, como assumimos  $d = 2$ , podemos escrever que, para dois fatores:

$$\begin{aligned} \ln(V_{t,T}) &= \bar{\mu}_1 + (X_{1t} - \bar{\mu}_1)e^{-\bar{\kappa}_1(T-t)} + \bar{\mu}_2 + (X_{2t} - \bar{\mu}_2)e^{-\bar{\kappa}_2(T-t)} \\ &\quad + \frac{1}{2} \frac{\sigma_1^2}{2\bar{\kappa}_1} \left(1 - e^{-2\bar{\kappa}_1(T-t)}\right) + \frac{1}{2} \frac{\sigma_2^2}{2\bar{\kappa}_2} \left(1 - e^{-2\bar{\kappa}_2(T-t)}\right) \\ &\quad + \frac{1}{2} \frac{\rho_{12}\sigma_1\sigma_2}{\bar{\kappa}_1 + \bar{\kappa}_2} \left(1 - e^{-(\bar{\kappa}_1+\bar{\kappa}_2)(T-t)}\right) + \frac{1}{2} \frac{\rho_{21}\sigma_2\sigma_1}{\bar{\kappa}_2 + \bar{\kappa}_1} \left(1 - e^{-(\bar{\kappa}_2+\bar{\kappa}_1)(T-t)}\right) \end{aligned}$$

Assim:

$$\begin{aligned} \ln(V_{t,T}) &= X_{1t}e^{-\bar{\kappa}_1(T-t)} + X_{2t}e^{-\bar{\kappa}_2(T-t)} + \bar{\mu}_1 \left(1 - e^{-\bar{\kappa}_1(T-t)}\right) + \bar{\mu}_2 \left(1 - e^{-\bar{\kappa}_2(T-t)}\right) \\ &\quad + \frac{\sigma_1^2}{4\bar{\kappa}_1} \left(1 - e^{-2\bar{\kappa}_1(T-t)}\right) + \frac{\sigma_2^2}{4\bar{\kappa}_2} \left(1 - e^{-2\bar{\kappa}_2(T-t)}\right) + \frac{\rho_{12}\sigma_1\sigma_2}{\bar{\kappa}_1 + \bar{\kappa}_2} \left(1 - e^{-(\bar{\kappa}_1+\bar{\kappa}_2)(T-t)}\right) \end{aligned} \quad (2.3.2)$$

Além disso, podemos notar que, no caso do contrato CMF com maturidade zero, temos o  $VIX$  (fazendo  $t = T$  na equação acima, pois estamos no vencimento):

$$\ln(VIX_t) = X_{1t} + X_{2t}$$

Escrevendo  $\mathbb{V}_t$  como o vetor que representa o  $\ln V_{t,T}$ , temos:

$$\mathbb{V}_t = \begin{bmatrix} \ln(VIX_t) \\ \ln(V_{t,T_1}) \\ \vdots \\ \ln(V_{t,T_n}) \end{bmatrix}$$

Pela formulação em B.3.2, a equação 2.3.2 pode ser representada na forma "Espaço" do Filtro de Kalman como:

$$\mathbb{V}_t = \mathbf{H}_t \mathbf{X}_t + \mathbf{d}_t + \epsilon_t$$

onde:

$$\mathbf{H}_t = \begin{bmatrix} 1 & 1 \\ e^{-\bar{\kappa}_1(T_1-t)} & e^{-\bar{\kappa}_2(T_2-t)} \\ \vdots & \vdots \\ e^{-\bar{\kappa}_1(T_n-t)} & e^{-\bar{\kappa}_2(T_n-t)} \end{bmatrix}, \mathbf{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix}, \mathbf{d}_t = \begin{bmatrix} 0 \\ A(T_1 - t) \\ \vdots \\ A(T_n - t) \end{bmatrix}, \epsilon_t \sim iid\mathcal{N}(0, \mathbb{Q}),$$

$$\mathbb{Q} = \begin{bmatrix} \sigma_{\epsilon_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\epsilon_n}^2 \end{bmatrix}$$

Na qual:

$$\begin{aligned} A(T_i - t) &= \bar{\mu}_1(1 - e^{-\bar{\kappa}_1(T_i-t)}) + \bar{\mu}_2(1 - e^{-\bar{\kappa}_2(T_i-t)}) + \frac{\sigma_1^2}{4\bar{\kappa}_1} \left(1 - e^{-2\bar{\kappa}_1(T_i-t)}\right) + \frac{\sigma_2^2}{4\bar{\kappa}_2} \left(1 - e^{-2\bar{\kappa}_2(T_i-t)}\right) \\ &\quad + \frac{\rho_{12}\sigma_1\sigma_2}{\bar{\kappa}_1 + \bar{\kappa}_2} \left(1 - e^{-(\bar{\kappa}_1+\bar{\kappa}_2)(T_i-t)}\right) \end{aligned}$$

E,  $T_i$  representa as respectivas maturidades constantes.

Por premissa do modelo, as dinâmicas das variáveis de estado ( $d = 2$ ) são dadas por:

$$\begin{aligned} dX_{1t} &= \kappa_1(\mu_1 - X_{1t})dt + \sigma_1dW_{1t} \\ dX_{2t} &= \kappa_2(\mu_2 - X_{2t})dt + \sigma_2dW_{2t} \end{aligned}$$

De modo que, discretizando-as, ficamos com:

$$\begin{aligned} X_{1t} - X_{1t-\Delta t} &= \kappa_1(\mu_1 - X_{1t-\Delta t})\Delta t + \Delta\nu_{1t} \\ X_{2t} - X_{2t-\Delta t} &= \kappa_2(\mu_2 - X_{2t-\Delta t})\Delta t + \Delta\nu_{2t} \end{aligned}$$

ou, ainda:

$$\begin{aligned} X_{1t} &= (1 - \kappa_1\Delta t)X_{1t-\Delta t} + \kappa_1\mu_1\Delta t + \Delta\nu_{1t} \\ X_{2t} &= (1 - \kappa_2\Delta t)X_{2t-\Delta t} + \kappa_2\mu_2\Delta t + \Delta\nu_{2t} \end{aligned}$$

Então, podemos reescrever as equações na forma estado do Filtro de Kalman da mesma maneira que em [B.3.3](#):

$$\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-\Delta t} + \mathbf{C} + \nu_t$$

onde:

$$\begin{aligned} \mathbf{X}_t &= \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 1 - \kappa_1\Delta t & 0 \\ 0 & 1 - \kappa_2\Delta t \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \kappa_1\mu_1\Delta t \\ \kappa_2\mu_2\Delta t \end{bmatrix}, \nu_t \sim iid\mathcal{N}(0, \mathbf{\Pi}), \\ \mathbf{\Pi} &= \begin{bmatrix} \frac{\sigma_1^2}{2\kappa_1} \left( 1 - e^{-2\kappa_1\Delta t} \right) & \frac{\rho_{12}\sigma_1\sigma_2}{\kappa_1 + \kappa_2} \left( 1 - e^{-(\kappa_1 + \kappa_2)\Delta t} \right) \\ \frac{\rho_{12}\sigma_1\sigma_2}{\kappa_1 + \kappa_2} \left( 1 - e^{-(\kappa_1 + \kappa_2)\Delta t} \right) & \frac{\sigma_2^2}{2\kappa_2} \left( 1 - e^{-2\kappa_2\Delta t} \right) \end{bmatrix} \end{aligned}$$

Dessa forma, com o entendimento completo da construção do modelo e a sua estrutura na forma espaço-estado, podemos, nos próximos capítulos, detalhar os aspectos práticos da implementação para calibragem, análise dos resultados e consequente aplicação aos produtos encontrados no mercado.



# Capítulo 3

## Resultados

Nesta seção, vamos detalhar um pouco mais o tratamento realizado nos dados para realizar a calibragem do modelo e, em seguida, comentar os resultados obtidos. Este capítulo está dividido como segue: na primeira parte, vamos falar das características dos dados da amostra e das possíveis formas de se obter as CMFs; em seguida, vamos colocar alguns aspectos do processo de implementação da calibragem e do teste alternativo proposto visando a complementar as análises feitas em Avellaneda and Papanicolaou [2018]; e, por último, vamos apresentar o valor dos parâmetros obtidos após a otimização, realizando um comparativo entre os erros.

### 3.1 Característica dos dados e obtenção das CMFs

Neste trabalho, utilizamos uma amostra com 510 observações (obtida na plataforma da *Bloomberg*). Esta base compreende o período entre 21 de abril de 2008 e 28 de janeiro de 2019 num intervalo semanal. Vale ainda destacar que os contratos futuros do VIX possuem sua negociação com o seguinte formato atualmente: 9 datas mensais de liquidação e 6 contratos adicionais que liquidam semanalmente (na parte inicial da curva).

Aqui, vamos concentrar o estudo apenas nos contratos mensais e, mais especificamente, por questões de volume de negociação, utilizou-se apenas o VIX à vista e os 7 contratos futuros subsequentes. Para capturar a evolução da estrutura a termo dos contratos futuros do VIX, é conveniente trabalhar com o conceito de Contratos Futuros com Maturidades Constantes (*Constant Maturity Futures - CMF*).

Dessa forma, podemos definir  $V_t^\tau$  como o preço teórico do contrato futuro que possui maturidade  $\tau$ , observado na data  $t$ .

Na prática,  $V_t^\tau$  é definido como uma interpolação linear entre os dois contratos mais próximos, um com maturidade imediatamente anterior e outro com maturidade imediatamente posterior à  $\tau$ .

Se  $\tau$  for mais curto que a maturidade do primeiro vencimento mensal,  $V^\tau$  será a interpolação linear entre o VIX à vista e o preço do primeiro contrato e, utilizando a notação tal que  $F_1$  e  $F_2$  são os preços dos contratos mensais (ou, o VIX e o primeiro contrato), com maturidades  $\tau_1$  e  $\tau_2$ , de modo que  $\tau_1 \leq \tau \leq \tau_2$ , temos:

$$V^\tau = \frac{\tau_2 - \tau}{\tau_2 - \tau_1} F_1 + \frac{\tau - \tau_1}{\tau_2 - \tau_1} F_2 \quad (3.1.1)$$

Caso  $F_1 = VIX$ , então  $\tau_1 = 0$ .

Apesar de ser a mais comum, a interpolação linear não é a única opção para esta etapa e outros métodos podem ser utilizados.

Aqui vale ainda destacar que nossa principal premissa de modelagem (assim como em Avelaneda and Papanicolaou [2018]) é que a série temporal dos CMFs é estacionária, ou seja, para cada  $\tau$  (maturidade constante) a série temporal  $V_t^\tau$ , para  $0 < t < \infty$  pode ser modelada como um processo estacionário.

Desta forma, somos capazes de construir a CMF a partir dos dados disponíveis para cada dia de negociação. No caso deste trabalho, para agilizar esta etapa, extraímos séries da *Bloomberg* que já possuíam esse tratamento em contratos futuros genéricos do VIX. Abaixo, apresentamos um gráfico dessa evolução temporal (Figura 3.1) e quatro dias diferentes da CMF que estão presentes na nossa amostra (Figura 3.2), apenas para dar um entendimento mais intuitivo sobre os dados utilizados. Vale ressaltar que, neste último, conseguimos observar dois dias (02/06/2014 e 04/05/2015) da estrutura a termo em *contango* e dois dias (05/02/2018 e 31/12/2018) em *backwardation*.

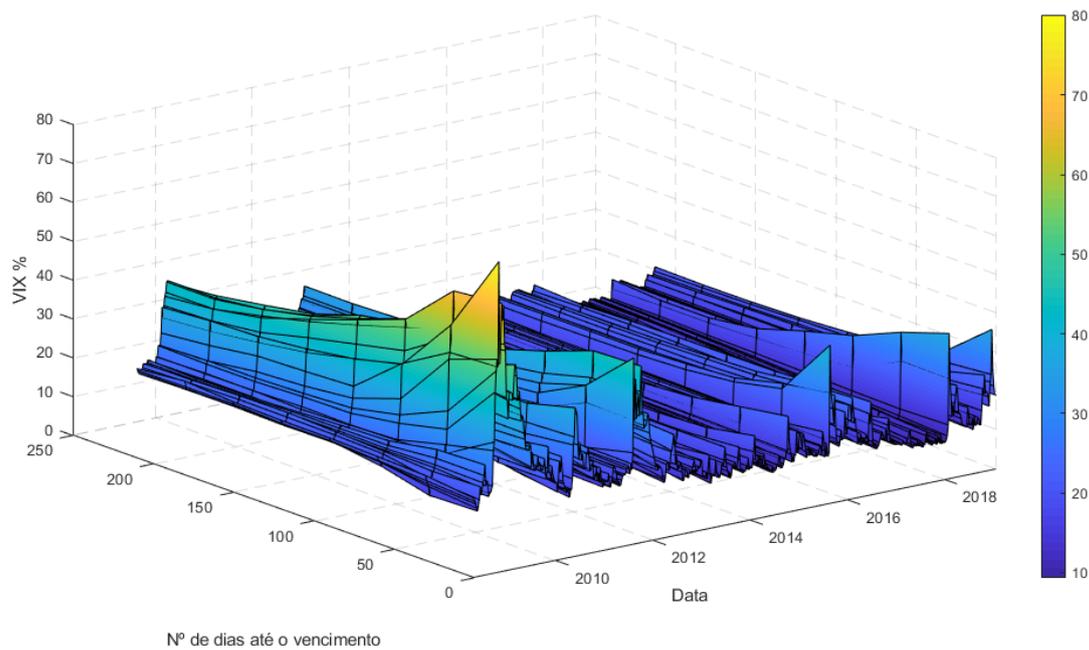


Figura 3.1: Evolução Temporal dos CMFs VIX após procedimento de Interpolação

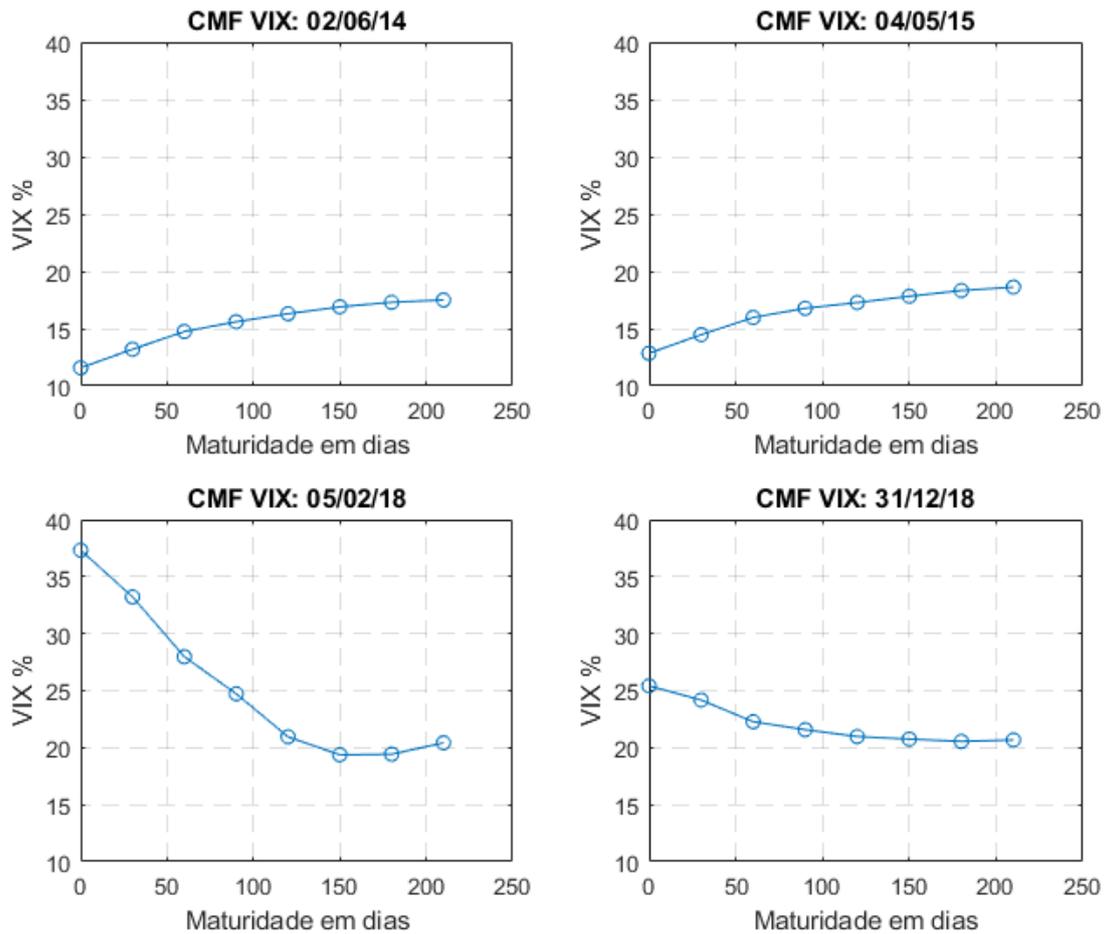


Figura 3.2: Estrutura a Termo do VIX para determinados dias da amostra

Depois desse tratamento para calcular a Estrutura a Termo dos Contratos Futuros com Maturidades Constantes do VIX, é possível observar também uma série que corresponde ao tempo ao vencimento de um determinado contrato futuro. Para o contrato mais próximo, esse vencimento varia de 1 até 30 dias corridos, e o mesmo raciocínio se aplica às demais séries de tempo ao vencimento (todas coerentes com as maturidades constantes escolhidas para os 7 primeiros contratos futuros).

Por fim, podemos também apresentar uma tabela com as principais estatísticas das séries das CMFs, inerentes à nossa amostra.

| Estatística         | VIX     | 30d     | 60d    | 90d    | 120d   | 150d   | 180d   | 210d   |
|---------------------|---------|---------|--------|--------|--------|--------|--------|--------|
| N° Obs.:            | 510     | 510     | 510    | 510    | 510    | 510    | 510    | 510    |
| mínimo (%)          | 9,34    | 9,88    | 11,33  | 12,25  | 13,23  | 13,78  | 14,23  | 14,83  |
| máximo (%)          | 80,06   | 67,95   | 56,29  | 52,72  | 48,63  | 46,44  | 44,49  | 44,06  |
| média (%)           | 19,85   | 20,18   | 20,96  | 21,50  | 21,87  | 22,19  | 22,48  | 22,76  |
| mediana (%)         | 16,84   | 17,30   | 18,48  | 19,24  | 19,68  | 20,03  | 20,34  | 20,54  |
| desvio-padrão (%)   | 10,03   | 8,96    | 7,96   | 7,28   | 6,78   | 6,43   | 6,15   | 5,96   |
| variância (ano) (%) | 15,98   | 12,73   | 10,06  | 8,41   | 7,30   | 6,55   | 6,01   | 5,64   |
| assimetria          | 2,37    | 2,19    | 1,86   | 1,63   | 1,43   | 1,27   | 1,16   | 1,10   |
| curtose             | 6,97    | 5,61    | 3,97   | 3,04   | 2,06   | 1,33   | 0,90   | 0,74   |
| Jarque-Bera         | 1523,76 | 1086,12 | 636,45 | 427,64 | 267,45 | 176,53 | 131,89 | 115,83 |

Tabela 3.1: Estatísticas das Séries dos CMFs

## 3.2 Estimação do modelo

A calibragem faz uso do filtro de Kalman, pois se trata de um modelo linear e Gaussiano. A utilização desta ferramenta está disponível em vários *softwares*, em geral no contexto relativo aos modelos espaço-estado. Mais especificamente, neste trabalho, utilizamos o *software* livre *RStudio*, com funções que executam o algoritmo do filtro já implementadas em pacotes. O pacote utilizado foi o **astsa** (*Applied Statistical Time Series Analysis*), que acompanha a referência Shumway and Stoffer [2017].

Destacamos que, na referência principal deste trabalho (Avellaneda and Papanicolaou [2018]), os autores realizam uma estimação utilizando as séries do VIX à vista em conjunto com os 7 primeiros vencimentos futuros (CMF). Além de replicar esse procedimento (original), pensou-se também em uma análise complementar de modo que realizamos uma outra estimação (alternativa), utilizando apenas os 7 vencimentos futuros, ou seja, excluindo-se da amostra a série histórica dos log-preços do VIX à vista. Vale ressaltar uma maior dificuldade para a convergência no caso em que a amostra continha os dados do VIX à vista. No caso alternativo, observamos uma redução no tempo de processamento para a convergência.

A justificativa para essa formulação alternativa vem do fato de que o VIX à vista não é diretamente operado pelos *traders* do mercado, ou seja, a exposição a essa classe de ativos (volatilidade) só é possível via ETPs, contratos futuros ou opções, todas elas dependentes dos preços futuros. Vale notar também (pela análise da tabela 3.1) que esta série é a que possui maior variação em toda a amostra, o que nos leva a pensar que sua utilização acaba trazendo mais erros para as estimativas totais.

Sendo assim, a não-inclusão dessa série na estimação para encontrar os respectivos parâmetros nos leva a pensar em uma possível redução nos erros das aproximações, sem influenciar na robustez do modelo.

Um argumento adicional vem do mercado de *commodities*, no qual o mesmo procedimento de filtração é realizado para se obter as estimativas dos parâmetros, e a negociação dos preços à vista das respectivas mercadorias também não é possível. Como nesse mercado realiza-se a estimação apenas nas séries dos contratos futuros e, a partir dos parâmetros obtidos, calcula-se um preço à vista, optou-se aqui por realizar este mesmo procedimento, com o benefício de que para o VIX, possuímos a série histórica realizada e podemos verificar a robustez do modelo.

Sendo assim, podemos colocar que, para o caso da estimação original (com a série histórica do VIX à vista), temos o vetor  $\Theta$  dos hiperparâmetros (com 19 parâmetros) que será estimado:

$$\Theta = \{\kappa_1, \kappa_2, \sigma_1, \sigma_2, p_1, p_2, q_1, q_2, \rho, \mu_1, \mu_2, \sigma_{\epsilon_1}, \sigma_{\epsilon_2}, \sigma_{\epsilon_3}, \sigma_{\epsilon_4}, \sigma_{\epsilon_5}, \sigma_{\epsilon_6}, \sigma_{\epsilon_7}, \sigma_{\epsilon_8}\}$$

E, de maneira similar, no caso da estimação alternativa (sem a série histórica do VIX à vista), temos o vetor  $\Theta$  (agora com 18 parâmetros):

$$\Theta = \{\kappa_1, \kappa_2, \sigma_1, \sigma_2, p_1, p_2, q_1, q_2, \rho, \mu_1, \mu_2, \sigma_{\epsilon_1}, \sigma_{\epsilon_2}, \sigma_{\epsilon_3}, \sigma_{\epsilon_4}, \sigma_{\epsilon_5}, \sigma_{\epsilon_6}, \sigma_{\epsilon_7}\}$$

Portanto, o processo de estimação do modelo resume-se agora ao seguinte problema de otimização: encontrar o vetor de hiperparâmetros  $\Theta$  que minimize a função de verossimilhança dada por B.3.9 a partir das seguintes restrições:

$$\min_{\Theta} \ln(L[y|\Theta]) \quad s.a. \quad \begin{cases} 0 < \kappa_i < \infty & i = \{1, 2\} \\ 0 < \sigma_i < \infty & i = \{1, 2\} \\ -\infty < p_i < \infty & i = \{1, 2\} \\ -\infty < q_i < \infty & i = \{1, 2\} \\ -1 \leq \rho \leq 1 \\ -\infty < \mu_i < \infty & i = \{1, 2\} \\ 0 < \sigma_{\epsilon_i} < \infty & i = \{1, 2, 3, 4, 5, 6, 7, 8\} \end{cases} \quad (3.2.1)$$

Dessa forma, podemos agora passar para uma análise mais detalhada dos resultados obtidos nessas otimizações.

### 3.3 Análise dos resultados

Como realizamos dois procedimentos de calibragem (um considerando os log-preços do VIX à vista dentro dos dados da amostra e o outro sem), vamos fazer duas análises separadas e, ao final, realizar um comparativo entre esses dois resultados. Sendo assim, vamos começar com os valores obtidos no caso em que o VIX à vista é considerado.

Na tabela 3.2 temos os resultados da otimização. Vale notar que o parâmetro de reversão à média  $\kappa_1$  associado ao primeiro fator é bem menor que  $\kappa_2$ , associado ao segundo fator. Acredito que este aspecto está associado à característica dos fatores que os mesmos representam. Por exemplo, como o fator  $X_2$ , pela análise de componentes principais feita em Avellaneda and Papanicolaou [2018] representa mudanças de inclinação na estrutura a termo, uma vez que a curva muda o seu comportamento de *contango* para *backwardation*, isso pode significar que o parâmetro  $\kappa_2$  irá exercer uma pressão maior para que a estrutura a termo volte para o seu nível de longo prazo, coerente com o fato de que, em períodos de crise, apesar de a curva alcançar níveis muito altos, ela não permanece durante muito tempo nesses níveis e tende a retornar de maneira mais rápida para a sua média. Além disso, no sentido de que as mudanças de inclinação acarretam maiores variações nos níveis do índice, a volatilidade do fator  $X_2$  também apresenta-se maior do que aquela associada a do fator  $X_1$ . Podemos ver também que as médias dos processos ( $\mu_1$  e  $\mu_2$ ) ficaram muito próximas. Por último, analisando o *p-valor* do teste-*t* bicaudal, para um intervalo de confiança de 95%, podemos verificar que, com exceção dos parâmetros ( $p_2$  e  $q_1$ ) — teste inconclusivo — todos os demais parâmetros do modelo apresentam

significância estatística, rejeitando a hipótese de nulidade para os mesmos.

| Parâmetro               | Valor Estimado | Desvio-padrão           | Estatística-t | P-Valor (0,025)                           |
|-------------------------|----------------|-------------------------|---------------|---|
| $\kappa_1$              | 0,3592         | 0,1301                  | 2,7615        | $5,9620 \times 10^{-3}$                   |
| $\kappa_2$              | 2,7256         | 0,1319                  | 20,6628       | $2,4757 \times 10^{-69}$                  |
| $\sigma_1$              | 0,2809         | 0,0537                  | 5,2323        | $2,4507 \times 10^{-7}$                   |
| $\sigma_2$              | 0,3787         | 0,0497                  | 7,6228        | $1,2246 \times 10^{-13}$                  |
| $p_1$                   | 3,1235         | 1,2389                  | 2,5212        | $1,2001 \times 10^{-2}$                   |
| $p_2$                   | -0,4046        | 1,1577                  | -0,3495       | <b><math>7,2683 \times 10^{-1}</math></b> |
| $q_1$                   | -1,6686        | 0,8820                  | -1,8918       | <b><math>5,9087 \times 10^{-2}</math></b> |
| $q_2$                   | -1,8117        | 0,3513                  | -5,1570       | $3,6004 \times 10^{-7}$                   |
| $\rho$                  | -0,5784        | 0,0887                  | -6,5235       | $1,6607 \times 10^{-10}$                  |
| $\mu_1$                 | 1,3812         | 0,0804                  | 17,1900       | $1,4729 \times 10^{-52}$                  |
| $\mu_2$                 | 1,3659         | 0,3543                  | 3,8552        | $1,3042 \times 10^{-4}$                   |
| $\sigma_{\epsilon_1}^2$ | 0,0158         | $4,5285 \times 10^{-5}$ | 347,8044      | 0,0000                                    |
| $\sigma_{\epsilon_2}^2$ | 0,0064         | $2,9632 \times 10^{-5}$ | 217,0337      | 0,0000                                    |
| $\sigma_{\epsilon_3}^2$ | 0,0008         | $4,3376 \times 10^{-6}$ | 188,6507      | 0,0000                                    |
| $\sigma_{\epsilon_4}^2$ | 0,0002         | $6,5260 \times 10^{-7}$ | 353,7492      | 0,0000                                    |
| $\sigma_{\epsilon_5}^2$ | 0,0001         | $3,6067 \times 10^{-7}$ | 290,9341      | 0,0000                                    |
| $\sigma_{\epsilon_6}^2$ | 0,0001         | $2,8864 \times 10^{-7}$ | 439,9852      | 0,0000                                    |
| $\sigma_{\epsilon_7}^2$ | 0,0001         | $2,6741 \times 10^{-7}$ | 345,1170      | 0,0000                                    |
| $\sigma_{\epsilon_8}^2$ | 0,0002         | $5,2819 \times 10^{-7}$ | 308,4680      | 0,0000                                    |

Tabela 3.2: Resultado da estimação original (amostra com o VIX à vista)

A figura 3.3 mostra as duas variáveis de estado  $X_1$  e  $X_2$  estimadas. Nela, podemos notar que o fator  $X_2$ , associado aos movimentos de inclinação da estrutura a termo é visivelmente mais volátil que o fator  $X_1$  associado ao nível da curva. Este gráfico corrobora com o resultado das estimações anteriormente colocadas.

Os parâmetros  $\lambda_1$  e  $\lambda_2$  (preços de risco de mercado) são obtidos pelas relações apresentadas em 2.2.1 e a figura 3.4 mostra a série temporal estimada para os mesmos. Aqui, vale notar que uma vez que a premissa é de que esse MPR é dependente do fator  $X_i$ , como o segundo fator é mais volátil que o primeiro,  $\lambda_2$  é mais volátil que  $\lambda_1$ .



Figura 3.3: Estimativas para as variáveis de estado  $X_1$  e  $X_2$  (amostra com o VIX à vista)



Figura 3.4: Estimativas para os preços de risco de mercado  $\lambda_1$  e  $\lambda_2$  (amostra com o VIX à vista)

Os erros expressos pelos RMSE, MAE e MAPE são apresentados na tabela 3.3. O erro percentual médio das observações suavizadas é de 2,51%. O maior erro observado é de 8,83%, referente ao VIX à vista. O contrato com menor erro é o CMF 180d, com 0,60%.

| Contrato    | RMSE   | MAE    | MAPE   |
|-------------|--------|--------|--------|
| VIX à vista | 3,3827 | 1,9284 | 0,0883 |
| CMF 30d     | 2,0436 | 1,2336 | 0,0557 |
| CMF 60d     | 0,7414 | 0,4296 | 0,0190 |
| CMF 90d     | 0,3576 | 0,2285 | 0,0098 |
| CMF 120d    | 0,2215 | 0,1445 | 0,0062 |
| CMF 150d    | 0,2458 | 0,1582 | 0,0070 |
| CMF 180d    | 0,2074 | 0,1386 | 0,0060 |
| CMF 210d    | 0,3192 | 0,2189 | 0,0091 |
| Média       | 0,9399 | 0,5600 | 0,0251 |

Tabela 3.3: Análise dos erros (amostra com o VIX à vista)

A figura 3.5 mostra o comparativo do VIX à vista observado com a respectiva estimativa calculada a partir dos parâmetros obtidos e a série suavizada. Podemos observar que o modelo se mostra aderente às observações.

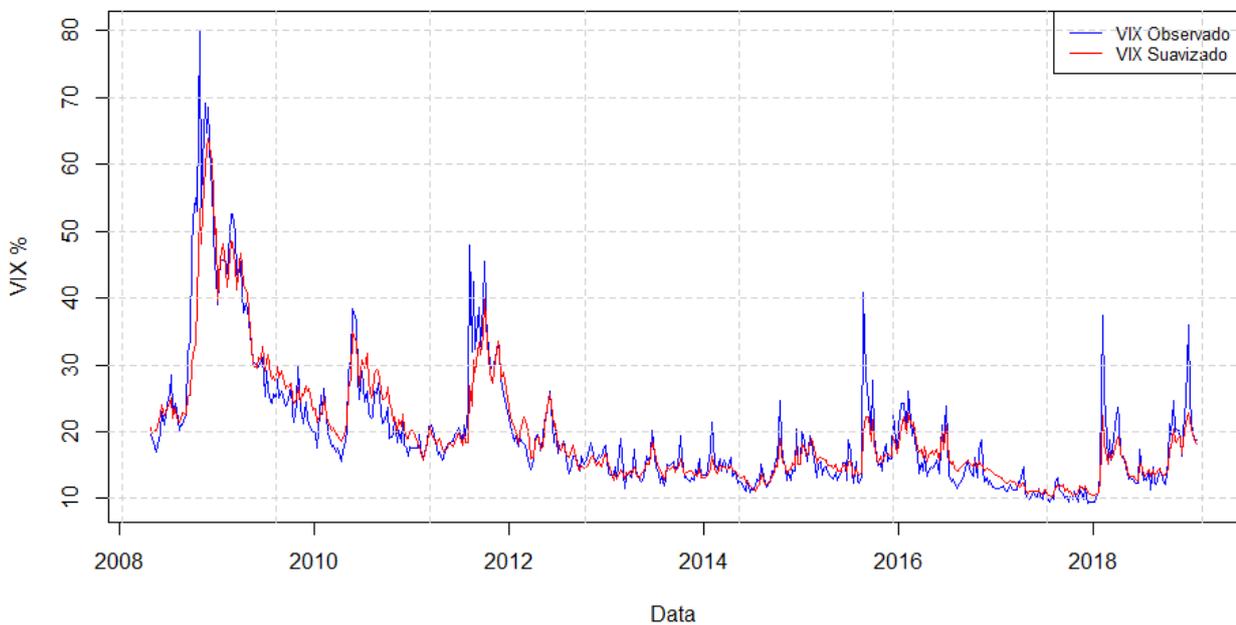


Figura 3.5: Comparativo VIX à Vista: Observado vs. Suavizado (amostra com o VIX à vista)

Por último, nesta primeira parte, podemos ver na figura 3.6 a comparação entre a estrutura a termo observada e a suavizada, em dias escolhidos da amostra. É possível ver que, a mesma apresenta-se bem coerente tanto em termos de dinâmica quanto em termos dos resultados obtidos na mensuração dos erros (na qual os vencimentos mais próximos apresentam maiores variações). Os demais gráficos comparativos, são apresentados no apêndice C.

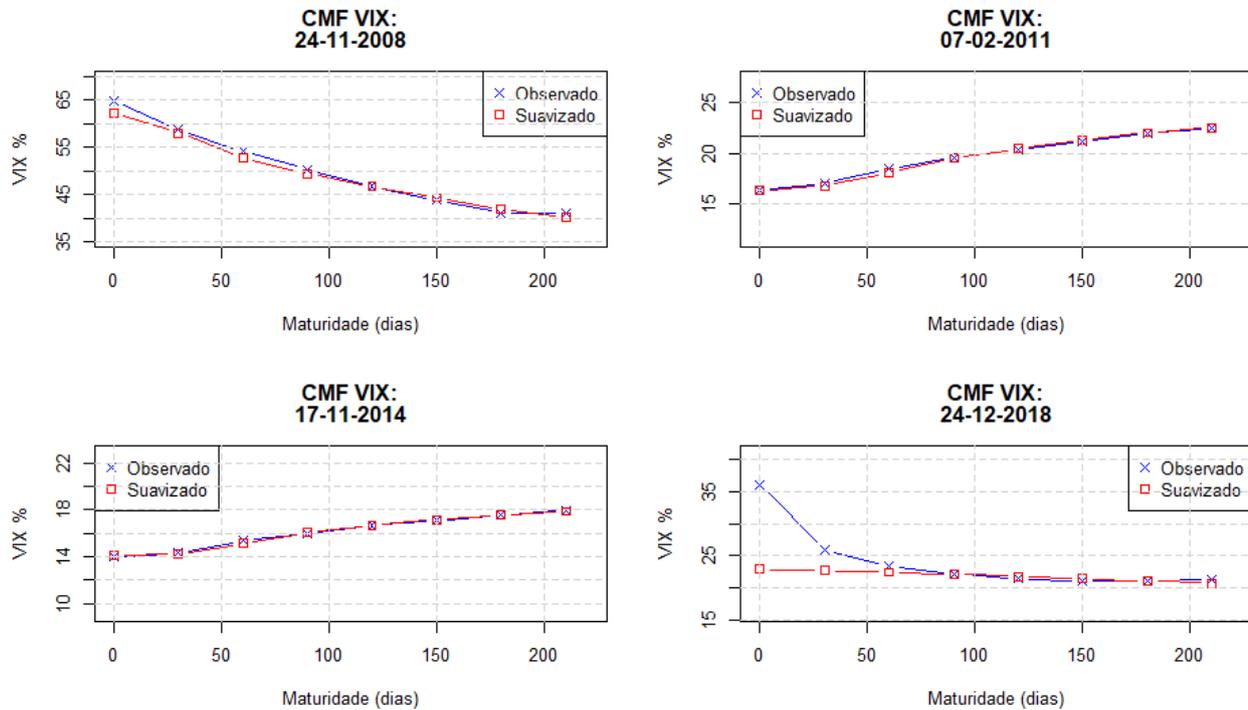


Figura 3.6: Comparativo da Estrutura a Termo (dias escolhidos da amostra com o VIX à vista)

Podemos passar agora à análise do caso em que o log-preço do VIX à vista não faz parte da amostra de dados para a calibragem. Basicamente, podemos dizer que nenhuma mudança muito expressiva em relação aos valores das estimativas dos parâmetros ocorreu. Contudo, ao olharmos para a significância dos parâmetros a partir do *teste-t*, vemos que, agora, rejeitamos a hipótese de nulidade para todos os parâmetros, com exceção de  $p_2$ . Os resultados da otimização são apresentados na tabela 3.4. Por conta dessa semelhança de valores, apresentamos os gráficos associados à essa análise no Apêndice C, mais especificamente as figuras C.7, C.8, C.9, C.10 e C.12.

Neste caso, o aspecto mais interessante está na medição dos erros e na hipótese de que, sem o VIX à vista na amostra utilizada para o procedimento de calibragem, as aproximações poderiam apresentar erros menores. Na tabela 3.5, podemos ver que o erro percentual médio das observações filtradas ficou em 1,71% (abaixo do que foi obtido anteriormente, como era esperado). O maior erro observado é de 6,41%, neste caso, referente ao CMF 30d. O contrato com menor erro permaneceu o CMF 180d, com 0,54%.

No geral, podemos observar que sem a utilização do log-preço do VIX à vista, os contratos com maturidade constante acima de 60 dias (90d, 120d, 150d, 180d e 210d) apresentaram uma redução em seus erros percentuais. Contudo, nos contratos mais curtos (30d e 60d) os erros

aumentaram (principalmente no caso do CMF 30d), mas sem anular os ganhos obtidos nas reduções dos erros dos contratos de maior prazo. Isso nos leva a pensar que, dependendo da maturidade do contrato que se quer aproximar, a escolha de quais vencimentos serão incluídos na amostra para calibragem deve ser um aspecto analisado com atenção.

Levantamos essa questão, pois como vimos na introdução deste trabalho, existem alguns ETNs e ETFs que buscam replicar a dinâmica mais longa da curva e, nessas aplicações, poderia ser mais interessante a inclusão de determinados contratos em detrimento de outros. Contudo, esta discussão não fará parte do objetivo deste trabalho, ficando ao leitor uma sugestão de tema de pesquisa futura.

| Parâmetro               | Valor Estimado | Desvio-Padrão           | Estatística-t | P-Valor (0,025)           |
|-------------------------|----------------|-------------------------|---------------|---------------------------|
| $\kappa_1$              | 0,5872         | 0,0724                  | 8,1107        | $3,7881 \times 10^{-15}$  |
| $\kappa_2$              | 2,8123         | 0,1049                  | 26,7990       | $2,4891 \times 10^{-99}$  |
| $\sigma_1$              | 0,3779         | 0,0346                  | 10,9088       | $4,8392 \times 10^{-25}$  |
| $\sigma_2$              | 0,4005         | 0,0132                  | 30,2562       | $7,9680 \times 10^{-116}$ |
| $p_1$                   | 2,5427         | 0,2747                  | 9,2558        | $5,8982 \times 10^{-19}$  |
| $p_2$                   | -0,1098        | 0,1934                  | -0,5678       | $5,7047 \times 10^{-1}$   |
| $q_1$                   | -1,8320        | 0,0896                  | -20,4465      | $2,8257 \times 10^{-68}$  |
| $q_2$                   | -1,2161        | 0,1127                  | -10,7921      | $1,3640 \times 10^{-24}$  |
| $\rho$                  | -0,4997        | 0,0238                  | -20,9997      | $5,5524 \times 10^{-71}$  |
| $\mu_1$                 | 1,3539         | 0,0436                  | 31,0433       | $1,6427 \times 10^{-119}$ |
| $\mu_2$                 | 1,2184         | 0,0819                  | 14,8798       | $7,7209 \times 10^{-42}$  |
| $\sigma_{\epsilon_1}^2$ | 0,0070         | $1,3231 \times 10^{-5}$ | 525,9351      | 0,0000                    |
| $\sigma_{\epsilon_2}^2$ | 0,0010         | $2,8059 \times 10^{-6}$ | 353,4038      | 0,0000                    |
| $\sigma_{\epsilon_3}^2$ | 0,0002         | $6,7669 \times 10^{-7}$ | 354,7496      | 0,0000                    |
| $\sigma_{\epsilon_4}^2$ | 0,0001         | $2,8533 \times 10^{-7}$ | 323,1370      | 0,0000                    |
| $\sigma_{\epsilon_5}^2$ | 0,0001         | $2,1763 \times 10^{-7}$ | 480,4263      | 0,0000                    |
| $\sigma_{\epsilon_6}^2$ | 0,0001         | $2,2251 \times 10^{-7}$ | 357,7685      | 0,0000                    |
| $\sigma_{\epsilon_7}^2$ | 0,0001         | $2,9629 \times 10^{-7}$ | 365,9486      | 0,0000                    |

Tabela 3.4: Resultado da estimação alternativa (amostra sem o VIX à vista)

| Contrato | RMSE   | MAE    | MAPE   |
|----------|--------|--------|--------|
| CMF 30d  | 2,1477 | 1,3715 | 0,0641 |
| CMF 60d  | 0,7679 | 0,4634 | 0,0207 |
| CMF 90d  | 0,3595 | 0,2216 | 0,0095 |
| CMF 120d | 0,2099 | 0,1404 | 0,0059 |
| CMF 150d | 0,2277 | 0,1543 | 0,0067 |
| CMF 180d | 0,1901 | 0,1266 | 0,0054 |
| CMF 210d | 0,2570 | 0,1736 | 0,0071 |
| Média    | 0,5942 | 0,3788 | 0,0171 |

Tabela 3.5: Análise dos erros (amostra sem o VIX à vista)

Agora, de posse dos parâmetros, podemos abordar alguns aspectos práticos que o modelo permite aplicar aos produtos existentes no mercado.

# Capítulo 4

## Aplicações aos ETPs

Nesta seção, vamos abordar com maior ênfase as estratégias que os ETPs realizam para replicar a rentabilidade dos contratos futuros de maturidade constante do VIX e derivar os respectivos índices de retorno, de modo a simular cenários para os produtos, a partir das estimativas dos parâmetros obtidos na seção anterior.

A estratégia mais comum na negociação (rolagem) de contratos futuros é quando as posições são transferidas de um vencimento para outro à medida que os contratos expiram. A estratégia mais simples é investir em uma posição comprada no primeiro contrato, reinvestir todos os lucros/perdas (dos ajustes) e, na data de vencimento do contrato, fechar a posição e, simultaneamente, abrir uma posição comprada pelo mesmo valor no contrato subsequente.

Vamos tomar como base a seguinte situação (que, como vimos no capítulo 1, é a estratégia mais adotada atualmente pelo mercado): um fundo de investimentos compra um número de contratos futuros com valor nocional igual ao respectivo patrimônio. Rebalanceia diariamente o valor investido de acordo com os ajustes (de lucro/prejuízo) dos contratos futuros e transfere a posição inteira para o próximo contrato (vencimento) na data de expiração da posição (contrato) atual. Além disso, assume-se que o caixa do fundo é remunerado à taxa  $r$  e desconsideramos os custos de chamada de margem dos contratos futuros e outras possíveis despesas. Vamos adotar a seguinte notação:

$F_{t,T_i}$  - Contrato futuro com maturidade em  $T_i$ , avaliado na data corrente (atual)  $t$ ;

$T_1$  - Data de expiração do primeiro contrato futuro;

$T_2$  - Data de expiração do contrato futuro subsequente à  $T_1$ ;

$r$  - Taxa de juros;

$dX_t = X_{t+dt} - X_{dt}$  - Avanço do processo  $X_t$ , mais especificamente, um dia de modificação da data  $t$  para a data  $t + dt$ ;

$I_t$  - Retorno Total do Índice (TRI) da estratégia.

Podemos então colocar que, sob essas condições, o valor total do fundo, no instante  $t$ , satisfaz:

$$\frac{dI_t}{I_t} = \begin{cases} \frac{dF_{t,T_1}}{F_{t,T_1}} + rdt & , se \quad t < T_2 - dt \\ \frac{dF_{t,T_2}}{F_{t,T_2}} + rdt & , se \quad t \geq T_2 - dt \end{cases} \quad (4.0.1)$$

Vamos ainda definir:

$$\phi(t) = \begin{cases} F_{t,T_1} & , se \quad t < T_2 - dt \\ F_{t,T_2} & , se \quad t = T_2 - dt \end{cases} \quad (4.0.2)$$

para representar o preço de liquidação do primeiro contrato no instante  $t$ . Assim:

$$\phi(t) = b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2}$$

na qual:

$$b(t) = \begin{cases} 1 & , se \quad t < T_2 \\ 0 & , se \quad t = T_2 \end{cases}$$

Vale lembrar que a medida de Dirac concentrada em  $x$  é:

$$\mu(S) = \begin{cases} 1 & , se \quad x \in S \\ 0 & , se \quad x \notin S \end{cases}$$

Portanto, a partir dessas definições, vamos analisar qual é a dinâmica de evolução de  $I_t$  (TRI), de acordo com as hipóteses assumidas. Primeiro, podemos ver que (aplicando a regra do produto):

$$\begin{aligned} \phi(t) &= b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2} \\ d\phi(t) &= b(t)dF_{t,T_1} + F_{t,T_1} \frac{\partial b(t)}{\partial t} dt + (1 - b(t))dF_{t,T_2} - F_{t,T_2} \frac{\partial b(t)}{\partial t} dt \\ d\phi(t) &= b(t)dF_{t,T_1} + (1 - b(t))dF_{t,T_2} + (F_{t,T_1} - F_{t,T_2}) \frac{\partial b(t)}{\partial t} dt \end{aligned}$$

Mas, o último termo só será diferente de zero quando  $t < T_2$ . Portanto, podemos reescrever como:

$$d\phi(t) = b(t)dF_{t,T_1} + (1 - b(t))dF_{t,T_2} + (F_{t,T_1} - F_{t,T_2})\delta(t - T_2)dt \quad (4.0.3)$$

Na qual  $\delta(\cdot)$  representa a medida de Dirac concentrada em  $x$ .

A partir desse resultado, podemos ainda desenvolver:

$$\begin{aligned} \frac{d\phi(t)}{\phi(t)} &= \frac{b(t)dF_{t,T_1}}{\phi(t)} + \frac{(1 - b(t))dF_{t,T_2}}{\phi(t)} + \frac{(F_{t,T_1} - F_{t,T_2})\delta(t - T_2)dt}{\phi(t)} \\ \frac{d\phi(t)}{\phi(t)} &= \frac{b(t)dF_{t,T_1}}{b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2}} + \frac{(1 - b(t))dF_{t,T_2}}{b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2}} + \frac{(F_{t,T_1} - F_{t,T_2})\delta(t - T_2)dt}{b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2}} \end{aligned}$$

De modo que, avaliando os casos da função, vamos ficar com:

Se  $t < T_2 \Rightarrow b(t) = 1$ , portanto:

$$\frac{d\phi(t)}{\phi(t)} = \frac{dF_{t,T_1}}{F_{t,T_1}} + \frac{(F_{t,T_1} - F_{t,T_2})\delta(t - T_2)dt}{F_{t,T_1}}$$

Se  $t = T_2 \Rightarrow b(t) = 0$ , portanto:

$$\frac{d\phi(t)}{\phi(t)} = \frac{dF_{t,T_2}}{F_{t,T_2}}$$

Mas, pela definição de  $\frac{dI_t}{I_t}$ , podemos reescrever:

$$\frac{d\phi(t)}{\phi(t)} = \frac{dI_t}{I_t} - rdt + \frac{(F_{t,T_1} - F_{t,T_2})}{\phi(t)}\delta(t - T_2)dt$$

ou, ainda:

$$\frac{dI_t}{I_t} = \frac{d\phi(t)}{\phi(t)} - \frac{(F_{t,T_1} - F_{t,T_2})}{\phi(t)}\delta(t - T_2)dt + rdt \quad (4.0.4)$$

Como sendo a dinâmica da evolução de TRI, a partir das hipóteses assumidas. Podemos ver que, supondo que o preço do contrato do primeiro vencimento siga uma distribuição estacionária, então heurísticamente, se a curva estiver em *contango*, rolar uma posição comprada fará com que o *trader* perca dinheiro (na média). Assim, de maneira inversa, ganhará dinheiro quando a curva estiver em *backwardation*.

Aqui, verifica-se ainda que, considerações semelhantes podem ser feitas a outras estratégias de rolagem, como uma estratégia que rola essas posições nos contratos no meio do mês, ou seja, quando  $\theta + t = \frac{T_1 + T_2}{2}$ .

Vamos agora estender a análise para o caso de rolagem de posições em tempo contínuo. Considerando uma função decrescente (ainda a ser determinada)  $b(t)$ , tal que:

$$b(t) = \begin{cases} b(T_1^+) = 1 \\ b(T_2) = 0 \end{cases} \quad (4.0.5)$$

Onde  $b(t)$  representa a fração do número total de contratos investidos no primeiro vencimento e  $(1 - b(t))$  é a fração investida no contrato referente ao segundo vencimento. O prazo médio (ponderado por contrato)  $\theta(t)$  será dado por:

$$\theta(t) = b(t)(T_1 - t) + (1 - b(t))(T_2 - t) \quad (4.0.6)$$

Tomando  $\theta(t) = \theta$  (constante), teremos:

$$\begin{aligned} \theta(t) &= b(t)(T_1 - t) + (1 - b(t))(T_2 - t) \\ \theta &= b(t)T_1 - \cancel{b(t)t} + T_2 - t - b(t)T_2 + \cancel{b(t)t} \\ \theta - T_2 + t &= b(t)(T_1 - T_2) \end{aligned}$$

Portanto:

$$b(t) = \frac{T_2 - t - \theta}{(T_2 - T_1)} \quad e \quad \frac{\partial b(t)}{\partial t} = -\frac{1}{(T_2 - T_1)}$$

Pela definição da CMF dada anteriormente, temos que:

$$\phi(t) = b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2} = V_t^\theta \quad (4.0.7)$$

de modo que podemos calcular:

$$\begin{aligned}
dV_t^\theta &= b(t)dF_{t,T_1} + F_{t,T_1} \frac{\partial b(t)}{\partial t} dt + (1 - b(t))dF_{t,T_2} - F_{t,T_2} \frac{\partial b(t)}{\partial t} dt \\
&= b(t)dF_{t,T_1} - \frac{F_{t,T_1}}{(T_2 - T_1)} dt + (1 - b(t))dF_{t,T_2} + \frac{F_{t,T_2}}{(T_2 - T_1)} dt \\
&= b(t)dF_{t,T_1} + (1 - b(t))dF_{t,T_2} + (F_{t,T_2} - F_{t,T_1}) \frac{dt}{(T_2 - T_1)}
\end{aligned}$$

agora, se tomarmos:

$$w_1(t) = \frac{b(t)F_{t,T_1}}{V_t^\theta} \quad e \quad w_2(t) = \frac{(1 - b(t))F_{t,T_2}}{V_t^\theta}$$

nós podemos escrever:

$$\frac{dV_t^\theta}{V_t^\theta} = w_1(t) \frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t) \frac{dF_{t,T_2}}{F_{t,T_2}} + \frac{(F_{t,T_2} - F_{t,T_1})}{V_t^\theta} \frac{dt}{(T_2 - T_1)}$$

mas, como temos:

$$\begin{aligned}
V_t^\theta &= \left( \frac{T_2 - t - \theta}{(T_2 - T_1)} \right) F_{t,T_1} + \left[ 1 - \left( \frac{T_2 - t - \theta}{(T_2 - T_1)} \right) \right] F_{t,T_2} \\
\ln(V_t^\theta) &= \ln \left[ \left( \frac{T_2 - t - \theta}{(T_2 - T_1)} \right) F_{t,T_1} \right] + \ln \left\{ \left[ 1 - \left( \frac{T_2 - t - \theta}{(T_2 - T_1)} \right) \right] F_{t,T_2} \right\} \\
\frac{\partial \ln(V_t^\theta)}{\partial \theta} &= \frac{1}{V_t^\theta} \left[ - \frac{F_{t,T_1}}{(T_2 - T_1)} + \frac{F_{t,T_2}}{(T_2 - T_1)} \right] dt \\
\frac{\partial \ln(V_t^\theta)}{\partial \theta} &= \frac{(F_{t,T_2} - F_{t,T_1})}{V_t^\theta} \frac{dt}{(T_2 - T_1)}
\end{aligned}$$

podemos então reescrever:

$$\frac{dV_t^\theta}{V_t^\theta} = w_1(t) \frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t) \frac{dF_{t,T_2}}{F_{t,T_2}} + \left[ \frac{\partial \ln(V_t^\tau)}{\partial \tau} \right]_{\tau=\theta} dt$$

A partir disso, vemos que o índice de retorno total (TRI) de uma estratégia de negociação (rolagem) contínua de compra de contratos futuros com maturidade constante média igual à  $\theta$ , satisfaz:

$$\begin{aligned}
\frac{dI_t^\theta}{I_t^\theta} &= w_1(t) \frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t) \frac{dF_{t,T_2}}{F_{t,T_2}} + r dt \\
&= \frac{dV_t^\theta}{V_t^\theta} - \left[ \frac{\partial \ln(V_t^\tau)}{\partial \tau} \right]_{\tau=\theta} dt + r dt
\end{aligned} \tag{4.0.8}$$

Então, o retorno diário do índice é a soma da variação percentual diária do CMF (com maturidade  $\theta$ ) e um termo de *drift* que vai depender da inclinação da curva CMF em  $\tau = \theta$ . Este termo de *drift* será negativo se a curva estiver em *contango*, e positivo, se a curva estiver em

*backwardation* em  $\tau = \theta$ . Os pesos  $w_1(t)$  e  $w_2(t)$  correspondem às proporções de ativos (valores em dólares) investidos em cada um dos contratos, a fim de manter um prazo de vencimento constantes igual a  $\theta$ . Para uma discussão mais aprofundada dessa derivação, recomendamos Bergomi [2015].

Uma vez que esse TRI corresponde à estratégia de uma posição comprada, de maneira análoga, podemos escrevê-lo para a estratégia de uma posição vendida:

$$\begin{aligned} \frac{dJ_t^\theta}{J_t^\theta} &= -w_1(t) \frac{dF_{t,T_1}}{F_{t,T_1}} - w_2(t) \frac{dF_{t,T_2}}{F_{t,T_2}} + rdt \\ &= -\frac{dV_t^\theta}{V_t^\theta} + \left[ \frac{\partial \ln(V_t^\tau)}{\partial \tau} \right]_{\tau=\theta} dt + rdt \end{aligned} \quad (4.0.9)$$

Em Avellaneda and Papanicolaou [2018] os autores desenvolvem também as derivações de outras estratégias, que utilizam mais de 2 contratos futuros para replicar maturidades constantes diferentes de 30 dias. Além disso, propõem uma análise na qual concluem que, sob certas suposições (de estacionariedade e ergodicidade), vender um ETP sobre o VIX, que tem como estratégia assumir e rolar uma posição comprada no primeiro vencimento, ou comprar um ETP sobre o VIX, que tem como estratégia assumir e rolar uma posição vendida no primeiro vencimento, ambos produzem lucros certos quanto  $t \rightarrow \infty$ .

Contudo, complementam que esses lucros teóricos não são isentos de risco. Pois, as equações também nos mostram que, se  $V_t^\tau$  for muito volátil, a variável aleatória  $\frac{V_t^\tau}{V_t^\tau}$  pode ser muito pequena, ainda que  $t$  seja grande. Além disso, o estado de *backwardation* associado a uma grande volatilidade pode fazer com que o valor de  $\frac{\partial \ln(V_t^\tau)}{\partial \tau}$  seja negativo. Dessa forma, o retorno associado às estratégias descritas anteriormente, pode ser negativo em qualquer intervalo de tempo  $[0, t]$ .

Voltando agora ao cálculo do lucro/perda teórica associado à estratégia de rolagem, podemos ver que, a partir do resultado encontrado em 4.0.8, e sabendo que, pelo modelo proposto em 2.2.9:  $V_t^T = V^\infty \exp\left(\sum_{i=1}^d e^{-\bar{\kappa}_i(T-t)} Y_{it} - \frac{1}{2} \sum_{ij=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j)(T-t)}\right)$  ou seja,  $V_t^T = f(Y_t, t)$ . Para calcular  $\frac{dV^\tau}{V^\tau}$  usamos Itô no caso multidimensional (conforme B.1.9):

Aqui, vale destacar que, como a estratégia possui uma maturidade constante  $\tau$ , temos  $V^\tau = f(Y_\tau) = V^\tau = V^\infty \exp\left(\sum_{i=1}^d e^{-\bar{\kappa}_i \tau} Y_{i\tau} - \frac{1}{2} \sum_{ij=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau}\right)$ . Ou seja, a maturidade da estratégia não varia. De modo que:

$$\begin{aligned} \frac{\partial f(Y_\tau)}{\partial t} &= 0 \\ \frac{\partial f(Y_\tau)}{\partial y_i} &= V_t e^{-\bar{\kappa}_i(T-t)} \\ \frac{\partial^2 f(Y_\tau)}{\partial y_i \partial y_j} &= V_t e^{-(\bar{\kappa}_i + \bar{\kappa}_j)(T-t)} \end{aligned}$$

E, ainda:

$$\langle dY_{i\tau}, dY_{j\tau} \rangle = \sigma_i \sigma_j \rho_{ij} dt$$

Assim, para uma maturidade média, igual a  $\tau$ , temos:

$$\begin{aligned} dV^\tau &= \left[ \sum_{i=1}^d V^\tau e^{-\bar{\kappa}_i \tau} \right] dY_{i\tau} + \left[ \frac{1}{2} \sum_{i,j=1}^d V^\tau e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \right] \langle dY_{i\tau}, dY_{j\tau} \rangle \\ \frac{dV^\tau}{V^\tau} &= \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} dY_{i\tau} + \frac{1}{2} \sum_{i,j=1}^d e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \sigma_i \sigma_j \rho_{ij} dt \end{aligned} \quad (4.0.10)$$

Agora, calculando  $\frac{\partial \ln V^\tau}{\partial \tau} dt$ , para uma maturidade média igual a  $\tau$ , temos::

$$\begin{aligned} \ln V^\tau &= \ln V^\infty + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} Y_{i\tau} - \frac{1}{2} \sum_{i,j=1}^d \frac{\sigma_i \sigma_j \rho_{ij}}{\bar{\kappa}_i + \bar{\kappa}_j} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \\ \frac{\partial \ln V^\tau}{\partial \tau} &= - \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} Y_{i\tau} \bar{\kappa}_i + \frac{1}{2} \sum_{i,j=1}^d \sigma_i \sigma_j \rho_{ij} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \end{aligned} \quad (4.0.11)$$

Substituindo os resultados encontrados em 4.0.10 e 4.0.11, ficamos com:

$$\begin{aligned} \frac{dI^\tau}{I^\tau} &= rdt + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} dY_{i\tau} + \frac{1}{2} \sum_{i,j=1}^d e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \sigma_i \sigma_j \rho_{ij} dt - \left[ - \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} Y_{i\tau} \bar{\kappa}_i + \frac{1}{2} \sum_{i,j=1}^d \sigma_i \sigma_j \rho_{ij} e^{-(\bar{\kappa}_i + \bar{\kappa}_j) \tau} \right] dt \\ &= rdt + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (dY_{i\tau} + Y_{i\tau} \bar{\kappa}_i dt) \end{aligned}$$

Mas podemos ainda somar e subtrair o termo  $\kappa_i Y_i dt$ , para obter:

$$\frac{dI^\tau}{I^\tau} = rdt + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (dY_{i\tau} + \kappa_i Y_i dt + (\bar{\kappa}_i - \kappa_i) Y_i dt)$$

E substituindo  $dY_{i\tau}$  (conforme definido em 2.2.8):

$$\frac{dI^\tau}{I^\tau} = rdt + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} \sigma_i dW_i^{\mathbb{P}} + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (\kappa_i \mu_i^Y + (\bar{\kappa}_i - \kappa_i) Y_i) dt$$

Somamos e subtraímos  $\bar{\kappa}_i \mu_i^Y dt$ , para reorganizar os termos e ficar com a mesma forma proposta para o TRI em Avellaneda and Papanicolaou [2018]:

$$\frac{dI^\tau}{I^\tau} = rdt + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} \sigma_i dW_i^{\mathbb{P}} + \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (\bar{\kappa}_i \mu_i^Y + (\bar{\kappa}_i - \kappa_i) (Y_i - \mu_i^Y)) dt \quad (4.0.12)$$

De modo que agora somos capazes de implementar essa E.D.E e, com os parâmetros obtidos na calibração do modelo, simular a rentabilidade esperada para as estratégias descritas nessa seção.

Diante disso, a partir da última data da base de dados utilizada para a estimação dos parâmetros (28/01/2019), simulamos 10.000 (dez mil) cenários de retornos diários (TRI) para um ano à frente (assumimos 252 dias úteis). Ou seja, de maneira mais prática nesse exemplo, simulamos as possíveis rentabilidades do ETN VXX, de modo que foi possível fazer a composição diária, com o intuito de obter um índice acumulado para o respectivo período e compará-lo com os dados realizados.

Como neste período ainda não temos os dados realizados até 28/01/2020 (1 ano a partir da última data da base), o período de dados realizados ficou menor que o simulado. Contudo, podemos observar na figura 4.1 que o Índice de Retorno Total Acumulado no período realizado está coerente com as simulações feitas e, ainda na figura 4.2, vemos que a média das simulações possui a característica esperada de um *drift* negativo por conta do estado de *contango* ser o mais comum e, conseqüentemente, gerar uma perda de rentabilidade na rolagem de posição do primeiro para o segundo contrato com vencimento mais próximo.

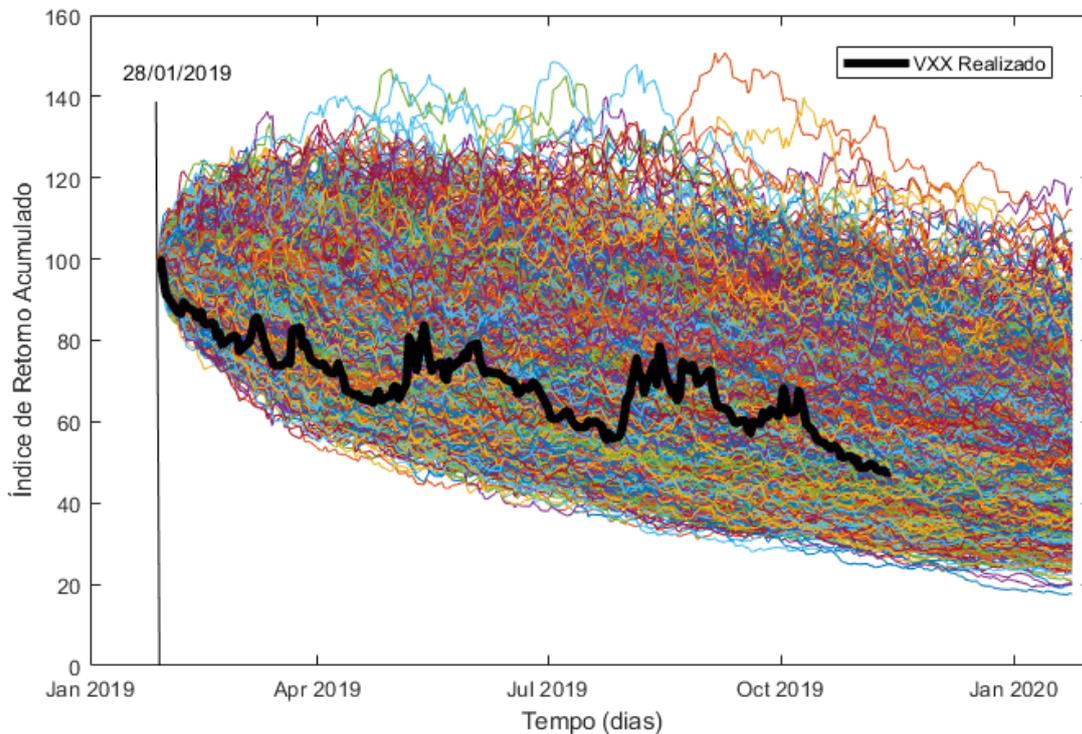


Figura 4.1: Comparativo entre as Simulações e o Realizado do ETN VXX

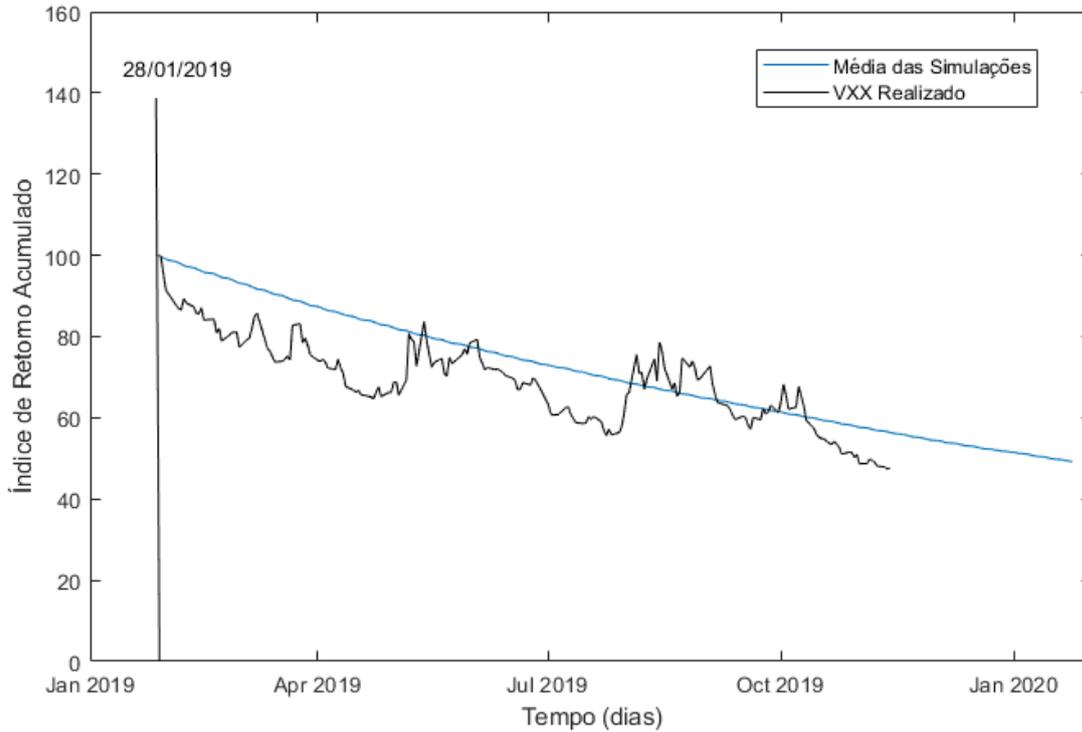


Figura 4.2: Comparativo entre a média das Simulações e o Realizado do ETN VXX

De maneira análoga ao que colocamos anteriormente, a rentabilidade da estratégia de assumir uma posição vendida nos 2 primeiros contratos futuros foi expressa por 4.0.9, de modo que, para uma maturidade média igual a  $\tau$ , seu TRI será dado por:

$$\frac{dJ^\tau}{J^\tau} = rdt - \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} \sigma_i dW_i^{\mathbb{P}} - \sum_{i=1}^d e^{-\bar{\kappa}_i \tau} (\bar{\kappa}_i \mu_i^Y + (\bar{\kappa}_i - \kappa_i)(Y_i - \mu_i^Y)) dt \quad (4.0.13)$$

Para esse exemplo, utilizamos o ETF SVXY para acompanhar os dados realizados e também é possível observar a coerência da sua evolução, principalmente quando olhamos a tendência da média das simulações na figura 4.4, com o comportamento inverso ao que temos no ETN VXX.

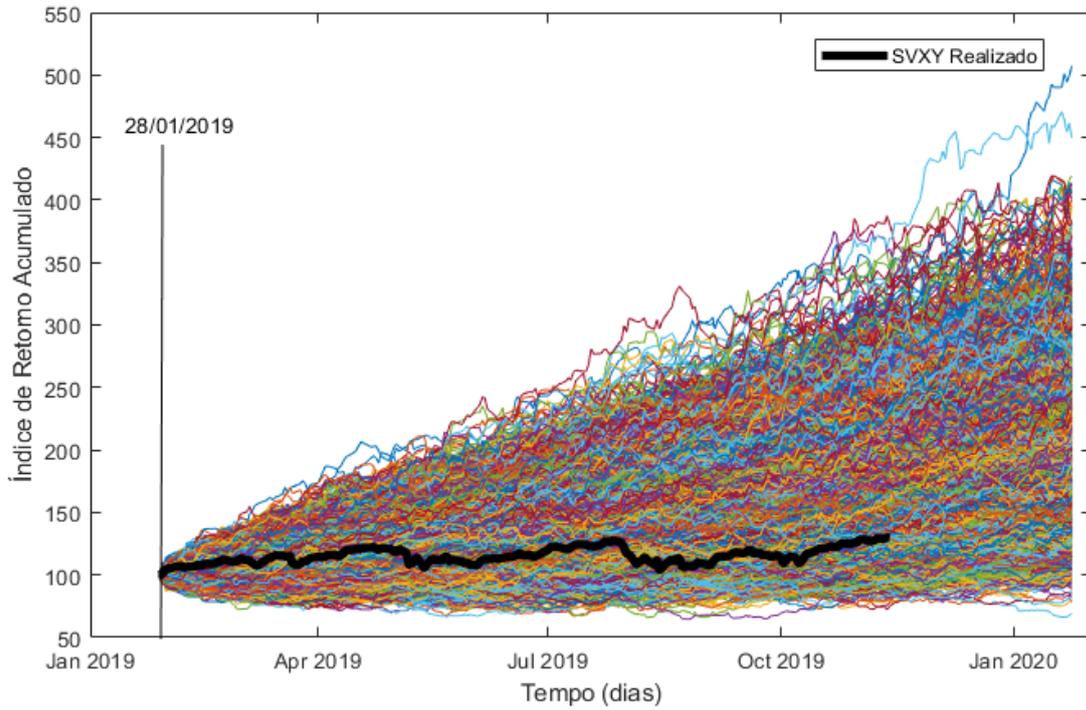


Figura 4.3: Comparativo entre as Simulações e o Realizado do ETF SVXY

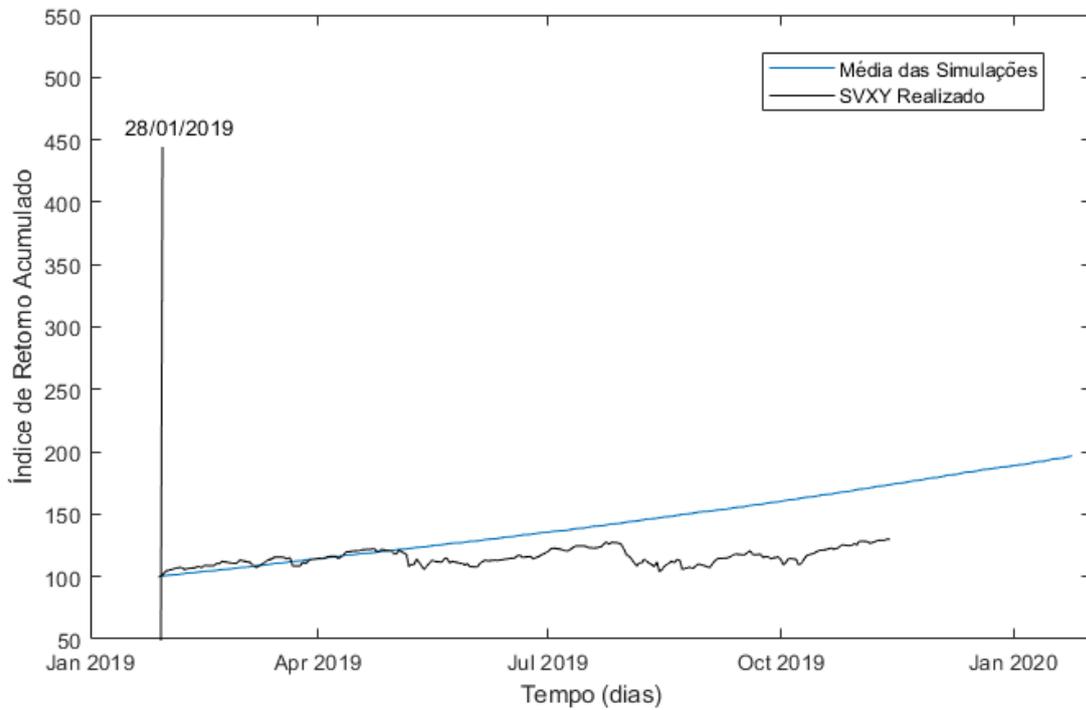


Figura 4.4: Comparativo entre a média das Simulações e o Realizado do ETF SVXY



# Capítulo 5

## Conclusão

Neste trabalho, mostramos as principais características do modelo multifatorial Gaussiano proposto para o VIX no artigo Avellaneda and Papanicolaou [2018]. Vimos que esse modelo proporciona um entendimento detalhado da dinâmica da estrutura a termo para os contratos futuros do VIX e sua aplicação prática teve como foco os ETPs que, para replicar a rentabilidade dos seus *benchmarks*, montam portfólios compostos por 2 ou mais futuros que são rebalanceados diariamente com o intuito de manter suas maturidades constantes.

Além da derivação do modelo, realizamos o procedimento de calibragem que, por se tratar de um modelo Gaussiano, utilizou-se o filtro de Kalman como ferramenta para obter estimativas dos parâmetros. Na implementação, é necessário escrever as equações na forma espaço-estado e, além disso, fizemos uma breve revisão dos conceitos atrelados a tópicos de finanças, probabilidade, cálculo estocástico e estatística, que nos permitiram ter o conhecimento necessário para entender todas as etapas e analisar os resultados.

Foi possível observar que os contratos mais curtos (principalmente o VIX à vista e o CMF 30d) apresentaram maiores erros, enquanto os contratos mais longos tiveram melhores aproximações. Tal fato já poderia ser esperado a partir de uma análise prévia das principais estatísticas da nossa amostra, pois os contratos mais curtos (devido à dinâmica do mercado) possuem maior variação. Mas, de maneira geral, constatamos que o modelo proposto no artigo original consegue capturar de forma bastante satisfatória a dinâmica de evolução da estrutura a termo do VIX.

Com o intuito de complementar as análises feitas no artigo original, fizemos uma estimação com uma amostra na qual não constava a série histórica dos log-preços do VIX à vista. Tal proposta baseou-se no argumento de que (assim como nos modelos de *commodities*) o preço à vista não é negociado no mercado e, portanto, ao retirar essa série da amostra, os erros na aproximação (suavização feita pelo Filtro de Kalman) poderiam ser menores. Vimos que essa melhora ocorreu nos contratos com vencimento acima de 60 dias. Esse ganho acabou sendo maior do que o aumento dos erros nos contratos mais curtos (30 dias e 60 dias), de modo que o erro médio percentual de toda a amostra, comparativamente, melhorou.

A estimação do modelo sem o VIX à vista envolveu um menor tempo computacional. A contribuição da dissertação com a estimação alternativa permite a simulação de instrumentos atrelados a contratos mais longos com menores erros.

Este aspecto nos faz pensar se poderia existir algum ganho no caso de o agente (investidor e/ou *trader*) utilizar uma estratégia de otimização diferente, de acordo com o objetivo do seu produto. Por exemplo, se o ETP busca replicar os rendimentos de médio prazo, a otimização com uma amostra sem o VIX à vista seria melhor do que aquela com a amostra completa? Mas, como este não era o objetivo do presente trabalho, deixamos como sugestão para pesquisas futuras.

De posse das estimativas dos parâmetros, fomos capazes de realizar simulações para analisar as rentabilidades futuras dos ETPs. Esse procedimento é de grande ajuda para o agente que busca/faz a gestão desse tipo de produto, pois permite ter uma ideia mais clara do risco ao qual está incorrendo e, a partir disso, traçar sua estratégia de proteção para o portfólio como um todo.

Por último, deixamos como tópicos para outras pesquisas futuras: inclusão de mais um fator (ficando com um total de 3), com o intuito de testar a melhoria nas aproximações obtidas nos contratos de maturidade mais curta; utilização de preços de risco de mercado diferentes do estipulado em 2.2.1, sem depender dos fatores, ou com funcionais diferentes, de modo a analisar as alterações geradas nas estimativas; inclusão de uma estrutura autoregressiva nos erros do modelo e, por fim, tentativas de estruturação de outros ETPs que não apenas os vistos aqui, por exemplo, ETPs similares às sugestões encontradas em Alexander and Korovilas [2012].

# Referências Bibliográficas

- Fernando Antonio Lucena Aiube. *Modelos quantitativos em finanças com enfoque em commodities*. Bookman Editora, 2013.
- Carol Alexander and Dimitris Korovilas. Understanding etns on vix futures. *Available at SSRN 2043061*, 2012.
- Marco Avellaneda and Andrew Papanicolaou. Statistics of vix futures and their applications to trading volatility exchange-traded products. *Journal of Investment Strategies*, 7(2):1–32, 2018.
- Lorenzo Bergomi. *Stochastic volatility modeling*. Chapman and Hall/CRC, 2015.
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. doi: 10.1086/260062.
- R Boroujerdi and K Fogertey. Etf’s: The rise of the machines. *Goldman Sachs Equity Research April*, 2015.
- Peter Carr and Dilip Madan. Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, pages 458–476, 2001.
- Peter Carr and Liuren Wu. Variance risk premiums. *The Review of Financial Studies*, 22(3): 1311–1341, 2008.
- Chicago Board Options Exchange CBOE. The cboe volatility index®—vix®. white paper, 2018.
- ETF.db. Etf database: The original & comprehensive guide to etfs, 2019. URL [https://etfdb.com/screener/#page=1&asset\\_class=volatility&volatility\\_terms=Mid%20Term%20Volatility%2CShort%20Term%20Volatility](https://etfdb.com/screener/#page=1&asset_class=volatility&volatility_terms=Mid%20Term%20Volatility%2CShort%20Term%20Volatility). acesso em 30 de Setembro de 2019.
- K Gregory and M Droumaguet. Navigating the vix etp market: Vix etps popular, but often misunderstood. *Goldman Sachs Options Research April*, 2017.
- K Gregory and A Timcenko. When to get long the vix: Assessing the risk-reward of long vix trades. *Goldman Sachs Options Research April*, 2016.
- Desmond J Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546, 2001.
- Joanne M Hill, Dave Nadig, and Matt Hougan. *A comprehensive guide to exchange-traded funds (ETFs)*. CFA Institute Research Foundation, 2015.

- John C Hull. *Opções, futuros e outros derivativos*. Bookman Editora, 2016.
- Peter E Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 2013.
- Ralf Korn and Elke Korn. *Option pricing and portfolio optimization: modern methods of financial mathematics*, volume 31. American Mathematical Soc., 2001.
- Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
- Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer Science & Business Media, 2004.
- Robert H Shumway and David S Stoffer. *Time series analysis and its applications: with R examples*. Springer, 2017.
- J Michael Steele. *Stochastic calculus and financial applications*, volume 45. Springer Science & Business Media, 2012.

# Apêndice A

## Fórmula do VIX

Neste apêndice vamos apresentar uma derivação da fórmula geral para cálculo do índice VIX, mostrando a sua ligação com a precificação dos contratos de *Swap* de Variância. Vamos usar como base a nota técnica 22 de Hull [2016].

Um *swap* de variância é um contrato para trocar a taxa de variância realizada  $\bar{V}$  entre o tempo 0 e o tempo  $T$  por uma taxa de variância predeterminada. A taxa de variância é o quadrado da volatilidade ( $\bar{V} = \bar{\sigma}^2$ ). Como veremos ao longo da demonstração, a taxa de variância entre o tempo 0 e o tempo  $T$  pode ser replicada usando um portfólio de opções de venda e de compra.

Assumindo que o preço do ativo tem a dinâmica de evolução dada pela seguinte EDE (movimento browniano geométrico):

$$dS = (r - q)Sdt + \sigma SdW \Leftrightarrow \frac{dS}{S} = (r - q)dt + \sigma dW \quad (\text{A.0.1})$$

na qual  $r$ ,  $q$  e  $\sigma$  são constantes e representam a taxa de juros livre de risco, uma taxa de pagamento de dividendos contínua e a volatilidade, respectivamente e,  $dW$  é um processo de Wiener com respeito à uma medida de probabilidade neutra ao risco.

Pelo lema de Itô (conforme descrito em B.1.7), se escolhermos  $f(S) = \ln(S)$ , podemos escrever:

$$\begin{aligned} f_t(t, S_t) &= \frac{\partial f(t, S_t)}{\partial t} = 0 \\ f_s(t, S_t) &= \frac{\partial f(t, S_t)}{\partial S_t} = \frac{1}{S} \\ f_{ss}(t, S_t) &= \frac{\partial^2 f(t, S_t)}{\partial S_t^2} = -\frac{1}{S^2} \end{aligned}$$

De modo que:

$$\begin{aligned} d[f(S)] &= \frac{\partial f(S)}{\partial S} dS + \frac{\partial f(S)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f(S)}{\partial S^2} \langle dS \rangle \\ d \ln(S) &= \frac{1}{S} (S(r - q)dt + S\sigma dW_t) - \frac{1}{2} \frac{1}{S^2} S^2 \sigma^2 dt \end{aligned}$$

Ficando então com:

$$d \ln(S) = \left( r - q - \frac{\sigma^2}{2} \right) dt + \sigma dW \quad (\text{A.0.2})$$

Agora, subtraindo A.0.1 de A.0.2, vamos obter:

$$\frac{dS}{S} - d \ln(S) = \cancel{(r - q)dt} + \cancel{\sigma dW} - \left( \cancel{r - q} - \frac{\sigma^2}{2} \right) dt - \cancel{\sigma dW}$$

Portanto,

$$\frac{\sigma^2}{2} dt = \frac{dS}{S} - d \ln(S) \quad (\text{A.0.3})$$

Agora, integrando do instante 0 (zero) até um determinado instante  $T$ :

$$\begin{aligned} \int_0^T \frac{\sigma^2}{2} dt &= \int_0^T \frac{dS}{S} - \int_0^T d \ln(S) \\ \frac{1}{2} \sigma^2 T &= \int_0^T \frac{dS}{S} - (\ln(S_T) - \ln(S_0)) \end{aligned}$$

De modo que:

$$\frac{1}{2} \sigma^2 T = \int_0^T \frac{dS}{S} - \ln \left( \frac{S_T}{S_0} \right) \quad (\text{A.0.4})$$

Mas, sabendo que a variância é o quadrado da volatilidade ( $\bar{V} = \sigma^2$ ), teremos:

$$\frac{1}{2} \bar{V} T = \int_0^T \frac{dS}{S} - \ln \left( \frac{S_T}{S_0} \right)$$

ou então, podemos ainda escrever que, a taxa de variação média realizada  $\bar{V}$  é dada por:

$$\bar{V} = \frac{2}{T} \int_0^T \frac{dS}{S} - \frac{2}{T} \ln \left( \frac{S_T}{S_0} \right) \quad (\text{A.0.5})$$

Tomando o valor esperado em relação à medida neutra ao risco  $\mathbb{Q}$ :

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}] = \frac{2}{T} \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \frac{dS}{S} \right] - \frac{2}{T} \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S_0} \right) \right] \quad (\text{A.0.6})$$

Vamos agora analisar separadamente cada um dos termos da equação anterior. Começando pelo primeiro termo do lado direito de A.0.6, vemos que é possível calcular a esperança proveniente da equação A.0.1. Assim:

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \frac{dS}{S} \right] = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T (r - q) dt \right] + \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \sigma dW \right]$$

Como  $r$  e  $q$  são constantes e sabemos também que a esperança da integral estocástica vale 0 (zero), ficaremos com:

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \frac{dS}{S} \right] = (r - q)T$$

Portanto, substituindo o resultado desse valor esperado na equação A.0.6:

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}] = \frac{2}{T}(r - q)T - \frac{2}{T}\mathbb{E}^{\mathbb{Q}}\left[\ln\left(\frac{S_T}{S_0}\right)\right] \quad (\text{A.0.7})$$

Além disso, da equação A.0.2, podemos desenvolver:

$$\begin{aligned} \ln(S_T) - \ln(S_0) &= \left(r - q - \frac{\sigma^2}{2}\right)dt + \sigma dW \\ \ln\left(\frac{S_T}{S_0}\right) &= \left(r - q - \frac{\sigma^2}{2}\right)dt + \sigma dW \\ S_T &= S_0 e^{\left(r - q - \frac{\sigma^2}{2}\right)dt + \sigma dW} \Rightarrow S_T = S_0 e^X \end{aligned}$$

Na qual, podemos descrever a variável aleatória  $X$  como:

$$X = \left(r - q - \frac{\sigma^2}{2}\right)dt + \sigma dW$$

E, por isso,  $X$  é Gaussiana, com:

$$\mathbb{E}^{\mathbb{Q}}[X] = \left(r - q - \frac{\sigma^2}{2}\right)T \quad e \quad Var[X] = \sigma^2 T$$

Dessa maneira, podemos calcular:

$$\mathbb{E}^{\mathbb{Q}}[S_T] = S_0 \exp\left[\mathbb{E}^{\mathbb{Q}}[X] + \frac{Var[X]}{2}\right] = S_0 e^{\left(r - q - \frac{\sigma^2}{2}\right)T + \frac{\sigma^2}{2}T} = S_0 e^{(r - q)T} = F_0$$

Na qual  $F_0$  é o preço a termo do ativo para um contrato com vencimento no tempo  $T$ .

Assim, tomando:

$$\ln\left(\frac{F_0}{S_0}\right) = \frac{\ln(S_0 e^{(r - q)T})}{\ln(S_0)} = \frac{\ln(S_0)(r - q)T}{\ln(S_0)} = (r - q)T$$

Podemos reescrever A.0.7 como:

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}] = \frac{2}{T} \ln\left(\frac{F_0}{S_0}\right) - \frac{2}{T}\mathbb{E}^{\mathbb{Q}}\left[\ln\left(\frac{S_T}{S_0}\right)\right] \quad (\text{A.0.8})$$

Vamos agora desenvolver o segundo termo à direita da equação acima. Primeiro, vamos considerar a seguinte integral e analisar o seu resultado para algum valor  $S^*$  de  $S$ :

$$\int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK$$

Quando  $S^* < S_T$ :

$$\int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK = 0$$

Quando  $S^* > S_T$ :

$$\begin{aligned} \int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK &= \int_{K=0}^{S^*} \frac{1}{K^2} (K - S_T, 0) dK \\ &= \left( \ln(K) + \frac{S_T}{K} \right) \Big|_{S_T}^{S^*} = \ln(S^*) + \frac{S_T}{S^*} - \left[ \ln(S_T) + \frac{S_T}{S_T} \right] \\ &= \ln\left(\frac{S^*}{S_T}\right) + \frac{S_T}{S^*} - 1 \end{aligned}$$

Em seguida, podemos considerar:

$$\int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK$$

Quando  $S^* > S_T$ :

$$\int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK = 0$$

Quando  $S^* < S_T$ :

$$\begin{aligned} \int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK &= \int_{K=S^*}^{S_T} \frac{1}{K^2} (S_T - K, 0) dK \\ &= \left( -\frac{S_T}{K} - \ln(K) \right) \Big|_{S^*}^{S_T} = -\frac{S_T}{S_T} - \ln(S_T) - \left[ -\frac{S_T}{S^*} - \ln(S^*) \right] \\ &= \ln\left(\frac{S^*}{S_T}\right) + \frac{S_T}{S^*} - 1 \end{aligned}$$

Portanto, destes resultados, segue que:

$$\int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK + \int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK = \ln\left(\frac{S^*}{S_T}\right) + \frac{S_T}{S^*} - 1$$

De modo que, para todos os valores de  $S^*$  vale a seguinte identidade:

$$\ln\left(\frac{S_T}{S^*}\right) = \frac{S_T}{S^*} - 1 - \int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK - \int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK \quad (\text{A.0.9})$$

Isso mostra que uma variável que paga o  $\ln(S_T)$  pode ser replicada usando um portfólio de derivativos. Este resultado pode ser usado em conjunto com a equação A.0.5 para fornecer um portfólio de replicação para  $\bar{V}$ . Tomando o valor esperado em um mundo neutro ao risco na equação A.0.9, podemos escrever:

$$\mathbb{E}^{\mathbb{Q}} \left[ \ln\left(\frac{S_T}{S^*}\right) \right] = \mathbb{E}^{\mathbb{Q}} \left[ \frac{S_T}{S^*} - 1 \right] - \mathbb{E}^{\mathbb{Q}} \left[ \int_{K=0}^{S^*} \frac{1}{K^2} \max(K - S_T, 0) dK \right] - \mathbb{E}^{\mathbb{Q}} \left[ \int_{K=S^*}^{\infty} \frac{1}{K^2} \max(S_T - K, 0) dK \right]$$

Mas, vale notar que:

$$\mathbb{E}^{\mathbb{Q}}[\max(K - S_T, 0)] = e^{RT} p(K)$$

É o preço de uma opção de venda (Put) com *strike*  $K$ .

$$\mathbb{E}^{\mathbb{Q}}[\max(S_T - K, 0)] = e^{RT} c(K)$$

É o preço de uma opção de compra (Call) com *strike*  $K$ . E, ainda:

$$\mathbb{E}^{\mathbb{Q}}[S_T] = F_0$$

Além disso,  $T$  representa a maturidade e  $R$  a taxa livre de risco até o respectivo vencimento. Com isso, vemos que:

$$\mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S^*} \right) \right] = \frac{F_0}{S^*} - 1 - \int_{K=0}^{S^*} \frac{1}{K^2} e^{RT} p(K) dK - \int_{K=S^*}^{\infty} \frac{1}{K^2} e^{RT} c(K) dK \quad (\text{A.0.10})$$

Sendo este um caso específico do resultado encontrado em Carr and Madan [2001], também conhecida como *decomposição de Carr-Madan*, ela nos mostra que qualquer *payoff* da forma  $f(S_T)$ , onde  $f$  é uma função de classe  $C^2$ , pode ser replicado por uma combinação estática de *calls*, *puts* e contratos a termo de mesmo *strike* e mesmo vencimento.

Agora, podemos colocar que:

$$\begin{aligned} \ln \left( \frac{S_T}{S_0} \right) &= \ln(S_T) - \ln(S_0) \\ &= \ln(S_T) - \ln(S^*) + \ln(S^*) - \ln(S_0) \\ &= \ln \left( \frac{S_T}{S^*} \right) + \ln \left( \frac{S^*}{S_0} \right) \end{aligned}$$

Agora, tomando o valor esperado na equação acima:

$$\mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S_0} \right) \right] = \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S^*} \right) \right] + \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S^*}{S_0} \right) \right]$$

Portanto, ficamos com:

$$\mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S_0} \right) \right] = \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S^*} \right) \right] + \ln \left( \frac{S^*}{S_0} \right) \quad (\text{A.0.11})$$

De modo que, agora, somos capazes de combinar as equações, [A.0.8](#), [A.0.10](#) e [A.0.11](#), para calcular:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[\bar{V}] &= \frac{2}{T} \ln \left( \frac{F_0}{S_0} \right) - \frac{2}{T} \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S_0} \right) \right] \\
&= \frac{2}{T} \ln \left( \frac{F_0}{S_0} \right) - \frac{2}{T} \left\{ \ln \left( \frac{S^*}{S_0} \right) + \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S^*} \right) \right] \right\} \\
&= \frac{2}{T} \ln \left( \frac{F_0}{S_0} \right) - \frac{2}{T} \ln \left( \frac{S^*}{S_0} \right) - \frac{2}{T} \mathbb{E}^{\mathbb{Q}} \left[ \ln \left( \frac{S_T}{S^*} \right) \right] \\
&= \underbrace{\frac{2}{T} \ln \left( \frac{F_0}{S_0} \right) - \frac{2}{T} \ln \left( \frac{S^*}{S_0} \right)}_{\frac{2}{T} \ln \left( \frac{F_0}{S^*} \right)} - \frac{2}{T} \left( \frac{F_0}{S^*} - 1 - \int_{K=0}^{S^*} \frac{1}{K^2} e^{RT} p(K) dK - \int_{K=S^*}^{\infty} \frac{1}{K^2} e^{RT} c(K) dK \right)
\end{aligned}$$

Com isso, podemos concluir que o valor esperado para a variância média de hoje ( $t = 0$ ) até uma determinada maturidade ( $t = T$ ) é dado por:

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}] = \frac{2}{T} \ln \left( \frac{F_0}{S^*} \right) - \frac{2}{T} \left( \frac{F_0}{S^*} - 1 \right) + \frac{2}{T} \left( \int_{K=0}^{S^*} \frac{1}{K^2} e^{RT} p(K) dK + \int_{K=S^*}^{\infty} \frac{1}{K^2} e^{RT} c(K) dK \right) \quad (\text{A.0.12})$$

Na qual,  $F_0$  é o preço a termo do ativo para um contrato com vencimento em  $T$ ,  $c(K)$  é o preço da opção de compra com *strike*  $K$  e maturidade  $T$ ,  $p(K)$  é o preço da opção de venda com *strike*  $K$  e maturidade  $T$ ,  $R$  é a taxa de juros livre de risco para um vencimento em  $T$  e  $S^*$  é qualquer valor de  $S$ .

Supondo que o preço de opções europeias com *strike*  $K_i$ , ( $1 \leq i \leq n$ ) são conhecidos e temos que  $K_1 < K_2 < K_3 < \dots < K_{n-1} < K_n$ . Além disso, definimos  $S^*$  igual ao primeiro preço de exercício abaixo de  $F_{0,T}$ , em seguida, aproximamos as integrais da seguinte forma:

$$\int_{K=0}^{S^*} \frac{1}{K^2} e^{RT} p(K) dK + \int_{K=S^*}^{\infty} \frac{1}{K^2} e^{RT} c(K) dK = \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) \quad (\text{A.0.13})$$

onde:

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}, \quad 2 \leq i \leq n-1. \quad \Delta K_1 = K_2 - K_1 \quad e \quad \Delta K_n = K_n - K_{n-1}$$

E, na função  $Q(K_i)$ :

Se,  $K_i < S^*$ , então a função  $Q(K_i)$  corresponde ao preço de uma Put com *strike*  $K_i$

Se,  $K_i > S^*$ , então a função  $Q(K_i)$  corresponde ao preço de uma Call com *strike*  $K_i$

Se,  $K_i = S^*$ , então a função  $Q(K_i)$  corresponde à média dos preços de uma Call e uma Put com *strike*  $K_i$

Assim, substituindo a equação A.0.13 em A.0.12, vamos obter:

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}] = \frac{2}{T} \ln \left( \frac{F_0}{S^*} \right) - \frac{2}{T} \left( \frac{F_0}{S^*} - 1 \right) + \frac{2}{T} \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i)$$

ou, ainda,

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}]T = 2 \ln \left( \frac{F_0}{S^*} \right) - 2 \left( \frac{F_0}{S^*} - 1 \right) + 2 \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) \quad (\text{A.0.14})$$

Agora, vamos olhar mais especificamente para os dois primeiros termos do lado direito da equação anterior. De acordo com a função  $\ln(N)$  e o seu polinômio de Maclaurin a seguir:

$$\ln(N) = (N - 1) - \frac{1}{2}(N - 1)^2 + O((N - 1)^2)$$

podemos obter:

$$\ln \left( \frac{F_0}{S^*} \right) = \left( \frac{F_0}{S^*} - 1 \right) - \frac{1}{2} \left( \frac{F_0}{S^*} - 1 \right)^2 + O \left( \left( \frac{F_0}{S^*} - 1 \right)^2 \right)$$

Isso significa que podemos fazer a seguinte aproximação:

$$\ln \left( \frac{F_0}{S^*} \right) = \left( \frac{F_0}{S^*} - 1 \right) - \frac{1}{2} \left( \frac{F_0}{S^*} - 1 \right)^2$$

ou, ainda,

$$\ln \left( \frac{F_0}{S^*} \right) - \left( \frac{F_0}{S^*} - 1 \right) = -\frac{1}{2} \left( \frac{F_0}{S^*} - 1 \right)^2 \quad (\text{A.0.15})$$

Assim, podemos usar [A.0.15](#) para reescrever [A.0.14](#) como segue:

$$\mathbb{E}^{\mathbb{Q}}[\bar{V}]T = -\left( \frac{F_0}{S^*} - 1 \right)^2 + 2 \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) \quad (\text{A.0.16})$$

Desta forma, o processo a ser realizado em um determinado dia para se calcular o índice VIX é calcular  $\mathbb{E}^{\mathbb{Q}}[\bar{V}]T$  para as opções negociadas no mercado e com vencimentos imediatamente anterior e posterior a 30 dias. A variância cumulativa esperada para os 30 dias, sob a medida neutra ao risco, é calculada a partir desses dois números usando interpolação. Em seguida, multiplica-se por  $\frac{365}{30}$ , e o índice é definido como a raiz quadrada desse resultado.

Isso significa que a variância acumulada esperada, sob a medida neutra ao risco, é calculada como:

$$\sigma^2 T = -\left( \frac{F_0}{K_0} - 1 \right)^2 + 2 \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i)$$

ou seja, o índice de volatilidade VIX se baseia na seguinte equação (colocando-a no mesmo formato que o encontrado em CBOE [2018]):

$$\sigma^2 = -\frac{1}{T} \left( \frac{F_0}{K_0} - 1 \right)^2 + \frac{2}{T} \sum_{i=1}^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) \quad (\text{A.0.17})$$

E, para o cálculo exato do índice, incluindo a interpolação e as respectivas regularizações temporais, temos:

$$VIX = 100 \times \sqrt{\left\{ T_1 \sigma_1^2 \left[ \frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}} \right] + T_2 \sigma_2^2 \left[ \frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}} \right] \right\} \times \frac{N_{365}}{N_{30}}} \quad (\text{A.0.18})$$

A explicação com exemplos e procedimentos operacionais detalhados de como calcular o índice a partir de dados reais é encontrada em CBOE [2018].

# Apêndice B

## Revisão Teórica

Neste apêndice, vamos apresentar algumas definições básicas de cálculo estocástico, estatística e outros tópicos de matemática e finanças que serão utilizados no decorrer do trabalho. As principais referências utilizadas nesse capítulo foram Aiube [2013], Korn and Korn [2001], Oksendal [2013], Shreve [2004] e Steele [2012].

### B.1 Probabilidade e Cálculo Estocástico

Nesta seção vamos definir alguns conceitos básicos de probabilidade e cálculo estocástico.

**Definição 2.1 (Espaço de Probabilidade)** *Um espaço de probabilidade é uma tripla  $(\Omega, \mathcal{F}, \mathbb{P})$ , no qual temos: i)  $\Omega$  é um conjunto não vazio, chamado de espaço amostral; ii)  $\mathcal{F}$  é uma  $\sigma$ -álgebra de subconjuntos de  $\Omega$ ; e iii)  $\mathbb{P}$  é uma medida de probabilidade.*

**Definição 2.2 (Filtração)** *Seja  $\Omega$  um conjunto não vazio. Seja  $T$  fixo, um número positivo, e assumamos que, para cada  $t \in [0, T]$  existe uma  $\sigma$ -álgebra  $\mathcal{F}_t$ . Assumamos também que, se  $s < t$ , então todo o conjunto em  $\mathcal{F}_s$  está também em  $\mathcal{F}_t$ . Então, chamamos a coleção de  $\sigma$ -álgebras  $\mathcal{F}_t$ ,  $0 \leq t \leq T$ , de filtração. Além disso, as  $\sigma$ -álgebras da filtração devem estar contidas em  $\mathcal{F}$ .*

**Definição 2.3 (Espaço de Probabilidade Filtrado)** *Um espaço de probabilidade filtrado é uma quádrupla  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ , onde  $(\Omega, \mathcal{F}, \mathbb{P})$  é o espaço de probabilidade e  $\{\mathcal{F}_t\}_{t \in [0, T]}$  é a filtração.*

**Definição 2.4 (Processo Estocástico)** *O conjunto  $\{(X_t, \mathcal{F}_t)\}_{t \in [0, T]}$  formado pela filtração  $\{\mathcal{F}_t\}_{t \in [0, T]}$  e uma família de variáveis aleatórias,  $X_t$ , tomando valores no  $\mathbb{R}^n$  com  $X_t$   $\mathcal{F}_t$ -mensurável, é chamado de processo estocástico com filtração  $\{\mathcal{F}_t\}_{t \in [0, T]}$ .*

Podemos complementar que, os processos estocásticos são de grande utilidade na modelagem de eventos aleatórios que evoluem ao longo do tempo (principalmente no estudo de Séries Temporais em Finanças). Portanto, é fundamental construir uma estrutura de informação que tenha essa mesma característica para estudarmos os mais diversos fenômenos.

O conceito matemático que descreve a dinâmica da propagação das informações ao longo do tempo é chamado filtração. E, uma forma mais intuitiva de descrever uma filtração é que ela funciona como um crivo que controla o fluxo de informações. Mais especificamente, no nosso trabalho, é importante saber que: i)  $\mathcal{F}_t$  representa a informação disponível no tempo  $t$ ; e ii)

A estrutura da definição assegura que a quantidade de informação cresce conforme o tempo evolui. Além disso, nenhuma informação é perdida com o decorrer do tempo (ou seja, qualquer informação disponível no momento  $s$  ainda está disponível no tempo  $t$ , uma vez que  $t \geq s$ ).

**Definição 2.5 (Martingal)** *Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade, seja  $T$  fixo e positivo, e seja  $\{\mathcal{F}_t\}_{t \in [0, T]}$ , uma filtração de sub- $\sigma$ -álgebras de  $\mathcal{F}$ . Considere o processo estocástico adaptado (ou seja, para cada  $t$ , a variável aleatória  $X_t$  é  $\mathcal{F}_t$ -mensurável)  $M_t$ ,  $0 \leq t \leq T$ , tal que  $\mathbb{E}[M_t] < \infty \forall t \in [0, T]$ . i) Se  $\mathbb{E}[M_t | \mathcal{F}_s] = M_s \forall 0 \leq s \leq t \leq T$ , dizemos que este processo é um martingal; ii) Se  $\mathbb{E}[M_t | \mathcal{F}_s] \geq M_s \forall 0 \leq s \leq t \leq T$ , dizemos que este processo é um submartingal; iii) Se  $\mathbb{E}[M_t | \mathcal{F}_s] \leq M_s \forall 0 \leq s \leq t \leq T$ , dizemos que este processo é um supermartingal.*

Desta definição, podemos ainda destacar que a noção de adaptabilidade pode ser interpretada como a incapacidade de prever eventos futuros. Dizemos que  $X_s$  é considerado conhecido em qualquer tempo  $t$  ( $t \geq s$ ) se  $\{X_t : t \geq 0\}$  é adaptado à filtração  $\{\mathcal{F}_t\}$  (ou seja, um processo estocástico  $\{X_t : t \geq 0\}$  adaptado à filtração implica que o valor de  $X_t$  é quase certamente determinado pela filtração  $\{\mathcal{F}_t\}$ ). Por outro lado, para cada  $t \geq s$ ,  $X_t$  em geral não é mensurável no tempo  $s$ . A variável  $X_t$  é considerada desconhecida porque as probabilidades dos eventos descritos por  $X_t$  não podem ser calculados com base na informação disponível antes do tempo  $t$ .

Além disso, quando analisamos a primeira condição da definição anterior, vemos que, no caso de um martingal, a melhor previsão para um evento futuro é o valor atual do processo. Ou seja, o ganho futuro esperado é zero e, por causa disso, os martingais modelam um jogo justo, que não favorece nenhum dos participantes (tornando-se, assim, uma ferramenta muito relevante no campo de apreçamento de ativos). Pode-se dizer que um jogo é justo quando os ganhos incrementais em qualquer estágio são nulos ao se condicionar aos resultados obtidos historicamente.

Na teoria moderna de Finanças, a propriedade martingal se torna uma condição crucial para que o mercado seja eficiente, ou seja, um mercado no qual a informação sobre preços passados é incorporada completamente e instantaneamente no preço atual do ativo. Esta eficiência de mercado nos diz que não é possível obter lucros excessivos em operações baseadas em tendências históricas dos preços (de modo que, a expectativa de mudança do preço futuro, condicionada ao histórico de preços, deve ser zero).

**Definição 2.6 (Movimento Browniano)** *Seja o espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ . Definimos um movimento Browniano como sendo um processo  $\{W_t\}_{t \geq 0}$  com caminhos (realizações)  $\omega$  contínuos, tal que: i)  $W_0 = 0$ ,  $\mathbb{P}$ -q.t.p.; ii)  $W_t - W_s \sim \mathcal{N}(0, t - s)$ ,  $s < t$ ; iii)  $W_t - W_s$  independentes de  $W_u - W_v$  para  $0 \leq v \leq u \leq s \leq t$ .*

**Definição 2.7 (Filtração Browniana)** *Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  o espaço de probabilidade no qual está definido o Movimento Browniano  $\{W_t\}_{t \geq 0}$ . A filtração do Movimento Browniano é uma coleção de  $\sigma$ -álgebras  $\{\mathcal{F}_t\}_{t > 0}$ , que satisfaz: i) **Acumula Informações** - Para  $0 \leq s < t$ , todo conjunto em  $\mathcal{F}_s$  está também em  $\mathcal{F}_t$ . Isso quer dizer que existe ao menos tanta informação disponível em um tempo posterior  $\mathcal{F}_t$  quanto em um tempo anterior  $\mathcal{F}_s$ ; ii) **Adaptatividade** - Para cada  $t \geq 0$ , o Movimento Browniano  $\{W_t\}_{t \geq 0}$  no instante  $t$  é  $\mathcal{F}_t$ -mensurável. De modo que, a informação disponível no tempo  $t$  é suficiente para avaliar o Movimento Browniano  $\{W_t\}_{t \geq 0}$  nesse tempo; iii) **Independência de Incrementos Futuros** - Para  $0 \leq t < u$ , todo incremento  $W_u - W_t$  é independente da filtração  $\mathcal{F}_t$ . Ou seja, qualquer incremento do Movimento*

Browniano após o tempo  $t$  é independente da informação disponível no tempo  $t$ .

**Teorema 2.1** *O Movimento Browniano  $\{W_t\}_{t \geq 0}$  é um martingal.*

A demonstração pode ser encontrada em Korn and Korn [2001].

**Definição 2.8 (Browniano d-dimensional)** *Um movimento Browniano d-dimensional é um processo estocástico com valores em  $\mathbb{R}^n$ ,  $W(t) = (W_1(t), W_2(t), \dots, W_d(t))$  tal que cada componente é um movimento Browniano unidimensional e  $W_i$  é independente de  $W_j$  se  $i \neq j$ .*

As definições descritas acima são alguns dos principais conceitos matemáticos utilizados na moderna teoria de apreçamento em finanças para trabalhar com o fluxo de informações e a evolução histórica dos dados. Contudo, além disso, gostaríamos de quantificar de alguma forma essa evolução, de modo a sermos capazes de entender melhor o mecanismo gerador dos dados e, como consequência, realizar previsões sobre o comportamento da variável que está sob análise. O matemático que propôs as ferramentas necessárias para essa tarefa foi Kiyosi Itô (1915-2008), e apresentamos abaixo mais alguns conceitos necessários para o desenvolvimento dos cálculos realizados neste trabalho.

**Definição 2.9 (Variação Quadrática)** *Dado o processo estocástico  $\{X_t\}_{t \in [0, T]}$ , definimos a variação quadrática  $\langle X \rangle_t$ , como*

$$\langle X_t, X_t \rangle = \lim_{\|\Pi\| \rightarrow 0} \sum_{i=1}^n (X_{t_i^n} - X_{t_{i-1}^n})^2 \quad (\text{B.1.1})$$

onde  $\Pi = \{0 = t_0^n < \dots < t_n^n = T\}$ .

**Definição 2.10 (Processo Simples)** *Um processo estocástico  $X = \{X_t\}_{t \in [0, T]}$  é chamado processo simples se existem números reais  $0 = t_0 < t_1 < \dots < t_p = T$ ,  $p \in \mathbb{N}$ , e variáveis aleatórias  $\phi_i : \Omega \rightarrow \mathbb{R}$ ,  $i = 0, 1, \dots, p$  com  $\phi_0$   $\mathcal{F}_0$ -mensurável e  $\phi_i$   $\mathcal{F}_{t_{i-1}}$ -mensurável tal que valha a seguinte representação (para cada  $\omega \in \Omega$ ):*

$$X_t(\omega) = \phi_0(\omega)1_{\{0\}}(t) + \sum_{i=1}^p \phi_i(\omega)1_{(t_{i-1}, t_i]}(t) \quad (\text{B.1.2})$$

**Definição 2.11 (Integral de Itô para processos simples)** *Seja  $\Pi = t_0, t_1, \dots, t_n$  uma partição de  $[0, T]$ , isto é,  $0 = t_0 < t_1 < \dots < t_n = T$ . Seja  $X_t$  um processo simples, onde  $X_t$  é constante em cada subintervalo  $[t_j, t_{j+1})$  e, de maneira genérica, temos  $t_k < t < t_{k+1}$ . Podemos definir a Integral de Itô (Integral Estocástica) como:*

$$I_t = \int_0^t X_u dW_u = \sum_{j=0}^{k-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] + X_{t_k} [W_t - W_{t_k}] \quad (\text{B.1.3})$$

**Definição 2.12 (Integral de Itô Generalizada)** *Vamos assumir agora, que  $X_t, t \geq 0$  é adaptado à filtração  $\{\mathcal{F}_t\}_{t \geq 0}$  e, que a seguinte condição de integrabilidade é válida:*

$$\mathbb{E} \int_0^T X_t^2 dt < \infty \quad (\text{B.1.4})$$

escolhemos uma sequência  $X_t^{(n)}$  de processos simples, de modo que, quando  $n \rightarrow \infty$  esse processo converge para  $X_t$ . Aqui, o conceito de convergência quer dizer:

$$\lim_{n \rightarrow \infty} \mathbb{E} \int_0^T |X_t^{(n)} - X_t|^2 dt = 0$$

Podemos então definir a Integral de Itô para o integrando com variações contínuas pela fórmula:

$$I_t = \int_0^t X_u dW_u = \lim_{n \rightarrow \infty} \int_0^t X_u^{(n)} dW_u, \quad 0 \leq t \leq T \quad (\text{B.1.5})$$

onde o limite é tomado em  $L^2(\Omega, d\mathbb{P})$ .

Mais detalhes sobre o espaço  $L^2$  e a escolha da sequência de processos simples pode ser encontrada em Korn and Korn [2001].

**Teorema 2.2 (Propriedades da Integral de Itô)** *Seja  $T$  uma constante positiva e seja  $X_t, 0 \leq t \leq T$ , um processo estocástico que satisfaz B.1.4. Então, a integral de Itô em B.1.5, possui as seguintes propriedades: i) **Continuidade**: Os caminhos de  $I_t$  são contínuos; ii) **Adaptatividade**: Para cada  $t$ ,  $I_t$  é  $\mathcal{F}_t$ -mensurável; iii) **Linearidade**: Se  $I_t = \int_0^t X_u dW_u$  e  $J_t = \int_0^t Y_u dW_u$ , então  $I_t \pm J_t = \int_0^t (X_u \pm Y_u) dW_u$ ; e também, para qualquer constante  $c$ ,  $cI_t = \int_0^t cX_u dW_u$ ; iv) **Martingal**:  $I_t$  é um martingal contínuo; v) **Isometria de Itô**:  $\mathbb{E}[(\int_0^t X_u dW_u)^2] = \mathbb{E}[\int_0^t X_u^2 du]$ ; vi) **Varição Quadrática**:  $\langle I_t, I_t \rangle = \int_0^t X_u^2 du$ .*

Todas essas demonstrações detalhadas podem ser encontradas em Shreve [2004].

**Teorema 2.3 (Fórmula — ou Lema — de Itô: Caso Unidimensional)** *Seja  $X_t, t \geq 0$ , um processo de Itô e seja  $f(t, X)$  uma função para a qual as derivadas parciais  $f_t(t, X) = \frac{\partial f(t, X)}{\partial t}$ ,  $f_x(t, X) = \frac{\partial f(t, X)}{\partial X}$  e  $f_{xx}(t, X) = \frac{\partial^2 f(t, X)}{\partial X^2}$  estão definidas e são contínuas. Então, para todo  $T \geq 0$ ,*

$$f(T, X_T) = f(0, X_0) + \int_0^T f_t(t, X_t) dt + \int_0^T f_x(t, X_t) dX_t + \frac{1}{2} \int_0^T f_{xx}(t, X_t) d\langle X_t, X_t \rangle \quad (\text{B.1.6})$$

ou, na forma diferencial (com  $0 \leq t \leq T$ ):

$$d[f(t, X_t)] = \frac{\partial f(t, X_t)}{\partial t} dt + \frac{\partial f(t, X_t)}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f(t, X_t)}{\partial X_t^2} \langle dX_t \rangle^2 \quad (\text{B.1.7})$$

A demonstração pode ser consultada também em Shreve [2004].

**Teorema 2.4 (Integral de Itô com Integrando Determinístico)** *Seja  $\{W_s\}_{s \geq 0}$ , um movimento Browniano, e seja  $\Delta_s$  uma função determinística no tempo. Defina  $I_t = \int_0^t \Delta_u dW_u$ . Para cada  $t \geq 0$ , a variável aleatória  $I_t$  é normalmente distribuída com valor esperado zero e variância  $\int_0^t \Delta^2 ds$ .*

A demonstração desse resultado está em Shreve [2004].

**Definição 2.13 (Variação Cruzada)** *Sejam  $X$  e  $Y$  dois processos de Itô tomando valores reais com representação,*

$$\begin{aligned} X_t &= X_0 + \int_0^t K_s ds + \int_0^t \mathbf{H}_s dW_s \\ Y_t &= Y_0 + \int_0^t L_s ds + \int_0^t \mathbf{M}_s dW_s \end{aligned}$$

então, a variação cruzada será dada por:

$$\langle X, Y \rangle = \sum_{j=1}^m \int_0^t H_j(s) M_j(s) ds \quad (\text{B.1.8})$$

**Teorema 2.5 (Fórmula de Itô para o caso multivariado)** *Sejam  $X_1, X_2, \dots, X_n$  processos de Itô univariados dados por:*

$$dX_i = \mu_i(X_t, t)dt + \sigma_i(X_t, t)dW_{it}, \quad i = 1, \dots, n$$

*Seja  $f(X_1, \dots, X_n, t)$ , onde  $f(\cdot)$  é contínua e diferenciável duas vezes em relação à  $x_1, \dots, x_n$  e uma vez em relação à  $t$  (com derivadas contínuas), então o diferencial  $df(\cdot)$  é dado por:*

$$df(X_1, \dots, X_n, t) = \frac{\partial f}{\partial t} dt + \sum_{i=1}^n \frac{\partial f}{\partial x_i} dX_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \langle dX_i, dX_j \rangle \quad (\text{B.1.9})$$

Maiores detalhes são dados em Aiube [2013] e Oksendal [2013].

## B.2 Conceitos de finanças e apreçamento neutro ao risco

Nesta seção vamos apresentar algumas das principais definições relacionadas ao apreçamento neutro ao risco utilizadas ao longo do trabalho. Segundo Aiube [2013], o ponto central na metodologia é a suposição da inexistência da possibilidade de arbitragem. Complementa ainda que, dado que não há arbitragem, existem fundamentalmente duas teorias bem definidas para realizarmos o apreçamento. A primeira é a metodologia clássica desenvolvida por Black and Scholes [1973], e a segunda faz uso da medida martingal equivalente, na qual iremos nos aprofundar neste trabalho. Uma discussão mais profunda sobre os motivos de utilização de uma em detrimento da outra é feito em Aiube [2013]. Além disso, definimos também o que é um contrato futuro e a respectiva fórmula para o cálculo do seu preço.

**Definição 2.14 (Medidas Equivalentes)** *Duas medidas de probabilidade  $\mathbb{P}$  e  $\mathbb{Q}$  em  $(\Omega, \mathcal{F})$  são equivalentes se elas estão em concordância com relação a quais conjuntos de  $\mathcal{F}$  têm probabilidade zero.*

**Definição 2.15 (Derivada de Radon-Nikodym)** *Seja o espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ . Seja  $\mathbb{Q}$  em  $(\Omega, \mathcal{F})$  uma medida equivalente à  $\mathbb{P}$ . Seja  $Z$  uma função não negativa, de tal modo que seja a densidade de  $\mathbb{Q}$  em relação à  $\mathbb{P}$ . Então,  $Z$  é denominada de derivada de Radon-Nikodym de  $\mathbb{Q}$  em relação à  $\mathbb{P}$ , ou seja:*

$$Z(\omega) = \frac{d\mathbb{Q}(\omega)}{d\mathbb{P}(\omega)} \quad (\text{B.2.1})$$

Seja  $A \in \mathcal{F}$ , podemos escrever que a probabilidade de ocorrer o evento, avaliado pela medida  $\mathbb{Q}$  é:

$$\mathbb{Q}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \quad (\text{B.2.2})$$

onde  $\omega \in \Omega$ . Além disso, é possível provar que  $\mathbb{E}^{\mathbb{P}}(Z) = 1$ .

**Teorema 2.6 (Teorema de Girsanov)** *Seja o espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$  e  $X_t \in \mathbb{R}, 0 \leq t \leq T$  um processo de Itô que possui a seguinte dinâmica:*

$$dX_t = \mu_t(\omega)dt + \sigma_t(\omega)dW_t$$

onde  $W_t \in \mathbb{R}, \mu_t(\omega) \in \mathbb{R}$  e  $\sigma_t(\omega) \in \mathbb{R}_+^*$ .

Suponha que existam os processos adaptados a  $\mathcal{F}_t$ ,  $\theta_t(\omega) \in \mathbb{R}$  e  $\alpha_t(\omega) \in \mathbb{R}$ , de modo que:

$$\sigma_t(\omega)\theta_t(\omega, t) = \mu_t(\omega) - \alpha_t(\omega), \quad (\text{B.2.3})$$

Suponha ainda a condição de Novikov:

$$\mathbb{E}^{\mathbb{P}} \left[ e^{\int_0^T \theta_u^2 du} \right] < \infty$$

válida para  $\theta_t(\omega)$  e  $\alpha_t(\omega)$ , de modo que:

$$\mathbb{E}^{\mathbb{P}} \left[ \exp \left( \frac{1}{2} \int_0^T \theta_u^2(\omega) du \right) \right] < \infty, \quad \mathbb{E}^{\mathbb{P}} \left[ \exp \left( \frac{1}{2} \int_0^T \alpha_u^2(\omega) du \right) \right] < \infty$$

Definimos ainda:

$$\tilde{W}_t(\omega) = W_t(\omega) + \int_0^t \theta_u(\omega) du \quad (\text{B.2.4})$$

$$Z_t(\omega) = \exp \left[ - \int_0^t \theta_u(\omega) dW_u - \frac{1}{2} \int_0^t \theta_u^2(\omega) du \right] \quad (\text{B.2.5})$$

Considere a medida de probabilidade  $\mathbb{Q}$  equivalente a  $\mathbb{P}$ , tal que:

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

Então, sob a medida de probabilidade  $\mathbb{Q}$ ,  $\tilde{W}_t(\omega)$  é um processo Browniano padrão. A medida  $\mathbb{Q}$  é denominada Medida Martingal Equivalente (MME  $\mathbb{Q}$ ). Ainda,  $X_t$  escrito em função de  $\tilde{W}_t(\omega)$  é:

$$dX_t = \alpha_t(\omega)dt + \sigma_t(\omega)d\tilde{W}_t(\omega). \quad (\text{B.2.6})$$

A demonstração deste teorema pode ser consultada em Oksendal [2013] e Shreve [2004]. Podemos ver que este resultado permite a troca da medida de probabilidade de  $\mathbb{P}$  para  $\mathbb{Q}$ . Esta mudança essencialmente remove a tendência do processo, que está presente na medida  $\mathbb{P}$ . Além do caso unidimensional, este teorema pode ser estendido (de maneira análoga) para o caso multivariado.

De acordo com Aiube [2013], a medida martingal equivalente é uma medida de probabilidade conveniente para o apreçamento de derivativos, pois dado que o processo estocástico do ativo subjacente está escrito sob esta medida, o cálculo do valor esperado fornece o valor do derivativo. De maneira mais geral, se o derivativo no vencimento vale  $\Lambda_t$ , então o seu valor será dado por:

$$\Lambda_t = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}\Lambda_T|\mathcal{F}_t], 0 \leq t \leq T. \quad (\text{B.2.7})$$

E, ainda podemos colocar que, caso a taxa livre de risco não seja constante no intervalo  $0 \leq t \leq T$ , ficamos com:

$$\Lambda_t = \mathbb{E}^{\mathbb{Q}}[e^{\int_0^T R_u du}\Lambda_T|\mathcal{F}_t], 0 \leq t \leq T. \quad (\text{B.2.8})$$

As equações B.2.7 e B.2.8 são denominadas *equações fundamentais de apreçamento*.

**Definição 2.16 (Estratégia)** *Uma estratégia  $q_t$  define a quantidade de cada ativo, no instante  $t$ , de um portfólio, tal que:*

$$q_t(\omega) = (q_t^1(\omega), q_t^2(\omega), \dots, q_t^n(\omega)),$$

onde  $q_t^i(\omega)$  é um processo adaptado a  $\mathcal{F}_t = \sigma(B_s, s \leq t)$ , e  $n$  é o número de ativos disponíveis no mercado.

Aqui, se assumimos que  $X_t^i$  representa o preço do  $i$ -ésimo ativo no instante  $t$ , o valor do portfólio definido por uma estratégia nesse instante será dado por:

$$R_t(\omega) = \sum_{i=1}^n q_t^i(\omega) X_t^i, \omega \in \Omega$$

**Definição 2.17 (Estratégia autofinanciável)** *Dizemos que uma estratégia  $q_t(\omega)$  é autofinanciável quando as alterações no valor  $R_t(\omega)$  ocorrem apenas por variações nos preços dos ativos que compõem o portfólio, de modo que:*

$$dR_t(\omega) = \sum_{i=1}^n q_t^i(\omega) dX_t^i, \omega \in \Omega$$

**Definição 2.18 (Estratégia de Hedging)** *Define-se uma estratégia de proteção (hedging) como aquela para qual existe uma estratégia autofinanciável  $q_t(\omega)$  que faz com que o valor do derivativo no vencimento  $\Lambda_T$  seja igual ao valor do portfólio (quase certamente), ou seja:*

$$R_T(\omega) = \sum_{i=1}^n q_T^i(\omega) X_T^i = \Lambda_T(\omega), q.c. \quad (\text{B.2.9})$$

**Definição 2.19 (Mercado Completo)** *Dizemos que um mercado é completo quando, para todos os derivativos do mercado existe uma estratégia de proteção (hedging) na qual B.2.9 é atendida.*

**Definição 2.20 (Arbitragem)** *Uma oportunidade de arbitragem resume-se em uma estratégia autofinanciável  $q_t(\omega)$  tal que: i)  $R_0 = 0$ ; ii)  $\mathbb{P}(R_t \geq 0) = 1$  e  $\mathbb{P}(R_t > 0) > 0$ .*

Em outras palavras, arbitragem é uma estratégia que permite que um agente comece sem capital e, em um tempo posterior  $T$ , esteja certo de que não haverá perda de dinheiro e ainda

haja uma probabilidade positiva de ter ganho de dinheiro.

**Definição 2.21 (Medida Neutra)** *Uma medida de probabilidade  $\mathbb{Q}$  é dita neutra ao risco se  $\mathbb{P}$  e  $\mathbb{Q}$  são equivalentes e se, sob  $\mathbb{Q}$ , o processo de preços descontados de cada ativo ( $X_t^i$ ) do mercado é martingal.*

**Teorema 2.7 (Primeiro Teorema Fundamental de Finanças)** *O sistema de preços que prevalece no mercado não admite arbitragem se, e somente se, existe a medida neutra.*

A demonstração encontra-se em Aiube [2013].

**Teorema 2.8 (Teorema da Representação Martingal — TRM)** *Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  um espaço de probabilidade, seja  $W_t$  um processo Browniano padrão neste espaço e  $\mathcal{F}_t$  a filtração natural gerada por este processo com  $t \in [0, T]$ . Seja  $M_t$  um processo martingal em relação à  $\mathcal{F}_t$ , ou seja,  $\mathbb{E}[M_t | \mathcal{F}_u] = M_u$ , para  $u < t$ . Existe um processo adaptado  $H_t$ , com  $\mathbb{E}[\int_0^T H_u^2 du] < \infty$ , tal que:*

$$M_t = M_0 + \int_0^t H_u dW_u \quad (\text{B.2.10})$$

Uma demonstração mais detalhada deste resultado é encontrada em Oksendal [2013], Shreve [2004].

**Teorema 2.9 (Segundo Teorema Fundamental de Finanças)** *Considere um modelo de mercado que possui uma medida de probabilidade neutra ao risco. Um mercado é completo se, e somente se, a medida de probabilidade neutra ao risco é única.*

Uma demonstração pode ser encontrada em Aiube [2013].

Ainda, de acordo com Aiube [2013], o primeiro teorema fundamental coloca as condições necessárias e suficientes para o apereamento, pois a existência da medida neutra e da garantia da replicação do preço do derivativo permitem tal processo; enquanto o segundo teorema fundamental assegura a unicidade da medida neutra  $\mathbb{Q}$  mediante a existência de um mercado completo, e o contrário também é verdade. De modo que, passa a ser uma situação mais restritiva do que a do primeiro teorema.

*"Então, se o mercado é completo, o preço do derivativo, a partir da equação geral do apereamento, é único. Desta forma, a estrutura de replicação do preço do derivativo define um preço único para o mesmo. A estrutura de replicação está garantida pela ausência de arbitragem e pelo TRM. O preço do derivativo está garantido pela existência da medida neutra. E a unicidade deste preço está garantida pela unicidade de  $\mathbb{Q}$  (mercado completo)." Aiube [2013]*

**Definição 2.22 (Contrato Futuro)** *O Contrato Futuro é um acordo feito entre dois agentes anônimos e intermediado por uma bolsa, na qual o agente que assume a posição comprada, está obrigado a adquirir da contraparte (o agente que assume a posição vendida) o ativo que será entregue no vencimento  $T$ . A troca de fluxos de caixa ocorre por marcação a mercado de finida pelo mecanismo da conta margem.*

Mais detalhes sobre a mecânica operacional dos mercados futuros podem ser consultados em Hull [2016].

**Definição 2.23 (Preço Futuro)** *Seja  $X_t$  o preço à vista de um ativo no instante  $t \in [0, T]$ . Denominamos por  $F_{t,T}$  o preço futuro do ativo em  $T$  e negociado em  $t$ . O preço futuro é o valor esperado, sob a MME, do preço à vista no vencimento:*

$$F_{t,T} = \mathbb{E}^{\mathbb{Q}}[X_T | \mathcal{F}_t], \quad 0 \leq t \leq T. \quad (\text{B.2.11})$$

*E possui as seguintes propriedades: i)  $F_{t,T}$  é  $\mathcal{F}_t$ -mensurável; ii) os custos de entrada e saída do contrato são nulos e as mesmas podem ocorrer em qualquer instante até o vencimento; iii) a cada intervalo de tempo  $[t_k, t_{k+1})$  é feita a marcação a mercado; iv) o valor de  $F_{t,T}$ , na expiração, é o preço à vista, ou seja,  $F_{T,T} = X_T$ .*

**Teorema 2.10 (O Preço Futuro é Martingal)** *O preço do contrato futuro é um martingal se na medida neutra ao risco  $\mathbb{Q}$  ele satisfaz  $F_{T,T} = X_T$  e se o valor de uma posição comprada (ou vendida) a ser realizada durante um intervalo de tempo é sempre zero.*

A demonstração desse resultado e a construção detalhada do preço futuro pode ser consultada em Shreve [2004].

Com isso, cobrimos os principais aspectos do referencial teórico que nos permite realizar os cálculos necessários no desenvolvimento deste trabalho.

## B.3 Filtro de Kalman

Nesta seção vamos apresentar os fundamentos do filtro de Kalman e alguns conceitos inerentes à metodologia de filtragem em finanças. Aqui, as principais referências utilizadas foram Aiube [2013] e Shumway and Stoffer [2017].

Calibrar (estimar) um modelo em economia e finanças quer dizer encontrar o valor dos hiperparâmetros associados às equações propostas que minimizam os erros entre os dados reais e os estimados. Esta etapa torna-se crucial para validar as hipóteses estruturadas no âmbito da modelagem. Além disso, a partir da estimação, conseguimos verificar a significância estatística dos hiperparâmetros calculados e, pelo processo de filtragem, conseguimos estimar as variáveis não observáveis. De maneira resumida, nessa etapa conseguimos identificar se o modelo está apto a reproduzir as principais propriedades do processo gerador dos dados reais.

No contexto de Kalman, as variáveis de estado são não observáveis, mas possuem uma dinâmica conhecida. Enquanto isso, as variáveis observáveis são aquelas que observamos diretamente e que são utilizadas para estimar as variáveis de estado.

De acordo com Aiube [2013] o filtro de Kalman é um conjunto de equações matemáticas que constitui um processo recursivo eficiente de estimação de uma variável não observável (*variável de estado*), uma vez que o seu erro quadrático é minimizado. Com os dados reais, denominados *variável de observação*, a variável não observável pode ser estimada eficientemente de tal forma que, podemos calcular os estados passados, o estado presente, e os estados futuros. É um procedimento aplicável quando os modelos estão escritos sob a forma *espaço-estado*. Além

disso, esta ferramenta permite a estimação dos parâmetros desconhecidos do modelo através da maximização da verossimilhança via decomposição do erro de previsão. Suas condições ideais de utilização abrangem problemas lineares e Gaussianos. Quando lidamos com problemas não lineares pode-se usar o filtro de Kalman estendido, que é uma expansão do problema não linear através da série de Taylor.

De acordo com Shumway and Stoffer [2017], uma formulação genérica das equações para o filtro de Kalman é dada por:

$$\begin{cases} y_t &= A_t x_t + \Gamma u_t + v_t \\ x_t &= \Phi x_{t-1} + \Upsilon u_t + w_t \end{cases} \quad (\text{B.3.1})$$

Por conta da característica recursiva do algoritmo é possível atualizar as medidas em tempo real, fazendo com que a sua utilização seja frequente em sistemas de controle e rastreamento no campo da engenharia. Além disso, a sua aderência às finanças deve-se muito aos mercados financeiros que produzem informações a cada instante. Dessa maneira, as variáveis não-observáveis podem ser estimadas a partir dos preços obtidos no mercado a cada momento que uma nova informação é incorporada.

Segundo Shumway and Stoffer [2017] olhando para o aspecto prático, um dos principais objetivos envolvidos na análise do modelo na forma espaço-estado é gerar estimadores para os dados não observados  $x_t$  de posse da informação disponível  $y_{1:s} = \{y_1, \dots, y_s\}$  até o tempo  $s$ . De modo que, as estimativas geradas na equação de estado formam um componente essencial na estimativa dos parâmetros, como um todo. Além disso, podemos distinguir o problema em 3 naturezas distintas. São elas:

Quando  $s < t$  (ou seja, temos dados disponíveis até um determinado período antes do tempo na qual desejamos saber o valor da variável), o problema é chamado de previsão (*forecasting* ou *prediction*);

Quando  $s = t$  (ou seja, temos dados disponíveis até o exato instante de tempo na qual desejamos saber o valor da variável), o problema é chamado de filtragem (*filtering*);

Quando  $s > t$  (ou seja, temos dados disponíveis até um determinado período após o tempo na qual desejamos saber o valor da variável), o problema é chamado de suavização (*smoothing*).

Além dessas 3 estimativas, nos interessa também quantificar a precisão das mesmas. Dessa forma, normalmente utilizamos os seguintes critérios: *i*) Raiz do Erro Quadrático Médio (RMSE); *ii*) Erro Absoluto Médio (MAE); *iii*) Erro Absoluto Percentual Médio (MAPE).

O modelo proposto pelo filtro de Kalman requer 2 equações (uma denominada equação de observação e a outra denominada equação de transição) que juntas são denominadas forma espaço-estado, que vamos descrever em maiores detalhes agora (seguindo bem de perto o desenvolvimento feito em Aiube [2013]):

Seja  $\mathbf{y}_t \in \mathbb{R}^N$  uma série temporal multivariada. Este vetor representa as variáveis observáveis, enquanto  $\mathbf{x}_t \in \mathbb{R}^m$  representa o vetor das variáveis não observáveis (de estado). Uma vez que esses vetores foram definidos, podemos relacioná-los através das equações de observação e

transição.

A equação de observação relaciona o vetor das observações com as variáveis de estado, da seguinte forma:

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{d}_t + \epsilon_t \quad t = 1, \dots, T^* \quad (\text{B.3.2})$$

onde:

$\mathbf{y}_t$  é um vetor  $N \times 1$ , onde  $N$  é o número de séries temporais;

$\mathbf{H}_t$  é uma matriz  $N \times m$ , onde  $m$  é o número de variáveis de estado;

$\mathbf{x}_t$  é um vetor  $m \times 1$ , na qual representa as variáveis de estado não observáveis;

$\mathbf{d}_t$  é um vetor  $N \times 1$ ;

$\epsilon_t$  é uma distribuição normal multivariada com média zero e matriz de covariância  $\mathbb{Q}$ , isto é  $\epsilon_t \sim NID(\mathbf{0}, \mathbb{Q})$ .

As variáveis de estado são geradas por um processo Markoviano de primeira ordem e sua equação é denominada equação de transição tal que:

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{c}_t + \nu_t \quad t = 1, \dots, T^* \quad (\text{B.3.3})$$

onde:

$\mathbf{A}_t$  é uma matriz  $m \times m$ ;

$\mathbf{c}_t$  é um vetor  $m \times 1$ ;

$\nu_t \sim NID(\mathbf{0}, \mathbb{T})$ .

Além disso, o vetor inicial de estado  $\mathbf{x}_0$  tem média  $\hat{\mathbf{x}}_0 = \mathbb{E}[\mathbf{x}_0]$  e matriz de covariância  $\mathbf{P}_0$ . Os erros (ruídos)  $\epsilon_t$  e  $\nu_t$  são não correlacionados entre si e não correlacionados com o estado inicial.

A derivação do filtro de Kalman não será colocada neste trabalho mas, a mesma pode ser encontrada em Aiube [2013]. Desse modo, após a definição do modelo com as respectivas equações de observação e transição vamos nos aprofundar no algoritmo do filtro.

Basicamente, implementa-se um processo recursivo que calcula as equações de evolução da estimação das variáveis de estado e de sua matriz de covariância.

Pode-se dividir as equações do filtro de Kalman em dois grupo, tais que:

Equações de atualização do tempo (ou equações de previsão) - responsáveis pelo avanço das variáveis de estado e das covariâncias. Aqui, calculamos as estimativas anteriores (*a priori*) para o próximo instante. Representam o avanço no tempo de  $t - 1$  para  $t$ .

Seja  $\hat{\mathbf{x}}_{t-1}$  o estimador ótimo de  $\mathbf{x}_{t-1}$  baseado nas informações até  $t - 1$ , incluindo  $\mathbf{y}_{t-1}$ . Dados  $\hat{\mathbf{x}}_{t-1}$  e  $\mathbf{P}_{t-1}$ , o estimador ótimo de  $\mathbf{x}_t$  é dado por:

$$\hat{\mathbf{x}}_t^- = \mathbf{A}_t \hat{\mathbf{x}}_{t-1} + \mathbf{c}_t \quad (\text{B.3.4})$$

A matriz de covariância dos erros das variáveis de estado é dada por:

$$\mathbf{P}_t^- = \mathbf{A}_t \mathbf{P}_{t-1} \mathbf{A}_t^\top + \mathbb{T} \quad (\text{B.3.5})$$

Equações de atualização da medição (ou equações de correção) — responsáveis pela retroalimentação, incorporando uma nova informação da variável observável nas estimativas anteriores para obter um ganho (ou melhoria) na estimação posterior (*a posteriori*). Quando uma nova observação  $\mathbf{y}_t$  é registrada, o estimador  $\hat{\mathbf{x}}_t$  de  $\mathbf{x}_t$  pode ser melhorado ou corrigido.

Assim, as equações de atualização (correção) das medições são:

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}_t^\top \left( \mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^\top + \mathbb{Q} \right)^{-1} \quad (\text{B.3.6})$$

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t^- + \mathbf{K}_t (\mathbf{y}_t - \mathbf{H}_t \hat{\mathbf{x}}_t^- - \mathbf{d}_t) \quad (\text{B.3.7})$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_t^- \quad (\text{B.3.8})$$

De modo que, esses dois grupos funcionam de maneira conjunta como um sistema de retroalimentação (conforme a figura B.1).

Assim, com todas as equações mapeadas, o primeiro passo do algoritmo é o cálculo de  $\mathbf{K}_t$  (também chamado de ganho de kalman, na equação B.3.6. Em seguida, a nova informação  $\mathbf{y}_t$  é incorporada à previsão anterior  $\hat{\mathbf{x}}_t^-$  juntamente com a matriz de ganho  $\mathbf{K}_t$  através da equação B.3.7, gerando a estimativa posterior  $\hat{\mathbf{x}}_t$ . Por fim, obtém-se a matriz de covariância dos erros a partir da equação B.3.8. Com isso, avançamos no tempo para o instante  $t + 1$  e damos início ao novo ciclo do algoritmo com  $\hat{\mathbf{x}}_t$  e  $\mathbf{P}_t$  como dados de entrada das equações B.3.4 e B.3.5, respectivamente.

Ainda segundo Aiube [2013], nos modelos em engenharia, os hiperparâmetros presentes nas matrizes  $\mathbf{A}_t$  e  $\mathbf{H}_t$  e nos vetores  $\mathbf{d}_t$  e  $\mathbf{c}_t$  das equações B.3.2 e B.3.3 normalmente são estabelecidos pelas leis da física que governam os respectivos fenômenos. Contudo no campo da economia e finanças tal fato não acontece, de maneira que os hiperparâmetros também precisam ser estimados. Neste caso, ele também mostra que é possível utilizar uma função de verossimilhança via decomposição dos erros de previsão para estimar  $\Theta$  (o vetor dos respectivos hiperparâmetros).

No caso em que o modelo é Gaussiano, a função de verossimilhança será dada por:

$$\ln[L(\mathbf{y}|\Theta)] = -\frac{NT^*}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T^*} \ln |\mathbb{F}_t| - \frac{1}{2} \sum_{t=1}^{T^*} \eta_t^\top \mathbb{F}_t^{-1} \eta_t \quad (\text{B.3.9})$$

onde:

$T^*$  é o número de observações do modelo,

$\eta_t = \mathbf{y}_t - \hat{\mathbf{y}}_t^-$  é o vetor dos erros de previsão para cada  $t$ ,

$\mathbb{F}_t = \mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^\top + \mathbb{Q}$  é a matriz de covariância.

Desta forma, em cada rodada do filtro, todos os estados assim como todas as observações são estimadas e, a partir disso, pode-se maximizar a função de verossimilhança B.3.9 com relação a  $\Theta$  utilizando um algoritmo otimizador. Assim, nesta maximização, calculamos um vetor de hiperparâmetros  $\Theta^{(1)}$  a partir de uma inicialização  $\Theta^{(0)}$ . Este "chute inicial", também deve conter o vetor inicial dos estados  $\hat{\mathbf{x}}_0$  e a matriz de covariância  $\mathbf{P}_0$ , que podem ser a média e a variância incondicional do vetor de estado. Por fim, o algoritmo otimizador será encerrado quando o critério de convergência em relação ao vetor  $\Theta$  for atingido, como por exemplo, quando a norma euclidiana com as diferenças entre duas iterações for inferior a uma tolerância previamente estabelecida  $\|\Theta^{(k)} - \Theta^{(k-1)}\| < tol$ .

Desta última rodada  $k$ , utilizamos o vetor  $\Theta^{(k)}$  para estimar o vetor final de estado  $\hat{\mathbf{x}}_t$ ,  $t = 1, \dots, T^*$  e da previsão da variável de observação  $\hat{\mathbf{y}}_t = \mathbf{H}_t \hat{\mathbf{x}}_t + \mathbf{d}_t$ . De modo que ficamos com as estimativas das observações  $\hat{\mathbf{y}}_t$  e podemos calcular os respectivos erros em relação aos dados reais  $\epsilon_t = \mathbf{y}_t - \hat{\mathbf{y}}_t$ . Com isso, será possível analisar a qualidade do processo de calibragem/estimação através dos testes diagnósticos. A seguinte figura tenta mostrar de maneira mais intuitiva como todo esse procedimento funciona:

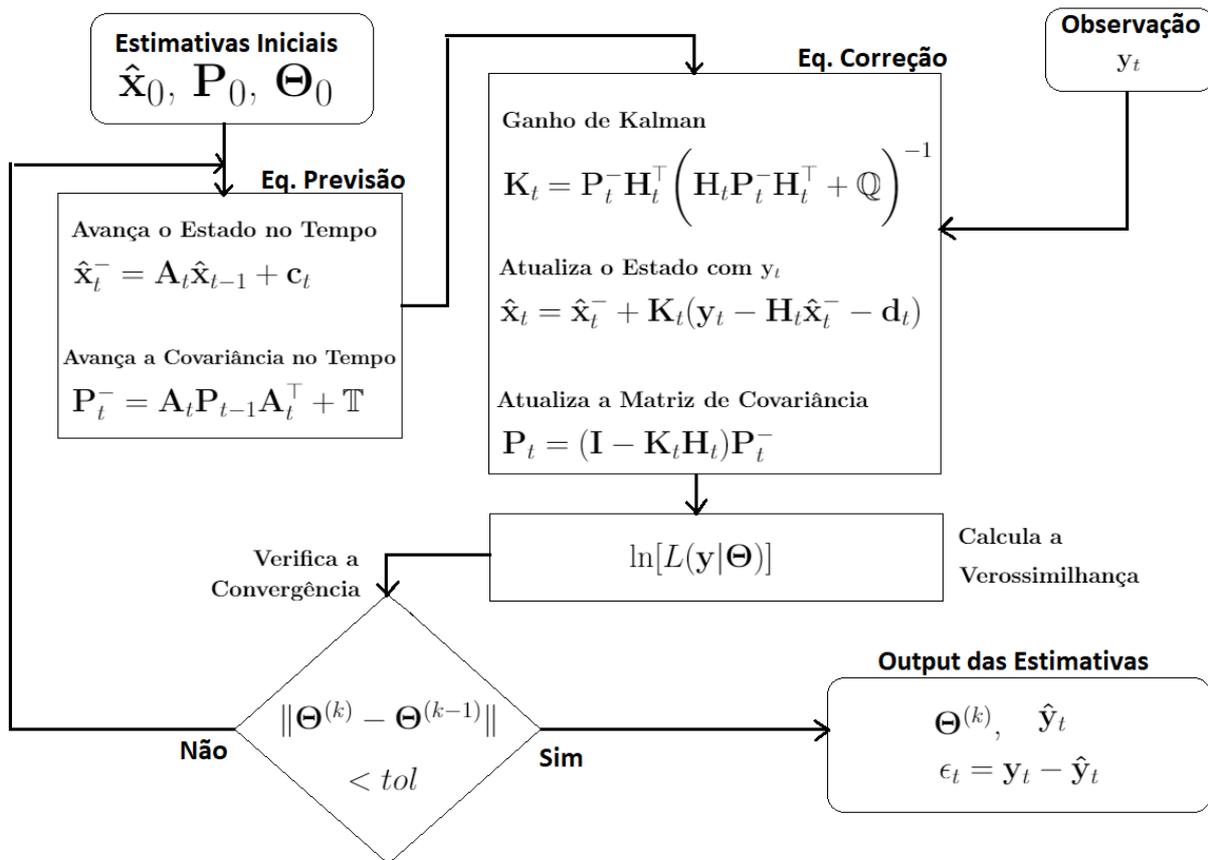


Figura B.1: Ilustração simplificada do processo de calibragem via Filtro de Kalman

E, dessa forma, colocamos as principais questões atreladas ao processo executado para

calibrar (estimar) o modelo proposto nesse trabalho.

## B.4 Método de Euler-Maruyama

Nesta seção, veremos uma forma de discretizar Equações Diferenciais Estocásticas (EDEs). De posse das estimativas dos parâmetros podemos, a partir desta metodologia, simular diversos cenários de evolução dos dados e analisar a coerência do modelo proposto. As principais referências e as demonstrações dos resultados são encontradas em Higham [2001] e Kloeden and Platen [2013].

O método de Euler-Maruyama é um método numérico que se propõe a resolver (discretizar) EDEs. Sabe-se que uma EDE pode ser escrita na forma integral da seguinte maneira:

$$X_t = X_0 + \int_0^t f(t, X_s)ds + \int_0^t g(t, X_s)dW_s, \quad 0 \leq t \leq T, \quad (\text{B.4.1})$$

onde,  $f : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $g : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$  são funções escalares e  $X_0$  é uma variável aleatória.

Sabemos também que a equação acima é solução da seguinte EDE:

$$dX_t = f(X_t)dt + g(X_t)dW_t, \quad X_0 = x_0 \quad 0 \leq t \leq T. \quad (\text{B.4.2})$$

Dessa forma, ao aplicar a metodologia numérica proposta na equação B.4.2 para o intervalo  $[0, T]$ , vamos primeiro discretizar o intervalo em  $L$  partes iguais, na qual  $L$  é um número inteiro e positivo. Portanto, seja  $\Delta t = \frac{T}{L}$  e  $\tau_j = j\Delta t$  e,  $X_j$  a notação utilizada para a aproximação numérica de  $X_{\tau_j}$ , podemos colocar o Método Euler-Maruyama como segue:

$$X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})(W_{\tau_j} - W_{\tau_{j-1}}), \quad j = 1, 2, \dots, L \quad (\text{B.4.3})$$

Assim, de maneira análoga, podemos colocar a forma integral de B.4.1 discretizada pelo respectivo método:

$$X_{\tau_j} = X_{\tau_{j-1}} + \int_{\tau_{j-1}}^{\tau_j} f(X_s)ds + \int_{\tau_{j-1}}^{\tau_j} g(X_s)dW_s \quad (\text{B.4.4})$$

Os incrementos  $W_{\tau_j} - W_{\tau_{j-1}}$  serão calculados com um passo  $\delta t$  de tal modo que  $\Delta t = R\delta t$ , com  $R \geq 1$ . Essa escolha vai garantir que o conjunto de pontos  $t_j$  onde o Browniano é discretizado contenha o conjunto de pontos  $\tau_j$ , onde a solução pelo método é calculada. De modo a garantir essa condição, podemos calcular os incrementos Brownianos conforme abaixo:

$$W_{\tau_j} - W_{\tau_{j-1}} = W_{jR\delta t} - W_{(j-1)R\delta t} = \sum_{k=R(j-1)+1}^{jR} dW_k \quad (\text{B.4.5})$$

Aqui, podemos complementar que, do movimento Browniano  $\{W_t\}_{t \geq 0}$ , esses incrementos são variáveis aleatórias gaussianas independentes com média  $\mathbb{E}[\Delta W_t] = 0$  e variância  $Var[\Delta W_t] = \Delta t$ .

Além disso, é válido apresentar os conceitos de convergência forte e fraca para a respectiva aproximação. Um método converge fortemente, com ordem  $\gamma$ , se existe uma constante  $C$ , tal que a seguinte condição é atendida:

$$\mathbb{E}|X_n - X_\tau| \leq C\Delta t^\gamma \quad (\text{B.4.6})$$

para qualquer  $\tau = n\Delta t \in [0, T]$  fixo e  $\Delta t$  suficientemente pequeno. Ainda, se  $f \in C^4$  e  $g \in C^4$ , ou seja, as duas funções possuem as quatro primeiras derivadas contínuas, a primeira derivada limitada e crescem no máximo em ordem polinomial, pode-se provar que o método irá convergir fortemente com ordem  $\gamma = \frac{1}{2}$ . Essa demonstração é encontrada em Kloeden and Platen [2013].

Supondo as condições acima verdadeiras, pode-se, ainda, definir para um determinado tempo  $t$  (fixo em  $[0, T]$ ) que o erro no sentido forte para um  $\Delta t$  pequeno, será dado por:

$$e_{\Delta t}^{forte} := \mathbb{E}|X_n - X_\tau| \leq C\Delta t^{\frac{1}{2}} \quad (\text{B.4.7})$$

Isso quer dizer que, se quisermos diminuir o erro 10 vezes, temos que diminuir o passo  $\Delta t$  em  $10^2$  vezes.

Por fim, podemos complementar dizendo que a convergência forte avalia a convergência dos caminhos de  $X_n - X_\tau$ , ou também a taxa de decaimento da média do erro conforme  $\Delta t$  tende para zero.

Por outro lado, podemos dizer que o método converge de maneira fraca a uma determinada ordem  $\gamma$  se existe uma constante  $C$  que respeita a seguinte equação para qualquer polinômio  $p$ :

$$|\mathbb{E}[p(X_n)] - \mathbb{E}[p(X_\tau)]| \leq C\Delta t^\gamma \quad (\text{B.4.8})$$

para qualquer  $\tau = n\Delta t \in [0, T]$  fixo e  $\Delta t$  suficientemente pequeno. Se  $f$  e  $g$  satisfazem condições apropriadas, o método de Euler-Maruyama possui ordem de convergência fraca  $\gamma = 1$ . Os detalhes são vistos nas referências indicadas no início desta seção.

Podemos então, levantar alguns cuidados que devem ser tomados ao escolher o tamanho do passo para o cálculo das Equações com o Método de Euler-Maruyama, de modo que: *i)* O passo não pode ser grande de modo a prejudicar a convergência do método; *ii)* O passo não pode ser de um tamanho próximo ao  $\epsilon$  da máquina, pois alguns erros de arredondamento farão o erro crescer ao invés de diminuir; *iii)* O tempo computacional aumenta exponencialmente conforme o inverso de  $\gamma$ , ou seja, para o caso com  $\gamma = \frac{1}{2}$ , ao diminuirmos o passo em 4 vezes, o tempo de cálculo da resposta deverá aumentar em 16 vezes, aproximadamente.

Dito isto, o *trade-off* deve ser atingido para um passo que respeite principalmente o primeiro e o último ponto colocados. Assim, a questão do  $\epsilon$  da máquina acaba não sendo tão relevante, pois muito antes de chegar neste aspecto, o tempo computacional já se apresentaria como um gargalo.

Dessa forma, cobrimos também o aspecto teórico relacionado à discretização de EDEs para realizar simulações. E, com tudo o que foi colocado, temos um arcabouço teórico suficiente para realizar a derivação do modelo proposto em Avellaneda and Papanicolaou [2018], implementar a sua calibragem e realizar simulações dos resultados de maneira a trazer mais robustez para aplicações práticas em produtos financeiros existentes no mercado.



# Apêndice C

## Gráficos: estimação

Neste apêndice vamos colocar mais alguns gráficos que acrescentam ao entendimento dos resultados do trabalho.

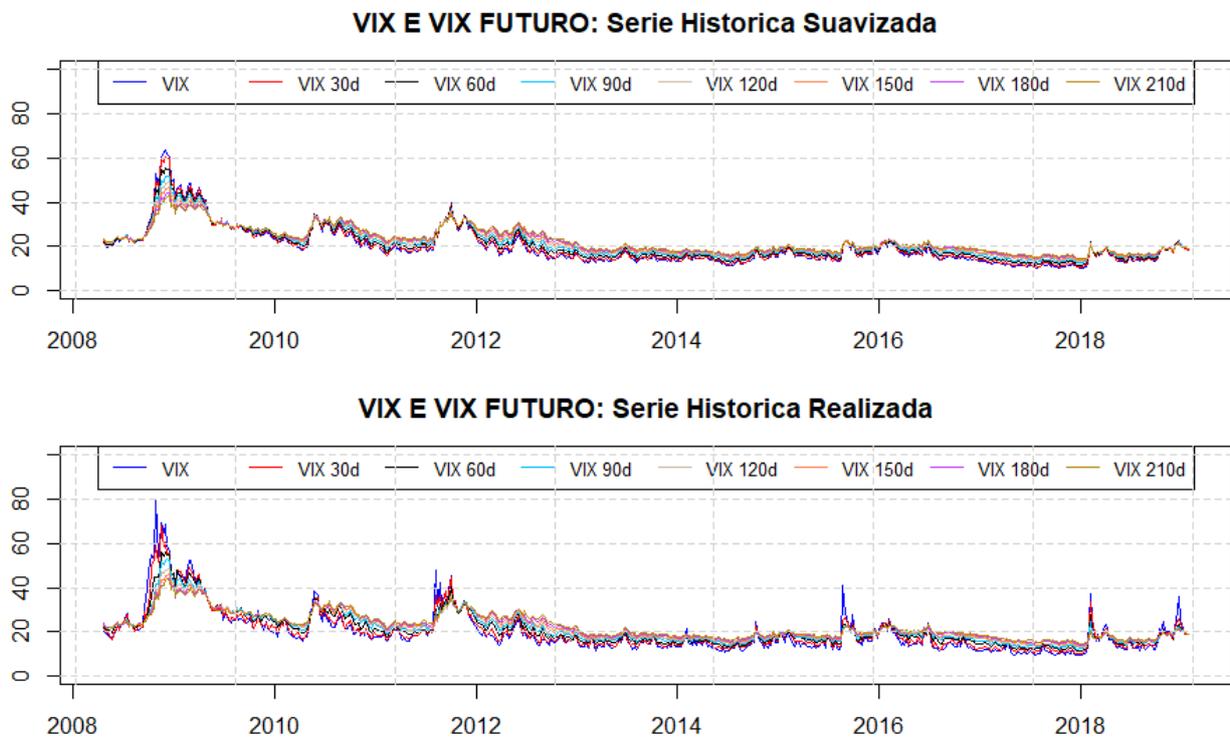


Figura C.1: Comparativo da Estrutura a Termo completa (amostra com o VIX à vista)

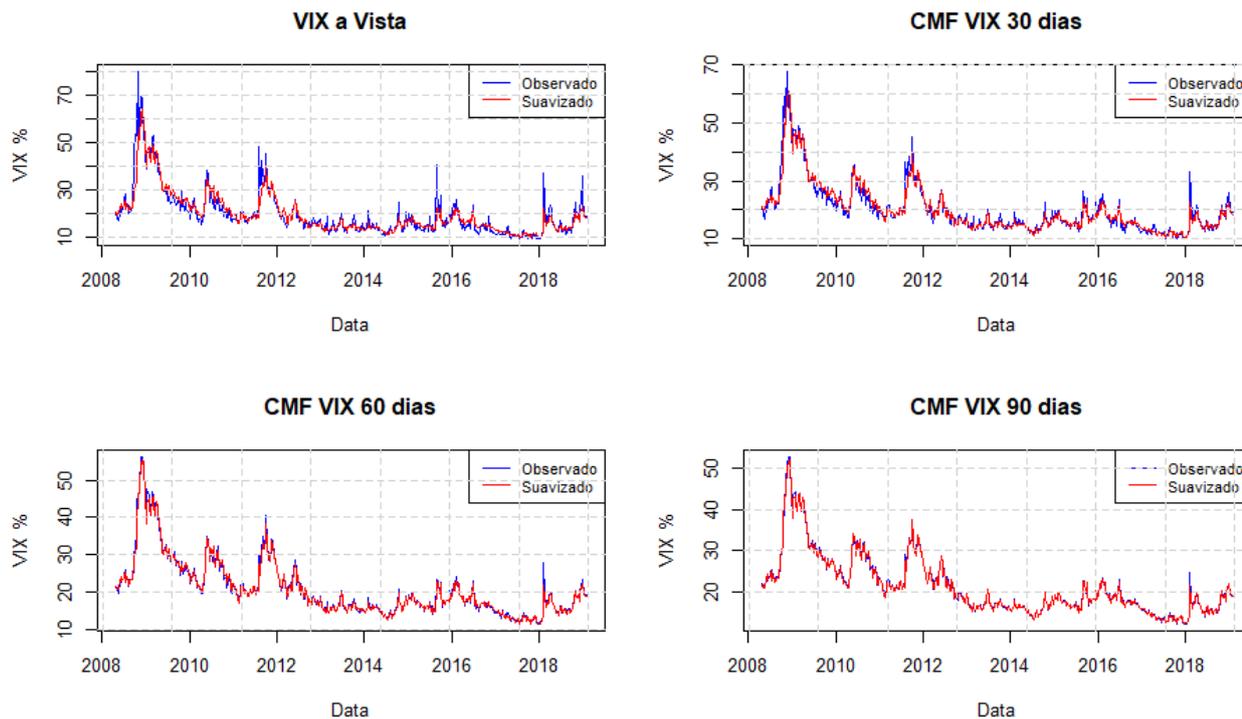


Figura C.2: Comparativo das estimativas do VIX à vista e 3 CMFs (amostra com o VIX à vista)

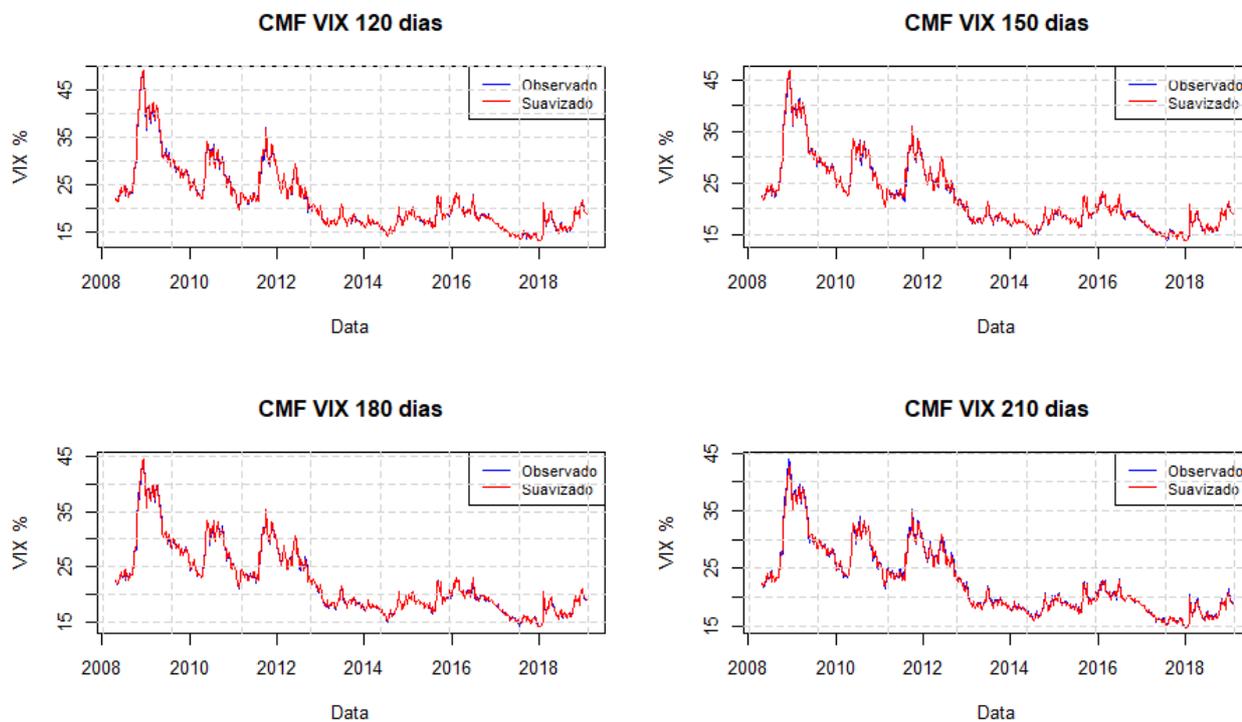


Figura C.3: Comparativo das estimativas de 4 CMFs (amostra com o VIX à vista)

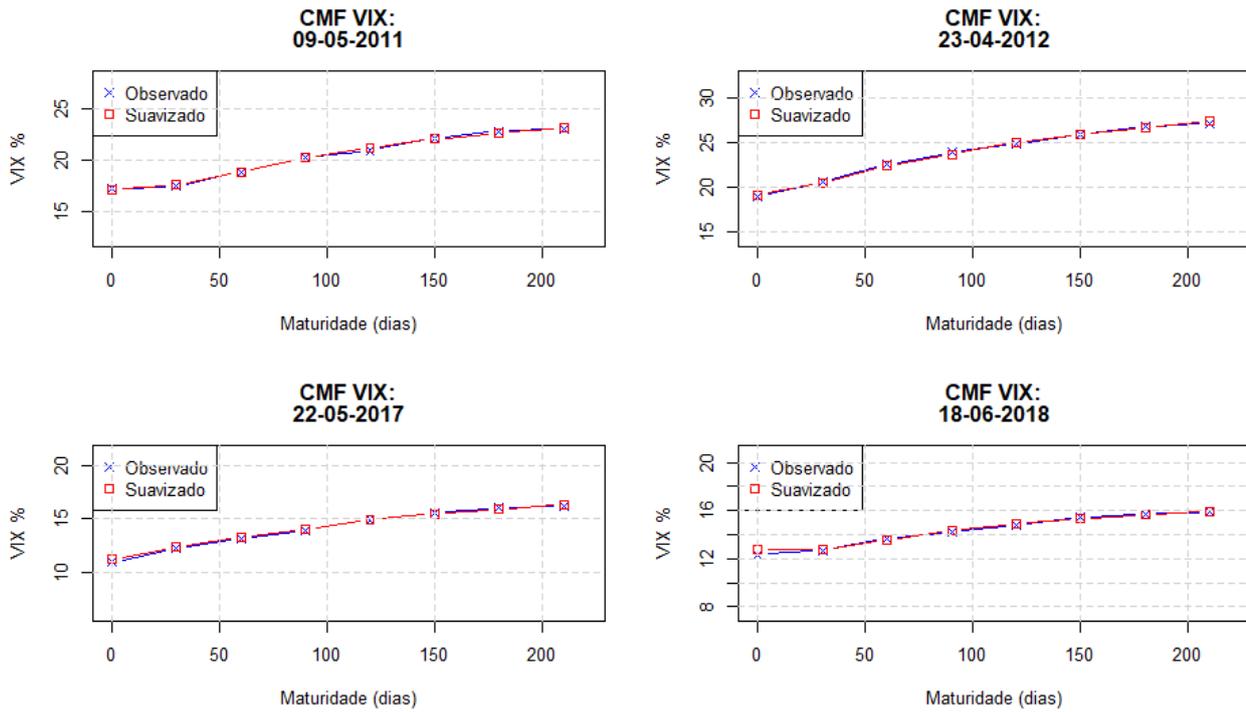


Figura C.4: Comparativo da estrutura a termo de dias em *Contango* (amostra com o VIX à vista)

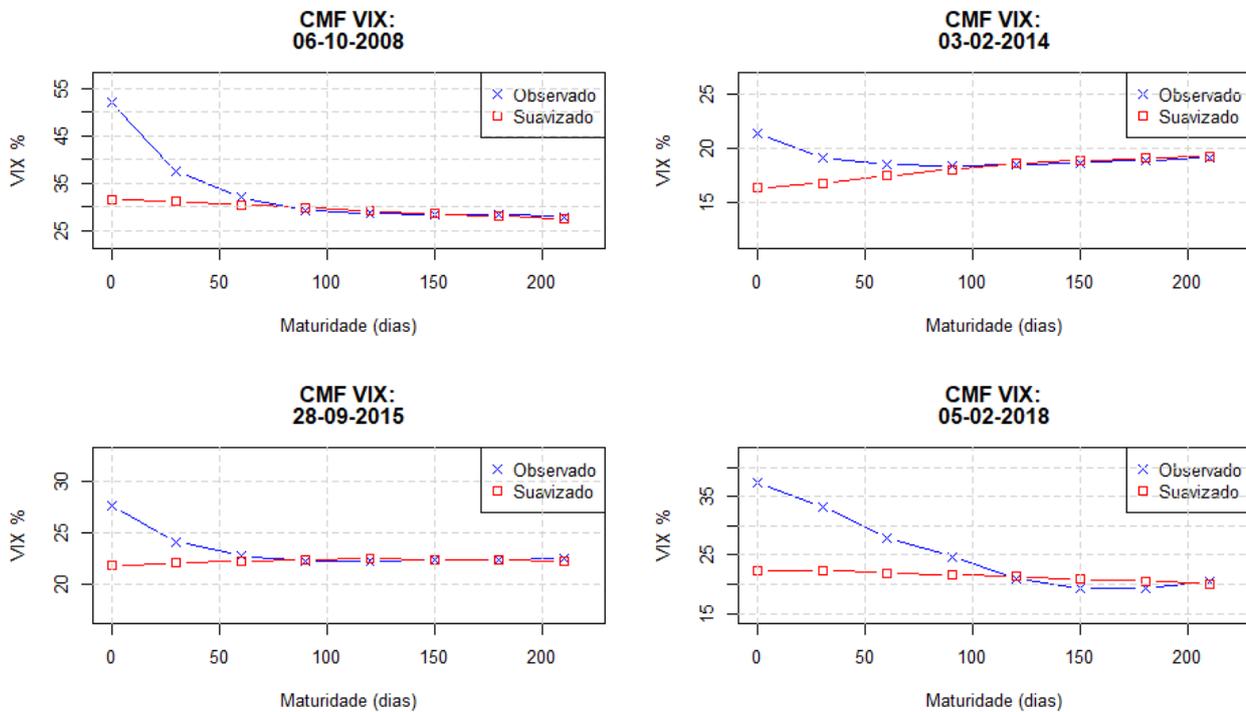


Figura C.5: Comparativo da estrutura a termo de dias em *Backwardation* (amostra com o VIX à vista)

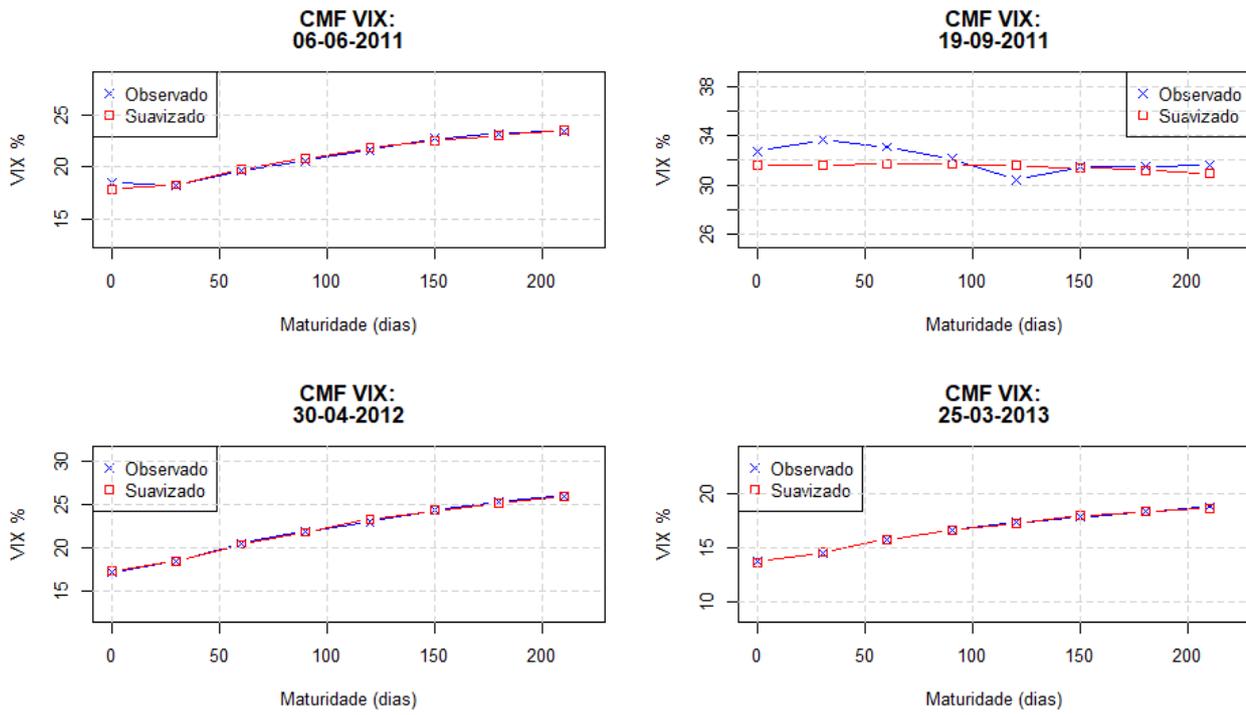


Figura C.6: Comparativo da estrutura a termo de dias escolhidos (amostra com o VIX à vista)



Figura C.7: Estimativas para as variáveis de estado  $X_1$  e  $X_2$  (amostra sem o VIX à vista)



Figura C.8: Estimativas para os preços de risco de mercado  $\lambda_1$  e  $\lambda_2$  (amostra sem o VIX à vista)

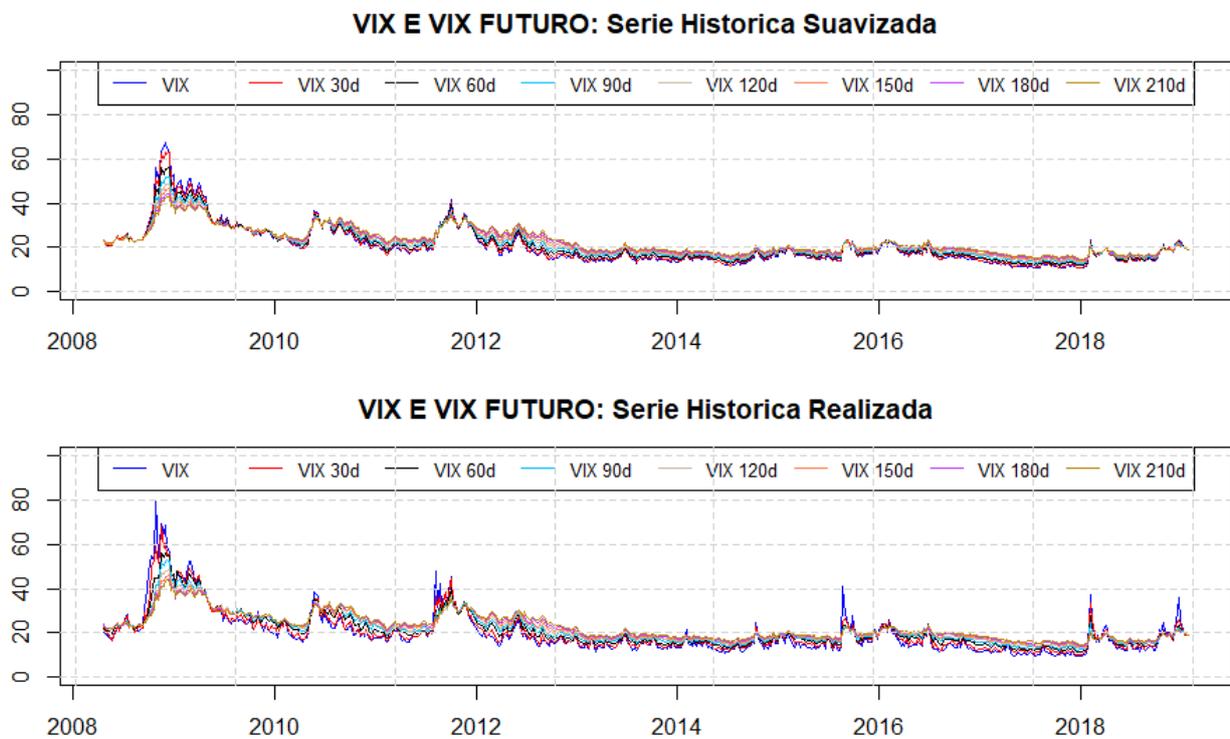


Figura C.9: Comparativo da Estrutura a Termo completa (amostra sem o VIX à vista)

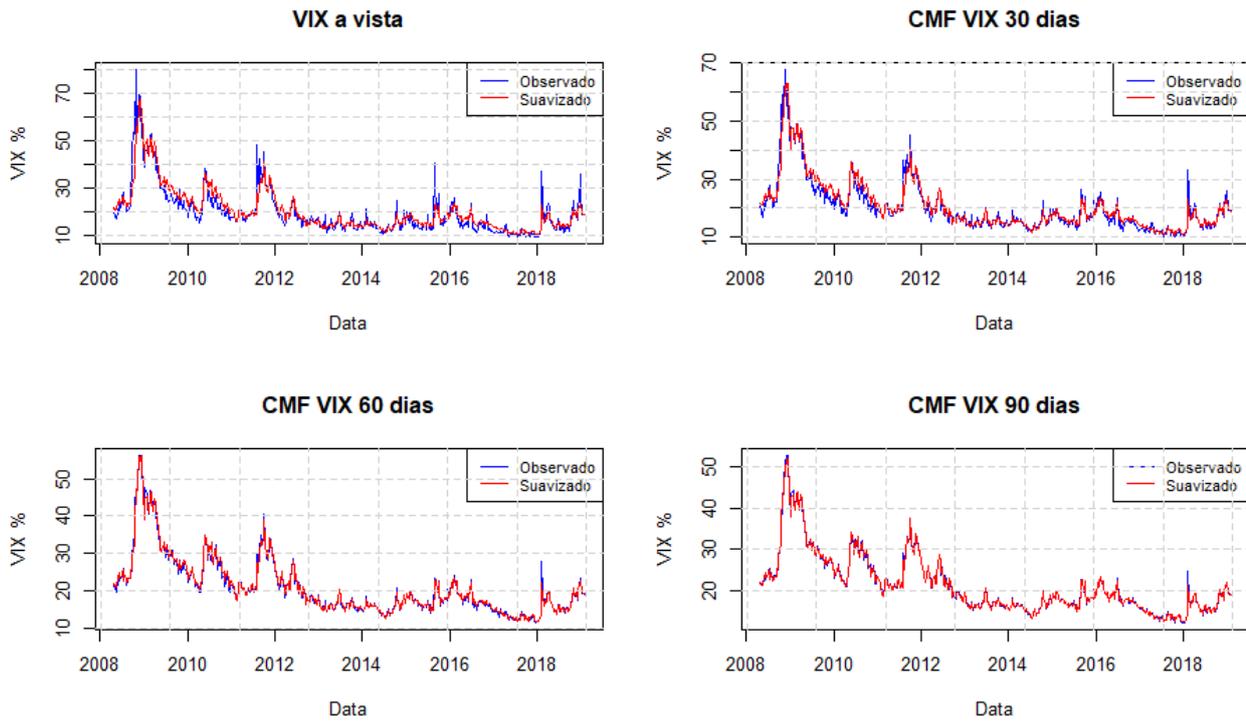


Figura C.10: Comparativo das estimativas do VIX à vista e 3 CMFs (amostra sem o VIX à vista)

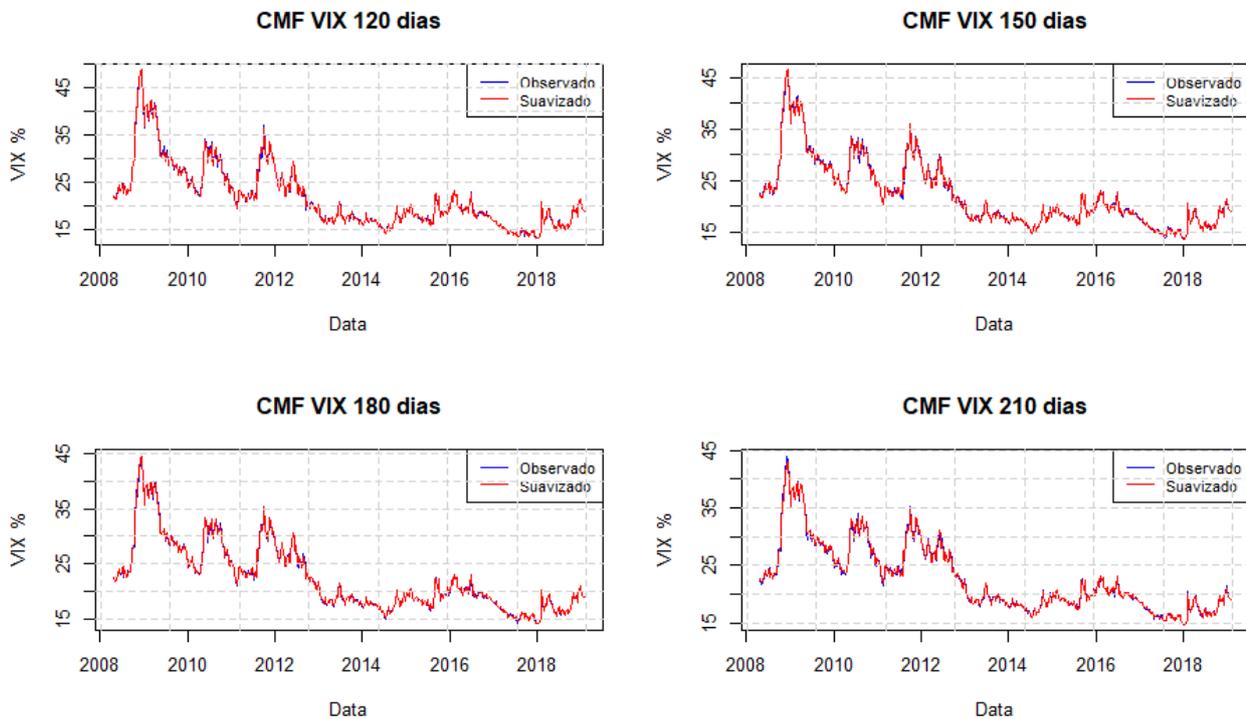


Figura C.11: Comparativo das estimativas de 4 CMFs (amostra sem o VIX à vista)

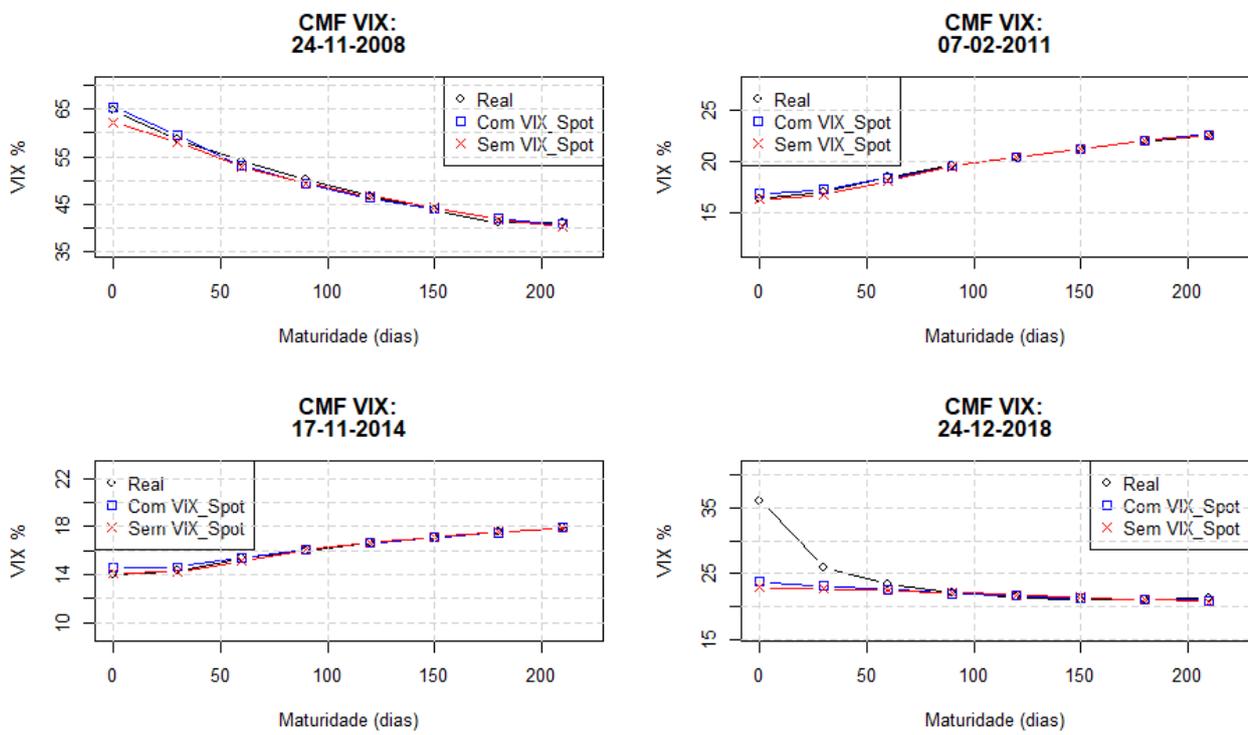


Figura C.12: Comparativo da estrutura a termo de dias escolhidos (amostras com e sem o VIX à vista)