



Instituto de Matemática Pura e Aplicada

Mestrado em Métodos Matemáticos em Finanças

Calibração do Modelo de Schwartz-Smith com Filtro de Kalman

Leonardo Lima da Silva Marotta

Orientador: Roberto Imbuzeiro Oliveira

Coorientador: Jorge Passamani Zubelli

IMPA
Fevereiro de 2011



Instituto de Matemática Pura e Aplicada

Mestrado em Métodos Matemáticos em Finanças

Calibração do Modelo de Schwartz-Smith com Filtro de Kalman

Autor:

Leonardo Lima da Silva Marotta

Orientador:

Roberto Imbuzeiro Oliveira

Coorientador:

Jorge Passamani Zubelli

Examinador:

Max Oliveira Souza

IMPA
Fevereiro de 2011

*Dedico este trabalho a minha esposa Michelle
e aos meus pais, Jorge e Lúcia.*

Agradecimentos

Gostaria de agradecer ...

a Deus por me guiar no caminho certo.

à minha querida e amada esposa Michelle, por ter me dado incondicional apoio em todos os momentos difíceis.

aos meus pais Jorge e Lúcia por todos os conselhos e carinho que tiveram comigo durante todo o tempo.

ao professor Zubelli pela paciência, pelas incontáveis dicas e conselhos para o término deste trabalho.

ao professor Roberto Imbuzeiro pelas correções e orientação deste trabalho.

aos meus colegas de turma Maristela, Rafael, Daniel, Arthur, Fábio e Henrique por terem compartilhado comigo as alegrias e dificuldades de conciliar trabalho, estudo e família.

Resumo

Este trabalho coloca em prática o modelo de Schwartz-Smith para modelagem de preços de commodities e utiliza o filtro de Kalman para obter as componentes não-observáveis em questão. O processo de otimização utilizado visa maximizar a função de verossimilhança extraída através do método do filtro de Kalman e assim obter os parâmetros ótimos que governam a dinâmica da commodity em estudo. Os dados utilizados são de futuros de Gás Natural, mais especificamente *Henry-Hub*, que por sua vez após a calibração do modelo, seu preço *Spot* é utilizado para valorar um opção de adiamento de investimento segundo a teoria de Opções Reais.

Palavras-Chave: Schwartz-Smith, Filtro de Kalman, *Henry-Hub*, Opções Reais.

Abstract

This work puts into practice the Schwartz-Smith model for commodity valuation and uses the Kalman filter method to extract the unobservable components in the model. The optimization process used tries to maximize the likelihood function extracted from the Kalman filter and thereby obtain the optimal parameters that govern the dynamics of the commodity under study. The data used are futures of Natural Gas, from Henry Hub. The model, the spot price is used to value an option to defer investment according to the theory of Real Options.

Keywords: Schwartz-Smith, Kalman Filter, *Henry-Hub*, Real Options.

Sumário

Lista de Figuras	vi
Lista de Tabelas	vii
1 Introdução	1
1.1 Mercado de Gás Natural	2
1.2 Estrutura da Dissertação	4
2 Revisão Bibliográfica	5
3 Revisão Teórica	12
3.1 Probabilidade e Cálculo Estocástico	12
3.2 Apreçamento Neutro ao Risco	16
3.3 Método de Euler-Maruyama	19
3.3.1 Convergência do Método de Euler-Maruyama	20
3.4 Filtro de Kalman	22
4 Modelo de Schwartz-Smith	26
4.1 Modelo na Medida Física	26
4.2 Modelo na Medida Neutra ao Risco	30
4.3 Valorando Contratos Futuros	32
5 Resultados	33

5.1	Pré-Processamento dos Dados	33
5.2	Simulação Henry-Hub	35
5.2.1	Simulação dos Processos χ_t e ξ_t	36
5.2.2	Simulação dos Futuros	37
5.3	Filtragem de Kalman	38
5.4	Calibragem do Modelo	43
5.5	Aplicação para Opções Reais	45
6	Conclusão	49
	Referências	51
	Apêndice A – Distribuição Log-Normal	53
	Apêndice B – Distribuição conjunta de χ_t e ξ_t	54
	Apêndice C – Figuras	56
	Anexo A – Códigos	59
A.1	Arquivo simChiXi2.m	59
A.2	Arquivo simSSfutPrices.m	59
A.3	Arquivo inicialsimSSfutPrices2.m	60
A.4	Arquivo testKalman.m	61
A.5	Arquivo logLikeGrad.m	62
A.6	Arquivo optimizeParamGrad.m	64
A.7	Arquivo SSSpotSim.m	66
A.8	Arquivo AmOptLSM.m	66

Lista de Figuras

1	Cadeia de Suprimento do LNG	3
2	Formato dos dados antes da rolagem	34
3	Formato dos dados depois da rolagem	34
4	Dados do Henry Hub após o pré-processamento	35
5	Tempo ao Vencimento (τ)	36
6	Realização dos processos χ_t e ξ_t com 1000 amostras	37
7	Realização dos processos χ_t e ξ_t com 10000 amostras	38
8	Superfície de $A(\tau)$	39
9	Preços Simulados para 24 maturidades e 1200 amostras	39
10	Preços Simulados vs Preços pós Filtragem de Kalman	41
11	Valor da Opção vs Taxa de Juros	47
12	Valor da Opção vs Valor do Investimento	48
13	Média do Processo χ_t com 1000 simulações e 10000 amostras	56
14	Média do Processo ξ_t com 1000 simulações e 10000 amostras	56
15	Variância do Processo χ_t com 1000 simulações e 10000 amostras	57
16	Variância do Processo ξ_t com 1000 simulações e 10000 amostras	57
17	Covariância dos Processos χ_t e ξ_t com 1000 simulações e 10000 amostras	58

Lista de Tabelas

1	Parâmetros para Simulação	37
2	Testes para o Filtro de Kalman	42
3	Resultado da Otimização para Dados Simulados	44
4	Resultado da Otimização para Dados Reais	45
5	Quadro comparativo entre opções reais e financeiras	46

1 Introdução

Um dos campos mais explorados em finanças na atualidade é a valoração de ativos reais. Muitas vezes utiliza-se os preços futuros de alguma commodity para precificá-los. A busca por modelos cada vez mais sofisticados faz com que a pesquisa para tal fim esteja na fronteira do conhecimento matemático aplicado às finanças. Sendo assim, é de suma importância entender o comportamento das commodities.

Diversos autores discutem as principais diferenças entre um ativo comum, como uma ação, e uma commodity (ver [1] e [2]). Em primeiro lugar, podemos notar que commodities, em geral, não são negociadas no mercado à vista fazendo com que seu preço *spot* não seja observável. No entanto é muito comum encontrar na literatura que os contratos futuros com vencimento mais próximo são uma *proxy* para o preço *spot*.

As commodities por serem bens fisicamente negociados apresentam o que chamamos de custo de carregamento (*cost of carry*) que pode estar relacionado à despesas financeiras, custos de estocagem, seguros, custos de obtenção de crédito, juros e impostos sobre propriedades. Sendo assim, é importante também entender o papel do retorno de conveniência (*convenience yield*) que nada mais é do que o prêmio cobrado pelo detentor da commodity física ao detentor do contrato futuro.

Outro ponto muito discutido é a característica de reversão à média dos preços das commodities. Isto pode ser explicado intuitivamente com conceitos básicos de oferta e demanda. Vamos supor que os preços de uma determinada commodity estejam maiores do que algum preço de equilíbrio hipotético. Então o suprimento desta commodity irá aumentar, novos produtores começam a produzir e produtores com custos maiores não deixam de operar e em algum momento este suprimento em excesso gerará uma pressão de queda nos preços. Por outro lado, quando estes

preços caem os produtores com custos mais elevados tendem a sair do mercado o que diminuirá o suprimento desta commodity e portanto gerará uma pressão de alta nos preços. É razoável supor que estas entradas e saídas de produtores levem um certo tempo, ou seja, os preços devem operar em patamares relativamente baixos ou relativamente altos por algum tempo em relação a um preço de equilíbrio e sempre revertendo a este equilíbrio. Outro ponto interessante é a mudança de patamares de produção em decorrência de fenômenos naturais que podem afetar significativamente a produção de uma certa commodity em uma determinada região. Um exemplo disso são as variações climáticas ocorridas nos últimos tempos, como consequência do El Niño e do efeito estufa gerando impacto na produção de determinadas commodities agrícolas. Vale a pena ressaltar que a evolução tecnológica pode aumentar o patamar de produção de uma determinada commodity: podemos citar o álcool de segunda geração produzido pela cana-de-açúcar que ao chegar em escala industrial elevará radicalmente os patamares de produção. Pelo lado da demanda, podemos citar o enriquecimento de países emergentes, tais como a China que se tornou uma grande consumidora de commodities. Estes são apenas alguns exemplos de que no longo prazo é possível enxergarmos uma mudança nos preços de equilíbrio de uma determinada commodity.

Este trabalho visa implementar o modelo de Schwartz-Smith [3] para preços futuros de gás natural. Consideramos aqui especificamente *Henry-Hub*, porém a metodologia é aplicável a outros problemas. Esta implementação é feita com o método clássico do filtro de Kalman que relaciona as variáveis de estado não observáveis com os preços futuros de diversas maturidades. Após isso é gerada uma função de verossimilhança que posteriormente é maximizada através de um processo de otimização, chegando-se nos parâmetros ótimos.

1.1 Mercado de Gás Natural

Os maiores consumidores de gás natural estão basicamente divididos em três grandes blocos, Estados Unidos, Europa e Japão.

Praticamente todo gás natural consumido nos EUA é produzido internamente, grande parte em uma região chamada *Henry-Hub* localizado no estado da

Louisiana. É neste local onde está construído um dos maiores hubs de dutos de gás natural do mundo que é operado pela Sabine Pipe Line LLC. Este local também é o ponto de precificação dos contratos futuros de gás natural negociados no NYMEX.

Para ter a intuição da precificação do gás natural é importante entender como estes grandes mercados consumidores estão conectados e como os preços de um mercado afetam o outro. Basicamente o gás natural pode ser transportado de duas formas diferentes, através de dutos e de navios carregadores de gás natural liquefeito (GNL ou LNG em inglês).

Em geral os dutos conectam estados de um determinado país ou até partes de um país próximo, como é o caso de alguns dutos do *Henry-Hub* que se conectam ao México. Para maiores distâncias a única alternativa disponível hoje é o transporte marítimo. Vamos ver brevemente como funciona a cadeia de suprimento do LNG.

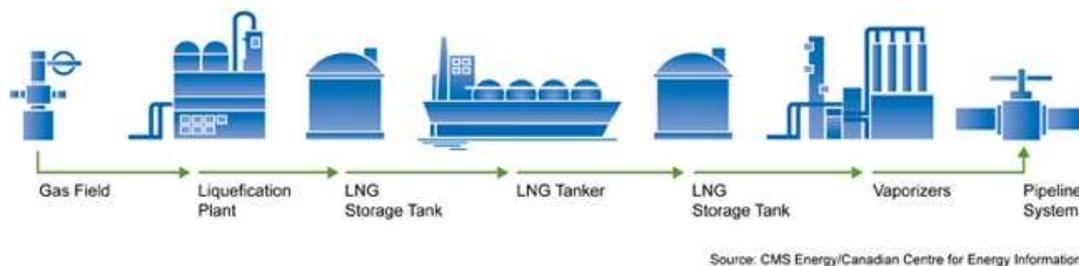


Figura 1: Cadeia de Suprimento do LNG

Como podemos ver na Figura 1 um determinado produtor localizado em um país distante produz o gás natural em um determinado campo on-shore ou off-shore, o transporta para uma planta de liquefação, armazena este gás que foi liquefeito, carrega o navio com LNG, o transporta até o país de destino, armazena no tanque de destino, passa por uma planta de gaseificação e entra no sistema interno de dutos. Uma observação importante é que um navio de LNG é capaz de diminuir o volume a quase um milésimo de seu volume original; este é um dos motivos pelo qual há uma razão econômica no transporte deste LNG. Outro ponto importante é que ao entrar no sistema de dutos do país de destino o gás natural tem o mesmo valor que o gás natural produzido internamente. Disto podemos concluir que para o explorador de um determinado campo de gás natural só valerá a pena vender para outros países se a soma do seu custo mais a soma de todos os custos da cadeia até o destino sejam menores que o preço ofertado pelo gás natural no destino. Está é uma das

peculiaridades mais importantes do mercado de gás natural.

Com o avanço da tecnologia nas plantas de liquefação e com o aumento da capacidade de transporte de LNG tornou-se viável economicamente levar esta commodity para vários destinos, principalmente para aqueles países que não possuem auto-suficiência na produção e para aqueles que vem mudando sua matriz energética para fontes mais limpas, comparada ao carvão e ao petróleo. Um bom exemplo disso é o Japão, que importa a maior parte do gás natural que consome.

Esta complexa cadeia deu incentivo a criação de novos instrumentos financeiros tais como o *cancelation option*, no qual o comprador tem o direito de cancelar o contrato de fornecimento de gás natural em datas pré-determinadas. Uma boa referência sobre o tema pode ser encontrada em [4].

Visto isso como motivação podemos começar os estudos para calibrar o modelo Schwartz-Smith.

1.2 Estrutura da Dissertação

Esta seção descreve muito resumidamente o que cada capítulo a seguir aborda.

O Capítulo 2 faz uma revisão de todos os trabalhos relacionados a este tema dos quais informações importantes foram extraídas. São citados trabalhos de diversos autores com modelos e metodologias diferentes para a identificação da dinâmica de preços de commodities.

O Capítulo 3 faz uma revisão geral da teoria e das ferramentas utilizadas nesta dissertação, passando por cálculo estocástico, apreçamento neutro ao risco, método de Euler-Maruyama para discretização dos modelos e filtragem de Kalman.

O Capítulo 4 entra no detalhe do modelo de Schwartz-Smith, onde os principais resultados são demonstrados.

O Capítulo 5 mostra os principais resultados na simulação dos processos não observáveis, dos preços futuros de commodities, da filtragem de Kalman e da calibração do modelo de Schwartz-Smith.

Por fim o Capítulo 6 descreve as principais conclusões desta dissertação.

2 Revisão Bibliográfica

Esta seção destina-se a fazer um apanhado geral de todos os artigos, teses e reports relevantes para esta dissertação.

Eduardo Schwartz em [5] aborda três modelos a respeito do comportamento estocástico dos preços das commodities levando em consideração reversão à média. O primeiro modelo é o mais simples de todos. Neste é considerado somente um fator, no qual o logaritmo do preço *spot* segue um processo de reversão à média. O segundo modelo leva em consideração um fator adicional, o *covenience yield* da commodity, no qual também é assumido que este segue um processo de reversão à média. O terceiro e último modelo leva em consideração além dos dois fatores supracitados, a estocasticidade da taxa de juros. Vale lembrar que os modelos apresentados são apenas uma das possibilidades de modelos de um, dois e três fatores.

O artigo também trata do filtro de Kalman para relacionar as variáveis observáveis com as não observáveis e para a calibração dos três modelos.

O Modelo de Um Fator (MUF) assume que o preço *spot* da commodity segue o processo estocástico descrito pela dinâmica.

$$dS = \kappa(\mu - \ln S)Sdt + \sigma SdZ \quad (2.1)$$

Definindo $X = \ln S$ e aplicando o lema de Itô chegamos a um processo de Ornstein-Uhlenbeck.

$$dX = \kappa(\alpha - X)dt + \sigma dZ \quad (2.2)$$

onde

$$\alpha = \mu - \frac{\sigma^2}{2\kappa} \quad (2.3)$$

O termo κ mede o grau de reversão à média para o logaritmo do preço *spot* de longo prazo, α . O termo σ caracteriza a volatilidade do processo e o dZ é a diferencial de um movimento Browniano padrão.

Passando a medida neutra ao risco [6] temos

$$dX = \kappa(\alpha^* - X)dt + \sigma dZ^* \quad (2.4)$$

onde $\alpha^* = \alpha - \lambda$, λ é o prêmio de risco de mercado. Neste artigo o prêmio de risco λ é suposto constante. Além disso, denota dz^* é o incremento do movimento Browniano na medida neutra ao risco.

Da equação (2.4) podemos extrair os seguintes resultados

$$E_0[X(T)] = e^{-\kappa T} X(0) + (1 - e^{-\kappa T})\alpha^* \quad (2.5)$$

$$Var_0[X(T)] = \frac{\sigma^2}{2\kappa}(1 - e^{-\kappa T}) \quad (2.6)$$

Como $X = \ln S$, o preço *spot* da commodity no tempo T possui uma distribuição log-normal.

Assumindo a taxa de juros constante temos que o preço do futuro da commodity com maturidade T é o valor esperado do preço da commodity no tempo T sob a medida neutra ao risco. Então, temos pela Equação (2.7) o valor do futuro da commodity no tempo T .

$$\begin{aligned} F(S, T) &= E[S(T)] = \exp(E_0[X(T)] + 1/2Var_0[X(T)]) \\ &= \exp \left[e^{-\kappa T} \ln S + (1 - e^{-\kappa T})\alpha^* + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa T}) \right] \end{aligned} \quad (2.7)$$

O lado direito da Equação (2.7) satisfaz a equação diferencial parcial

$$1/2\sigma^2 S^2 F_{SS} + \kappa(\mu - \lambda - \ln S)SF_S - F_T = 0 \quad (2.8)$$

com condição de contorno terminal $F(S, 0) = S$.

O Modelo de Dois Fatores (MDF) considera o preço *spot* da commodity (fator 1: S) e o *convenience yield* instantâneo (fator 2: δ). Esses dois fatores seguem o processo estocástico descrito nas Equações (2.9) e (2.10).

$$dS = (\mu - \delta)Sdt + \sigma_1 S dZ_1 \quad (2.9)$$

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dZ_2 \quad (2.10)$$

onde os incrementos do movimento Browniano são correlacionados segundo a Equação (2.11).

$$dZ_1 dZ_2 = \rho dt \quad (2.11)$$

Note que se o valor de δ fosse determinístico, seguindo a relação $\delta(S) = \kappa \ln S$, teríamos então que o Modelo de Dois Fatores se reduziria ao Modelo de um Fator.

Mais uma vez fazendo $X = \ln S$ chegamos a Equação (2.12).

$$dX = (\mu - \delta - 1/2\sigma_1^2)dt + \sigma_1 dZ_1 \quad (2.12)$$

Neste modelo a commodity é tratada como um ativo que paga um dividendo estocástico δ . Assim, o termo de *drift* ajustado ao risco será $r - \delta$. Novamente é considerado um prêmio de risco de mercado associado λ que é constante. O processo estocástico para os fatores deste modelo sob a medida martingal equivalente pode ser expresso como nas Equações (2.13), (2.14) e (2.15).

$$dS = (r - \delta)Sdt + \sigma_1 S dZ_1^* \quad (2.13)$$

$$d\delta = [\kappa(\alpha - \delta) - \lambda]dt + \sigma_2 dZ_2^* \quad (2.14)$$

$$dZ_1^* dZ_2^* = \rho dt \quad (2.15)$$

Os preços futuros satisfazem a seguinte equação diferencial parcial com condição inicial em $T = 0$ $F(S, \delta, 0) = S$

$$1/2\sigma_1^2 S^2 F_{SS} + \sigma_1 \sigma_2 \rho S F_{S\delta} + 1/2\sigma_2^2 F_{\delta\delta} + (r - \delta) S F_S + (\kappa(\alpha - \delta) - \lambda) F_\delta - F_T = 0 \quad (2.16)$$

É possível mostrar que a solução da Equação (2.16) com condição inicial em $T = 0$ é da forma da Equação (2.17).

$$F(S, \delta, T) = S \exp \left[-\delta \frac{1 - e^{-\kappa T}}{\kappa} + A(T) \right] \quad (2.17)$$

onde

$$A(T) = \left(r - \hat{\alpha} + \frac{1}{2} \frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1 \sigma_2 \rho}{\kappa} \right) T + \frac{1}{4} \sigma_2^2 \frac{1 - e^{-2\kappa T}}{\kappa^3} + \left(\hat{\alpha} \kappa + \sigma_1 \sigma_2 \rho - \frac{\sigma_2^2}{\kappa} \right) \frac{1 - e^{-\kappa T}}{\kappa^2} \quad (2.18)$$

e

$$\hat{\alpha} = \alpha - \frac{\lambda}{\kappa} \quad (2.19)$$

O Modelo de Três Fatores (MTF) considera as variáveis de estado como sendo o preço *spot*, o *convenience yield* instantâneo e a taxa de juros instantânea. Assumindo que a taxa de juros segue um processo de Ornstein-Uhlenbeck, como no modelo de Vasicek, podemos estender o modelo de dois fatores para o de três fatores. Utilizando o processo estocástico conjunto descrito nas Equações (2.13) e (2.14) e incluindo a taxa de juros estocástica teremos as Equações (2.20), (2.21), (2.22) e (2.23) que governam o MTF.

$$dS = (r - \delta)Sdt + \sigma_1 S dZ_1^* \quad (2.20)$$

$$d\delta = \kappa(\hat{\alpha} - \delta)dt + \sigma_2 dZ_2^* \quad (2.21)$$

$$dr = \alpha(m^* - r)dt + \sigma_3 dZ_3^* \quad (2.22)$$

$$\begin{aligned} dZ_1^* dZ_2^* &= \rho_1 dt, \\ dZ_2^* dZ_3^* &= \rho_2 dt, \\ dZ_1^* dZ_3^* &= \rho_3 dt \end{aligned} \quad (2.23)$$

Onde α e m^* são, respectivamente, a velocidade de reversão à média da taxa de juros e a taxa de juros de curto prazo ajustada ao risco. Os preços futuros devem obedecer a equação diferencial parcial (2.24) conforme visto nos modelos anteriores.

$$\begin{aligned} 1/2\sigma_1^2 S^2 F_{SS} + 1/2\sigma_2^2 F_{\delta\delta} + 1/2\sigma_3^2 F_{rr} + \sigma_1\sigma_2\rho_1 F_{\delta r} + \sigma_1\sigma_3\rho_3 S F_{Sr} \\ + (r - \delta)S F_S + \kappa(\hat{\alpha} - \delta)F_\delta + \alpha(m^* - r)F_r - F_T = 0 \end{aligned} \quad (2.24)$$

com condição de contorno inicial $F(S, \delta, r, 0) = S$. A solução da Equação (2.24) esta na forma da Equação (2.25).

$$F(S, \delta, r, T) = S \exp \left[\frac{-\delta(1 - e^{-\kappa T})}{\kappa} + \frac{r(1 - e^{-\alpha T})}{\alpha} + C(T) \right] \quad (2.25)$$

onde $C(T)$ é mostrado em [5].

Em resumo o artigo implementa três modelos que levam em conta a natureza de reversão à média dos preços das commodities de formas diferentes. O autor encontra dificuldades em modelar preços futuros com maturidades mais longas, pois na ocasião em que o artigo foi escrito ainda não existiam dados públicos disponíveis com maturidades mais longas. Uma observação interessante é que a volatilidade do retorno das commodities modeladas por apenas um fator converge para zero

em maturidades muito longas, o que não acontece para os modelos com dois e três fatores no qual a volatilidade assume um valor estável para maturidades longas diferente de zero. Os resultados em [5] mostraram que o Modelo de Um Fator não apresentou uma boa aderência para os preços das commodities avaliadas. O Modelo de Dois Fatores apresentou resultados muito melhores com boa aderência aos preços das commodities. O Modelo com taxa de juros estocástica apresentou resultados semelhantes ao MDF.

Um segundo trabalho importante foi escrito por Eduardo Schwartz e James Smith [3], no qual é apresentado o bem difundido modelo de Schwartz-Smith. Nesta dissertação o Capítulo 4 trata especificamente deste modelo, portanto apenas uma breve introdução será feita.

Este artigo desenvolve um modelo de dois fatores, no qual um deles tenta explicar a reversão à média de curto prazo dos preços. O outro fator tenta explicar o preço de equilíbrio de longo prazo das commodities. Apesar destes dois fatores não serem observáveis, eles podem ser estimados através dos preços *spot* e preços dos futuros da commodity em estudo com a implementação do filtro de Kalman. Um ponto muito interessante é que este modelo não utiliza o *convenience yield* como fator a ser modelado. Ainda é mostrado que há uma equivalência entre o modelo de Schwartz-Smith e o chamado modelo Gibson-Schwartz (adaptação do Modelo de Dois Fatores discutido no artigo anterior).

Outro trabalho muito interessante foi escrito por Manoliu e Tompaidis [7]. Ele estuda uma generalização da idéia do artigo anterior e mostra um modelo multifator para modelar futuros ligados à energia. Esta formulação depois é reduzida para um e dois fatores similares ao visto em [5] e [3] no qual é modelado preços de gás natural. Os autores também utilizam o filtro de Kalman em conjunto com estimação por máxima verossimilhança. Mais uma vez nos resultados, o modelo de dois fatores obtém resultados melhores do que o modelo de um fator. Além disso os autores incluem uma sazonalidade anual na modelagem dos preços.

A dissertação desenvolvida por Ana Luiza Carvalho [8] propõe uma abordagem diferente para calibração do modelo de Schwartz-Smith. Neste trabalho foi feita a maximização da função de verossimilhança e mínimos quadrados não lineares para

a calibração do modelo ao invés de decomposição do erro de previsão com filtro de Kalman. O método consiste basicamente em dois passos. No primeiro passo tenta encontrar os processos não observáveis através de mínimos quadrados não-lineares. O segundo passo utiliza os processos encontrados no passo anterior para gerar novas estimativas dos parâmetros via maximização da função de verossimilhança. O processo pára quando atinge o número máximo de iterações ou quando a condição de convergência é atingida. O método foi aplicado para dados artificiais obtendo bons resultados. A robustez foi testada iniciando o problema com parâmetros diferentes através de uma perturbação crescente nos mesmos.

A tese de doutorado de Fernando Aiube [9] compara a utilização do método do filtro de Kalman linear para a estimação dos parâmetros com o filtro de partículas, cuja principal vantagem esta calcada na não linearidade e não gaussianidade. Além disso o autor estende o modelo de Schwartz-Smith com a utilização de saltos seguindo um processo gaussiano e de Poisson. Este incremento possibilita capturar o excesso de curtose apresentado em retornos financeiros. O filtro de partículas apresenta resultados ligeiramente inferiores aos do filtro de Kalman, e ainda com tempo computacional superior.

3 Revisão Teórica

Este capítulo destina-se a apresentar os principais resultados teóricos e as ferramentas necessárias para a implementação do modelo de Schwartz-Smith e baseia-se nas referências [10] e [6].

Basicamente serão apresentadas definições e teoremas úteis para o desenvolvimento deste trabalho.

3.1 Probabilidade e Cálculo Estocástico

Nesta seção serão apresentados alguns resultados importantes de probabilidade e cálculo estocástico.

Definição 3.1.1 (Espaço de Probabilidade). *Um espaço de probabilidade é uma tripla $(\Omega, \mathcal{F}, \mathbb{P})$, no qual temos:*

- Ω é um conjunto não-vazio, chamado de espaço amostral;
- \mathcal{F} é uma σ -álgebra de subconjuntos de Ω ;
- \mathbb{P} é uma medida de probabilidade.

Definição 3.1.2 (Filtração). *Seja Ω um conjunto não-vazio. Seja T fixo, um número positivo, e assumamos que para cada $t \in [0, T]$ existe uma σ -álgebra $\mathcal{F}(t)$. Assumamos também que se $s < t$, então todo conjunto em $\mathcal{F}(s)$ está também em $\mathcal{F}(t)$. Então chamamos a coleção de σ -álgebras $\mathcal{F}(t)$, $0 \leq t \leq T$, de filtração. Além disso, as σ -álgebras da filtração devem estar contidas em \mathcal{F} .*

Definição 3.1.3 (Processo Adaptado). *Seja Ω um espaço amostral não vazio e uma filtração $\mathcal{F}(t)$, $0 \leq t \leq T$. Seja $X(t)$ uma coleção de variáveis aleatórias indexadas*

por $t \in [0, T]$. Falamos que $X(t)$ é um processo adaptado se para cada t , a variável aleatória $X(t)$ é $\mathcal{F}(t)$ -mensurável.

Definição 3.1.4 (Martingale). *Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade, seja T fixo e positivo, e seja $\mathcal{F}(t)$, $0 \leq t \leq T$, uma filtração de sub- σ -álgebras de \mathcal{F} . Considere o processo estocástico adaptado $M(t)$, $0 \leq t \leq T$, tal que $\mathbb{E}[|M(t)|] < \infty$, $\forall t \in [0, T]$.*

- (a) *Se $\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s)$ para todo $0 \leq s \leq t \leq T$, dizemos que este processo é um martingale,*
- (b) *Se $\mathbb{E}[M(t)|\mathcal{F}(s)] \geq M(s)$ para todo $0 \leq s \leq t \leq T$, dizemos que este processo é um submartingale,*
- (c) *Se $\mathbb{E}[M(t)|\mathcal{F}(s)] \leq M(s)$ para todo $0 \leq s \leq t \leq T$, dizemos que este processo é um supermartingale.*

Definição 3.1.5 (Variação Quadrática). *Seja $f(t)$ uma função definida para $0 \leq t \leq T$. A variação quadrática de f até o tempo T é dada por*

$$[f, f](T) = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2,$$

onde $\Pi = t_0, t_1, \dots, t_n$ e $0 = t_0 < t_1 < \dots < t_n = T$.

Definição 3.1.6 (Movimento Browniano). *Seja $(\Omega, \mathcal{F}, \mathbb{P})$ o espaço de probabilidade. Para cada $\omega \in \Omega$, suponha que existe uma função contínua $W(t)$ para $t \geq 0$ que satisfaça $W(0) = 0$ e que dependa de ω . Então $W(t)$, $t \geq 0$, é um Movimento Browniano se para todo $0 = t_0 < t_1 < \dots < t_m$ os incrementos*

$$W(t_1) - W(t_0), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$$

são independentes e cada um desses incrementos sejam normalmente distribuídos com

$$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0$$

$$\text{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i$$

Definição 3.1.7 (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 σ -álgebras $\mathcal{F}(t)$, $t \geq 0$, que satisfaz:*

- (a) **(Acumula Informações)** *Para $0 \leq s < t$, todo conjunto em $\mathcal{F}(s)$ está também em $\mathcal{F}(t)$. Em outras palavras, existe no mínimo tanta informação disponível em um tempo posterior $\mathcal{F}(t)$ quanto existe em um tempo anterior $\mathcal{F}(s)$;*
- (b) **(Adaptatividade)** *Para cada $t \geq 0$, o movimento Browniano $W(t)$ no tempo t é $\mathcal{F}(t)$ -mensurável. Em outras palavras, a informação disponível no tempo t é suficiente para avaliar o movimento Browniano $W(t)$ neste tempo;*
- (c) **(Independência de Incrementos Futuros)** *Para $0 \leq t < u$, todo incremento $W(u) - W(t)$ é independente de $\mathcal{F}(t)$. Em outras palavras, qualquer incremento do movimento Browniano após o tempo t é independente da informação disponível no tempo t .*

Teorema 3.1.1. *O movimento Browniano é um martingale*

Demonstração. Seja $0 \leq s \leq t$ dados. Então

$$\begin{aligned} \mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[(W(t) - W(s)) + W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)|\mathcal{F}(s)] + \mathbb{E}[W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)] + W(s) \\ &= W(s) \end{aligned}$$

□

Definição 3.1.8 (Integral de Itô para processos simples). *Seja $\Pi = t_0, t_1, \dots, t_n$ uma partição de $[0, T]$, i.e., $0 = t_0 < t_1 < \dots < t_n = T$. Seja $\Delta(t)$ um processo simples, onde $\Delta(t)$ é constante em cada subintervalo $[t_j, t_{j+1})$. Vamos agora definir a Integral de Itô como segue abaixo:*

$$I(t) = \int_t^0 \Delta(u) dW(u) = \sum_{j=0}^{k-1} \Delta(t_j) [W(t_{j+1}) - W(t_j)] + \Delta(t_k) [W(t) - W(t_k)].$$

Definição 3.1.9 (Integral de Itô Generalizada). *Assumindo que $\Delta(t)$, $t \geq 0$, é adaptado a filtração $\mathcal{F}(t)$, $t \geq 0$ e a condição de integrabilidade*

$$\mathbb{E} \int_0^T \Delta^2(t) dt < \infty$$

vamos escolher uma sequência¹ $\Delta_n(t)$ de processos simples tal que quando $n \rightarrow \infty$ esse processo convirja para $\Delta(t)$. Por convergência entende-se que

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

Definimos então a Integral de Itô para o integrando com variações contínuas pela fórmula

$$\int_0^t \Delta(u) dW(u) = \lim_{n \rightarrow \infty} \int_0^t \Delta_n(u) dW(u), \quad 0 \leq t \leq T.$$

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

Teorema 3.1.2. *Seja T uma constante positiva e seja $\Delta(t)$, $0 \leq t \leq T$, um processo estocástico adaptado que satisfaça 3.1.*

$$\mathbb{E} \left[\int_0^T \Delta^2(t) dt \right] < \infty \tag{3.1}$$

Então $I(t) = \int_0^t \Delta(u) dW(u)$ possui as seguintes propriedades:

- (a) **(Continuidade)** *Os caminhos de $I(t)$ são contínuos;*
- (b) **(Adaptatividade)** *Para cada t , $I(t)$ é $\mathcal{F}(t)$ -mensurável;*
- (c) **(Linearidade)** *Se $I(t) = \int_0^t \Delta(u) dW(u)$ e $J(t) = \int_0^t \Gamma(u) dW(u)$, então $I(t) \pm J(t) = \int_0^t (\Delta(u) \pm \Gamma(u)) dW(u)$; e também, para qualquer constante c , $cI(t) = \int_0^t c\Delta(u) dW(u)$;*
- (d) **(Martingale)** *$I(t)$ é um martingale contínuo;*

¹A existência de tal sequência pode ser demonstrada. Veja [6]

(e) (**Isometria de Itô**) $\mathbb{E}[I^2(t)] = \mathbb{E}[\int_0^t \Delta^2(u)du]$;

(f) (**Varição Quadrática**) $[I, I](t) = \int_0^t \Delta^2(u)du$.

Demonstração. Veja [10] Capítulo 4, Seção 3. □

Teorema 3.1.3 (Fórmula de Itô). *Seja $X(t)$, $t \geq 0$, um processo de Itô e seja $f(t, x)$ uma função para a qual as derivadas parciais $f_t(t, x)$, $f_x(t, x)$ e $f_{xx}(t, x)$ estão definidas e são contínuas. Então, para todo $T \geq 0$,*

$$f(T, X(T)) = f(0, X(0)) + \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))dX(t) + \frac{1}{2} \int_0^T f_{xx}(t, X(t))d[X, X](t)$$

ou na forma diferencial

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))dX(t)dX(t)$$

Demonstração. Veja [6], Capítulo 2, Seção 3 □

Teorema 3.1.4 (Integral de Itô com Integrandos Determinísticos). *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 $\int_0^t \Delta^2 ds$.*

Demonstração. Veja [10], Capítulo 4, Seção 4 □

3.2 Aprecamento Neutro ao Risco

Nesta seção serão enunciados os principais resultados de aprecamento neutro ao risco utilizados no desenvolvimento deste trabalho.

Definição 3.2.1 (Derivada de Radon-Nokodým). *Seja o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e uma variável aleatória não-negativa Z que satisfaz $\mathbb{E}[Z] = 1$. Podemos definir uma nova medida de probabilidade $\tilde{\mathbb{P}}$ pela fórmula*

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

Dizemos que Z é a derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ com respeito a \mathbb{P} , e escrevemos como

$$Z = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}$$

Definição 3.2.2 (Medida Martingale Equivalente). *Duas medidas \mathbb{P} e $\tilde{\mathbb{P}}$ são ditas equivalentes se ambas possuem os mesmos conjuntos com medida zero. Se além disso, o processo de preços descontados for um martingale na medida $\tilde{\mathbb{P}}$ dizemos que $\tilde{\mathbb{P}}$ é uma medida martingale equivalente a \mathbb{P}*

Teorema 3.2.1 (Girsanov). *Seja $W(t)$, $0 \leq t \leq T$, um movimento Browniano no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, e seja \mathcal{F} , $0 \leq t \leq T$, a filtração Browniana. Seja $\Theta(t)$, $0 \leq t \leq T$, um processo adaptado a filtração Browniana. Defina*

$$Z(t) = \exp \left\{ - \int_0^t \Theta(u) dW(u) - \frac{1}{2} \int_0^t \Theta^2(u) d(u) \right\}$$

$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u) d(u)$$

e assumamos que

$$\mathbb{E} \left[\int_0^T \Theta^2(u) Z^2(u) du < \infty \right]$$

Faça $Z = Z(T)$. Então $\mathbb{E}[Z] = 1$ e na medida de probabilidade $\tilde{\mathbb{P}}$, o processo $\tilde{W}(t)$, $0 \leq t \leq T$, é um movimento Browniano.

Demonstração. Veja [10], capítulo 5, seção 2. □

Teorema 3.2.2 (Teorema de Representação Martingal). *Seja $W(t)$, $0 \leq t \leq T$, um movimento Browniano em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, e seja $\mathcal{F}(t)$, $0 \leq t \leq T$, a filtração gerada por este movimento Browniano. Seja $M(t)$, $0 \leq t \leq T$, um martingale com respeito a esta filtração (i.e., $\forall t$, $M(t)$ é $\mathcal{F}(t)$ -mensurável e para*

$0 \leq s \leq t \leq T$, $\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s)$ e $\mathbb{E}[M^2(t)] < \infty$, $\forall t \in [0, T]$. Então existe um processo adaptado $\Gamma(u)$, $0 \leq u \leq T$, tal que

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u), \quad 0 \leq t \leq T$$

Demonstração. Veja [10], capítulo 5, seção 3. □

Definição 3.2.3 (Arbitragem). *Uma oportunidade de arbitragem resume-se em um processo de riqueza auto-financiado admissível $X(t)$ que satisfaça as seguintes condições:*

(a) $X(0) = 0$ e para algum $T > 0$ ainda satisfaça

(b) $\mathbb{P}\{X(T) \geq 0\} = 1$, $\mathbb{P}\{X(T) > 0\} > 0$.

Arbitragem é uma maneira de negociação que permita que um agente comece sem capital e em um tempo posterior T esteja certo que não haverá perda de dinheiro e ainda tenha uma probabilidade positiva de ter ganho dinheiro.

Teorema 3.2.3 (Primeiro Teorema Fundamental da Precificação de Ativos). *Se o modelo de mercado possui uma medida de probabilidade neutra ao risco, então ele não admite arbitragem.*

Demonstração. Veja [10], Capítulo 5, Seção 4. □

Teorema 3.2.4 (Segundo Teorema Fundamental da Precificação de Ativos). *Considere um modelo de mercado que possui uma medida de probabilidade neutra ao risco. O modelo é completo se, e somente se, a medida de probabilidade neutra ao risco é única.*

Demonstração. Veja [10], Capítulo 5, Seção 4. □

Definição 3.2.4 (Contrato Futuro). *O preço futuro de um ativo cujo valor no tempo T é $S(T)$ é dado pela fórmula*

$$Fut_S(t, T) = \tilde{\mathbb{E}}[S(T)|\mathcal{F}(t)], \quad 0 \leq t \leq T.$$

Teorema 3.2.5. *O preço do contrato futuro é martingale se na medida neutra ao risco $\tilde{\mathbb{P}}$, ele satisfaz $Fut_S(T, T) = S(T)$ e se o valor de uma posição longa (ou curta) a ser realizada durante um intervalo de tempo é sempre zero.*

Demonstração. Veja [10], Capítulo 5, Seção 6. □

3.3 Método de Euler-Maruyama

Uma equação diferencial estocástica pode ser escrita na forma integral como segue:

$$X(t) = X(0) + \int_0^t f(t, X(s))ds + \int_0^t g(t, X(s))dW(s), \quad 0 \leq t \leq T, \quad (3.2)$$

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. Temos que a Equação (3.2) é solução da Equação Diferencial Estocástica (3.3).

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad X(0) = X_0 \quad 0 \leq t \leq T. \quad (3.3)$$

Para aplicar o método numérico na Equação (3.3) em $[0, T]$ vamos discretizar o intervalo em L partes iguais onde L é um número inteiro positivo. Seja $\Delta t = T/L$ e $\tau_j = j\Delta t$. A aproximação numérica para $X(\tau_j)$ será denotada por X_j . O Método de Euler-Maruyama segue a forma:

$$X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})(W(\tau_j) - W(\tau_{j-1})), \quad j = 1, 2, \dots, L. \quad (3.4)$$

Analogamente, podemos ver a forma integral da Equação (3.2) discretizada pelo método de Euler-Maruyama.

$$X(\tau_j) = X(\tau_{j-1}) + \int_{\tau_{j-1}}^{\tau_j} f(X(s))ds + \int_{\tau_{j-1}}^{\tau_j} g(X(s))dW(s). \quad (3.5)$$

Os incrementos $W(\tau_j) - W(\tau_{j-1})$ serão calculados com um passo δt de tal modo que $\Delta t = R\delta t$, com $R \geq 1$. Isto garantirá que o conjunto de pontos t_j onde o Browniano é discretizado contenha o conjunto de pontos τ_j , onde a solução pelo método de Euler-Maruyama é calculada. Para garantir tal condição podemos calcular os incrementos Brownianos como na Equação 3.6.

$$W(\tau_j) - W(\tau_{j-1}) = W(jR\delta t) - W((j-1)R\delta t) = \sum_{k=R(j-1)+1}^{jR} dW_k. \quad (3.6)$$

3.3.1 Convergência do Método de Euler-Maruyama

Nesta seção serão apresentadas noções de convergência forte e convergência fraca para o Método de Euler-Maruyama.

Um método converge fortemente com ordem γ se existe uma constante C tal que a equação 3.7 seja satisfeita.

$$E|X_n - X(\tau)| \leq C\Delta t^\gamma \quad (3.7)$$

para qualquer $\tau = n\Delta t \in [0, T]$ fixo e Δt suficientemente pequeno. Se $f \in C^4$, $g \in C^4$, ou seja, suas quatro primeiras derivadas contínuas e suas primeiras derivadas limitadas e crescerem no máximo em ordem polinomial, pode-se provar que o método de Euler-Maruyama converge fortemente com ordem $\gamma = 1/2$. Para um discussão intuitiva sobre esta convergência, consultar [11]. Para algo mais formal, consultar [12].

Supondo que as condições acima sejam satisfeitas podemos definir para um t fixo em $[0, T]$ o erro no sentido forte para um Δt suficientemente pequeno (Equação 3.8).

$$e_{\Delta t}^{forte} := E|X_n - X(\tau)| \leq C\Delta t^{1/2} \quad (3.8)$$

Esta convergência forte de ordem $1/2$ significa que se desejarmos diminuir o erro 100 vezes teremos que diminuir o passo Δt em 100^2 vezes. Em resumo, podemos

dizer 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 Δt tende para zero.

O conceito agora explorado é o de convergência fraca, onde podemos avaliar a convergência do métodos para os momentos de X_n e $X(\tau)$ conforme Δt tende para zero. Sendo assim, um método é dito convergente fraco com ordem γ se existe uma constante C que respeita a Equação (3.9) para todo polinômio p .

$$|E[p(X_n)] - E[p(X(\tau))]| \leq C\Delta t^\gamma \quad (3.9)$$

Para qualquer $\tau = n\Delta t \in [0, T]$ fixo e Δt suficientemente pequeno. Novamente temos que se f e g satisfizerem as condições apropriadas o Método de Euler-Maruyama possui ordem de convergência fraca γ igual a 1. Veja [11] e [12] para maiores detalhes.

Alguns cuidados devem ser tomados na escolha do passo para o cálculo das equações com o Método de Euler-Maruyama, são eles:

- O passo não pode ser grande de modo a prejudicar a convergência do método;
- O passo não pode ser tomado com magnitude próxima ao *epsilon* da máquina, pois alguns erros de arredondamento farão o erro crescer ao invés de diminuir;
- O tempo computacional aumenta exponencialmente conforme o inverso de γ , ou seja, para o caso com $\gamma = 1/2$ se diminuirmos o passo em 4 vezes o tempo de cálculo da resposta deverá aumentar na ordem de 16 vezes.

O *trade-off* deve ser atingido para um passo que não seja grande a ponto de gerar erros significativos e nem pequeno a ponto de inviabilizar o cálculo por tempo computacional e requisitos de memória. A questão da magnitude próxima ao *epsilon* da máquina não é tão relevante, pois muito antes de chegar a este ponto o tempo computacional já seria um gargalo.

3.4 Filtro de Kalman

Esta seção fará uma breve introdução ao filtro de Kalman, baseada na referência [13].

Em primeiro lugar é importante dar uma breve definição do que são variáveis de estado e variáveis observáveis. As variáveis de estado no contexto de Kalman são não-observáveis possuindo uma dinâmica conhecida. As variáveis observáveis são aquelas que observamos diretamente e que são utilizadas para estimar as variáveis de estado.

O Filtro de Kalman é um procedimento recursivo para calcular estimativas das variáveis de estado não observáveis baseado em variáveis observáveis relacionadas à essas variáveis não observáveis. Esta técnica é muito útil em várias práticas da engenharia e finanças, pois o mesmo pode ser implementado para operar em tempo real dada sua natureza recursiva.

Dada uma distribuição a priori do valor inicial das variáveis de estado e um modelo descrevendo a verossimilhança das observações como função dos valores reais, o Filtro de Kalman gera uma distribuição a posteriori destas variáveis de estado de acordo com a regra de Bayes. Além disso, o filtro é uma ferramenta eficiente para minimizar o erro quadrático da estimação.

Seja $\mathbf{y}_t \in \mathbb{R}^N$ o vetor das variáveis observáveis e $\mathbf{x}_t \in \mathbb{R}^m$ o vetor das variáveis de estado. Conhecidos estes vetores podemos relacioná-los através das equações de observação e transição. A primeira delas visa relacionar o vetor de observações com as variáveis de estado que não são observáveis. A Equação (3.10) mostra esta relação.

$$\mathbf{y}_t = \mathbf{Z}_t \mathbf{x}_t + \mathbf{d}_t + \boldsymbol{\epsilon}_t, \quad t = 1, 2, \dots \quad (3.10)$$

onde:

- \mathbf{y}_t é um vetor $N \times 1$, onde N é o número de séries temporais;
- \mathbf{Z}_t é uma matriz $N \times m$, onde m é o número de variáveis de estado;

- \mathbf{x}_t é um vetor $m \times 1$, que representa as variáveis de estado não observáveis;
- \mathbf{d}_t é um vetor $N \times 1$;
- $\boldsymbol{\epsilon}_t$ é um vetor $N \times 1$, serialmente não correlacionado com média zero e matriz de covariância \mathbf{H}_t .

As variáveis de estado são geradas através de um processo Markoviano de primeira ordem como pode ser visto na Equação (3.11).

$$\mathbf{x}_t = \mathbf{T}_t \mathbf{x}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t, \quad t = 1, 2, \dots \quad (3.11)$$

onde:

- \mathbf{T}_t é uma matriz $m \times m$;
- \mathbf{c}_t é uma vetor $m \times 1$;
- \mathbf{R}_t é uma matriz $m \times g$, onde g é o número de componentes de incerteza associadas as variáveis de estado não observáveis;
- $\boldsymbol{\eta}_t$ é um vetor $g \times 1$, serialmente não correlacionado com média zero e matriz de covariância \mathbf{Q}_t .

É importante ressaltar que o vetor inicial de estado não observável \mathbf{x}_0 possui média $\hat{\mathbf{x}}_0$ e matriz de covariância \mathbf{P}_0 . Além disso, os ruídos $\boldsymbol{\epsilon}_t$ e $\boldsymbol{\eta}_t$ são não correlacionados entre si e não correlacionados com o estado inicial.

Uma vez definido o modelo com as equações de observação e transição podemos agora verificar o algoritmo de Kalman para estimação das variáveis não observáveis. Basicamente implementa-se um processo recursivo que calcula as equações de evolução da estimação das variáveis de estado e de sua matriz de covariância, como pode ser visto nas equações 3.12 e 3.13.

$$\hat{\mathbf{x}}_t^- = \mathbf{T}_t \hat{\mathbf{x}}_{t-1} + \mathbf{c}_t \quad (3.12)$$

onde:

- $\hat{\mathbf{x}}_{t-1}$ é o estimador ótimo² de \mathbf{x}_{t-1} ;
- $\hat{\mathbf{x}}_t^-$ é a evolução no tempo t baseado no estimador ótimo do tempo $t - 1$.

$$\mathbf{P}_t^- = \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}_t' + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t' \quad (3.13)$$

onde:

- \mathbf{P}_t^- é a matriz de covariância de $\hat{\mathbf{x}}_t^-$ baseado na matriz de covariância anterior \mathbf{P}_{t-1} .

Agora de posse das equações de previsão podemos calcular as equações de atualização das medições. São elas (3.14), (3.15) e (3.16).

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{Z}_t' (\mathbf{Z}_t \mathbf{P}_t^- \mathbf{Z}_t' + \mathbf{H}_t)^{-1} \quad (3.14)$$

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t^- + \mathbf{K}_t (\mathbf{y}_t - \mathbf{Z}_t \hat{\mathbf{x}}_t^- - \mathbf{d}_t) \quad (3.15)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{Z}_t) \mathbf{P}_t^- \quad (3.16)$$

De posse de todas as equações descritas acima é possível implementar o algoritmo do filtro de Kalman e a cada rodada atualizar a estimativa para as variáveis de estado não observáveis.

A implementação do filtro é uma tarefa relativamente simples, no entanto para minimizar o erro quadrático das previsões se faz necessário encontrar o estimador ótimo para $\hat{\mathbf{x}}_t$, o que implica em um processo de otimização que muitas vezes é complicado de resolver. No caso em que o modelo é Gaussiano, é possível mostrar que a função verossimilhança dado um vetor de hiperparâmetros Θ é dada pela Equação (3.17).

²Entende-se por estimador ótimo aquele que minimiza o erro médio quadrático entre a média da distribuição condicional de $\hat{\mathbf{x}}_t$ e \mathbf{x}_t

$$\ln(L[y|\Theta]) = -\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^T \ln(\det(\mathbf{F}_t)) - \frac{1}{2}\sum_{t=1}^T \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t \quad (3.17)$$

onde:

- T é o número de amostras do modelo;
- $\mathbf{v}_t = \mathbf{y}_t - \mathbf{y}_{t|t-1}$ é o vetor dos erros de previsão para cada t ;
- $\mathbf{F}_t = \mathbf{Z}_t \mathbf{P}_t^- \mathbf{Z}_t' + \mathbf{H}_t$ é a matriz de covariância.

Dizemos que a Equação (3.17) é a decomposição dos erros de previsão. A cada rodada do Filtro de Kalman é possível calcular a função de verossimilhança e portanto podemos implementar um algoritmo otimizador para maximizar esta função dado o vetor de hiperparâmetros Θ .

4 Modelo de Schwartz-Smith

Este capítulo visa detalhar o modelo de Schwartz-Smith e demonstrar os principais resultados obtidos no artigo [3]. Basicamente começaremos com uma breve explicação intuitiva do modelo de dois fatores, depois entraremos na modelagem propriamente dita e no apreçamento de futuros de commodities na medida neutra ao risco.

4.1 Modelo na Medida Física

Seja S_t o preço *spot* de uma commodity no tempo t . A idéia fundamental é dividir o preço *spot* em dois fatores estocásticos, sendo $\ln(S_t) = \chi_t + \xi_t$, onde χ_t se refere ao desvio de curto prazo dos preços e ξ_t o preço de equilíbrio. Os desvios de curto prazo (χ_t) são modelados como um processo de Ornstein-Uhlenbeck com reversão à média para zero. A Equação (4.1) mostra este comportamento.

$$d\chi_t = -\kappa\chi_t dt + \sigma_\chi dZ_\chi \quad (4.1)$$

O preço de equilíbrio (ξ_t) segue um movimento Browniano seguindo a Equação (4.2).

$$d\xi_t = \mu_\xi dt + \sigma_\xi dZ_\xi \quad (4.2)$$

Além disso temos que os incrementos Brownianos das Equações (4.1) e (4.2) são correlacionados segundo a Equação (4.3).

$$dZ_\chi dZ_\xi = \rho_{\chi\xi} dt \quad (4.3)$$

Podemos interpretar que as mudanças que afetam o fator de desvio de curto prazo (χ_t) são de natureza transitória, tais como variações climáticas e falhas no fornecimento da commodity. Já para o fator de equilíbrio (ξ_t) espera-se que seja de natureza duradoura tais como a mudança na capacidade produtiva e até a descoberta de uma nova tecnologia que afetem estruturalmente o mercado.

Para encontrar a solução do processo que modela as variações de preços de curto prazo χ_t basta aplicar a Fórmula de Itô na função $f(x, t) = e^{\kappa t} \chi_t$. Isto resulta na Equação (4.4).

$$\begin{aligned} d(e^{\kappa t} \chi_t) &= \kappa e^{\kappa t} \chi_t + e^{\kappa t} d\chi_t \\ &= \kappa e^{\kappa t} \chi_t + e^{\kappa t} (-\kappa \chi_t + \sigma_\chi dZ_\chi) \\ &= e^{\kappa t} \sigma_\chi dZ_\chi \end{aligned} \quad (4.4)$$

Integrando ambos os lados da Equação (4.4) de t até T , teremos a Equação (4.5).

$$\begin{aligned} e^{\kappa T} \chi_T - e^{\kappa t} \chi_t &= \sigma_\chi \int_t^T e^{\kappa s} dZ_\chi(s) \\ e^{\kappa T} \chi_T &= e^{\kappa t} \chi_t + \sigma_\chi \int_t^T e^{\kappa s} dZ_\chi(s) \end{aligned} \quad (4.5)$$

Simplificando (4.5) chegamos na solução χ_T como pode ser visto em (4.6).

$$\chi_T = e^{-\kappa(T-t)} \chi_t + \sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s) \quad (4.6)$$

Uma vez encontrada a solução da equação de desvios de curto prazo χ_T vamos calcular a média e a variância condicionados a χ_t .

$$\begin{aligned} \mathbb{E}[\chi_T | \chi_t] &= \mathbb{E} \left[e^{-\kappa(T-t)} \chi_t + \sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s) | \chi_t \right] \\ &= \mathbb{E} \left[e^{-\kappa(T-t)} \chi_t | \chi_t \right] + \sigma_\chi e^{-\kappa T} \mathbb{E} \left[\int_t^T e^{\kappa s} dZ_\chi(s) | \chi_t \right] \\ &= e^{-\kappa(T-t)} \chi_t + \sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s) \\ &= e^{-\kappa(T-t)} \chi_t \end{aligned} \quad (4.7)$$

A Equação (4.7) é obtida pois $\int_t^T e^{\kappa s} dZ_\chi(s)$ por se tratar de uma integral estocástica é martingal.

Agora vamos calcular a variância de χ_T através da Equação (4.8) utilizando a Isometria de Itô.

$$\begin{aligned}
\text{Var}[\chi_T|\chi_t] &= \text{Var}\left[e^{-\kappa(T-t)}\chi_t + \sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s)|\chi_t\right] \\
&= \text{Var}\left[\sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s)|\chi_t\right] \\
&= \sigma_\chi^2 e^{-2\kappa T} \int_t^T e^{2\kappa s} ds \\
&= \left(1 - e^{-2\kappa(T-t)} \frac{\sigma_\chi^2}{2\kappa}\right)
\end{aligned} \tag{4.8}$$

Seguindo passos similares vamos calcular a solução do processo ξ_t . O preço de equilíbrio segue um movimento Browniano simples e pode ser resolvida integrando ambos os lados da Equação (4.2). Daí obtemos

$$\begin{aligned}
\int_t^T d\xi_s &= \int_t^T \mu_\xi ds + \sigma_\xi \int_t^T dZ_\xi(s) \\
\xi_T - \xi_t &= \mu_\xi(T-t) + \sigma_\xi \int_t^T dZ_\xi(s) \\
\xi_T &= \xi_t + \mu_\xi(T-t) + \sigma_\xi \int_t^T dZ_\xi(s)
\end{aligned} \tag{4.9}$$

Utilizando os mesmos argumentos vamos calcular os momentos condicionais de ξ_T . A média é dada por

$$\begin{aligned}
\mathbb{E}[\xi_T|\xi_t] &= \mathbb{E}\left[\xi_t + \mu_\xi(T-t) + \sigma_\xi \int_t^T dZ_\xi(s)|\xi_t\right] \\
&= \xi_t + \mu_\xi(T-t)
\end{aligned} \tag{4.10}$$

onde novamente vemos que o único termo estocástico é $\sigma_\xi \int_t^T dZ_\xi(s)$ e portanto ξ_T possui a mesma distribuição do Browniano Z_ξ .

Utilizando a Isometria de Itô novamente teremos que a variância do processo ξ_T é dada pela Equação (4.11).

$$\begin{aligned}
\text{Var}[\xi_T|\xi_t] &= \text{Var}\left[\xi_t + \mu_\xi(T-t) + \sigma_\xi \int_t^T dZ_\xi(s)|\xi_t\right] \\
&= \text{Var}\left[\sigma_\xi \int_t^T dZ_\xi(s)|\xi_t\right] \\
&= \sigma_\xi^2 \int_t^T ds \\
&= \sigma_\xi^2(T-t)
\end{aligned} \tag{4.11}$$

O modelo de Schwartz-Smith assume que os Brownianos Z_χ e Z_ξ possuem coeficiente de correlação $\rho_{\chi\xi}$, portanto devemos então calcular a covariância dos processos. Este resultado é obtido na Equação (4.12).

$$\begin{aligned}
\text{Cov}[\chi_T, \xi_T] &= \mathbb{E}[(\chi_T - \mathbb{E}[\chi_T])(\xi_T - \mathbb{E}[\xi_T])] \\
&= \mathbb{E}\left[\sigma_\chi e^{-\kappa T} \int_t^T e^{\kappa s} dZ_\chi(s) \sigma_\xi \int_t^T dZ_\xi(s)\right] \\
&= \sigma_\chi e^{-\kappa T} \sigma_\xi \mathbb{E}\left[\int_t^T e^{\kappa s} dZ_\chi(s) \int_t^T dZ_\xi(s)\right] \\
&= \sigma_\chi e^{-\kappa T} \sigma_\xi \rho_{\chi\xi} \int_t^T e^{\kappa s} ds \\
&= (1 - e^{-\kappa(T-t)}) \frac{\sigma_\chi \sigma_\xi \rho_{\chi\xi}}{\kappa}
\end{aligned} \tag{4.12}$$

No apêndice B será provado que dado χ_t e ξ_t , os processos χ_T e ξ_T possuem distribuição conjunta normal com média e matrix de covariância dadas por:

$$\mathbb{E}[(\chi_T, \xi_T)] = [e^{-\kappa(T-t)}\chi_t, \xi_t + \mu_\xi(T-t)] \tag{4.13}$$

$$\text{Cov}[(\chi_T, \xi_T)] = \begin{bmatrix} (1 - e^{-2\kappa(T-t)}) \frac{\sigma_\chi^2}{2\kappa} & (1 - e^{-\kappa(T-t)}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \\ (1 - e^{-\kappa(T-t)}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} & \sigma_\xi^2(T-t) \end{bmatrix} \tag{4.14}$$

Lembrando que o logaritmo do preço *spot* no tempo t é a soma dos processos χ_t e ξ_t , então temos que o preço S_t possui distribuição log-normal. Para achar uma expressão fechada para o preço *spot* vamos calcular a média (4.15) e a variância

(4.16) do seu logaritmo e depois utilizar uma relação conveniente que relaciona a média e a variância de variáveis aleatórias normais e log-normais (a demonstração está no apêndice A).

$$\begin{aligned}
\mathbb{E}[\ln(S_T)|S_t] &= \mathbb{E}[\chi_T + \xi_T | (\chi_t, \xi_t)] \\
&= \mathbb{E}[\chi_T | \chi_t] + \mathbb{E}[\xi_T | \xi_t] \\
&= e^{-\kappa(T-t)}\chi_t + \mu_\xi(T-t)
\end{aligned} \tag{4.15}$$

$$\begin{aligned}
\text{Var}[\ln(S_T)|S_t] &= \text{Var}[\chi_T | \chi_t] + \text{Var}[\xi_T | \xi_t] + 2\text{Cov}[(\chi_T, \xi_T)] \\
&= (1 - e^{-2\kappa(T-t)})\frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2(T-t) + 2(1 - e^{-\kappa(T-t)})\frac{\sigma_\chi\sigma_\xi\rho_{\chi\xi}}{\kappa}
\end{aligned} \tag{4.16}$$

Agora vamos utilizar a relação dada na (4.17) para finalmente calcular a média do preço *spot* que posteriormente será utilizada para calcular o preço futuro da commodity em estudo.

$$\begin{aligned}
\mathbb{E}[S_T] &= \exp(\mathbb{E}[\ln(S_T)] + \frac{1}{2}\text{Var}[\ln(S_T)]) \quad \text{ou} \\
\ln(\mathbb{E}[S_T]) &= \mathbb{E}[\ln(S_T)] + \frac{1}{2}\text{Var}[\ln(S_T)]
\end{aligned} \tag{4.17}$$

Sendo assim teremos que a média do preço *spot* será dada pela equação 4.18.

$$\begin{aligned}
\ln(\mathbb{E}[S_T]) &= e^{-\kappa(T-t)}\chi_t + \mu_\xi(T-t) + \\
&+ \frac{1}{2}\left((1 - e^{-2\kappa(T-t)})\frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2(T-t) + 2(1 - e^{-\kappa(T-t)})\frac{\sigma_\chi\sigma_\xi\rho_{\chi\xi}}{\kappa} \right)
\end{aligned} \tag{4.18}$$

4.2 Modelo na Medida Neutra ao Risco

Vamos agora desenvolver a versão neutra ao risco do modelo de Schwartz-Smith que será utilizada para valorar os contratos futuros de commodities. A idéia principal é introduzir dois parâmetros adicionais constantes que serão interpretados como os prêmios de risco de mercado para cada variável¹. Sendo assim λ_χ , λ_ξ e

¹Estes prêmios de risco foram considerados constantes por questões de simplificação do modelo assim como na referência [3]. Uma generalização pode ser adotada de modo que o prêmio de risco

juntamente com o teorema de Girsanov desenvolvemos o modelo dado por:

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dZ_\chi^* \quad (4.19)$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dZ_\xi^* \quad (4.20)$$

Novamente temos que dZ_χ^* e dZ_ξ^* são incrementos Brownianos correlacionados com $dZ_\chi^* dZ_\xi^* = \rho_{\chi\xi} dt$. Agora o processo neutro ao risco associado aos desvios de curto prazo (χ_t) é Ornstein-Uhlenbeck revertendo a média $-\lambda_\chi/\kappa$ e o processo neutro ao risco para os preços de equilíbrio continua sendo um movimento Browniano com um *drift* $\mu_\xi^* \equiv \mu_\xi - \lambda_\xi$. Nota-se o aparecimento dos prêmios de risco de mercado quando trocamos da medida física para a medida neutra ao risco.

Da mesma forma que no modelo original é provado que dado χ_t e ξ_t , a distribuição conjunta de χ_T e ξ_T é normal com o vetor de média e matriz de covariância da seguinte forma:

$$\mathbb{E}^*[(\chi_T, \xi_T)] = [e^{-\kappa(T-t)}\chi_t - (1 - e^{-\kappa(T-t)})\lambda_\chi/\kappa, \xi_t + \mu_\xi^*(T - t)], \quad (4.21)$$

$$\mathbb{C}ov^*[(\chi_T, \xi_T)] = Cov[(\chi_T, \xi_T)], \quad (4.22)$$

onde $*$ denota que as médias e as variâncias são tomadas na medida neutra ao risco. Neste processo neutro ao risco o logaritmo do preço *spot*, $\ln(S_t) = \chi_t + \xi_t$, possui distribuição normal com os parâmetros das equações (4.23) e (4.24).

$$\mathbb{E}^*[\ln(S_T)|S_t] = e^{-\kappa(T-t)}\chi_t + \xi_t - (1 - e^{-\kappa(T-t)})\lambda_\chi/\kappa + \mu_\xi^*(T - t) \quad (4.23)$$

$$\mathbb{V}ar^*[\ln(S_T)|S_t] = \mathbb{V}ar[\ln(S_T)|S_t] \quad (4.24)$$

seja uma função afim do preço da commodity.

4.3 Valorando Contratos Futuros

Para a valoração dos contratos futuros vamos considerar $F_{t,T}$ como sendo o preço do contrato futuro no tempo t , com maturidade no tempo T ($0 \leq t \leq T$). No contexto de valoração neutra ao risco temos que os preços dos contratos futuros são iguais à expectativa do futuro do preço *spot* no processo neutro ao risco, assumindo que as taxas de juros são determinísticas. Podemos assim escrever uma expressão para o valor dos contratos futuros a saber

$$\begin{aligned}
 \ln(F_{t,T}) &= \ln(\mathbb{E}^*[S_{t,T}]) \\
 &= \mathbb{E}^*[\ln(S_{t,T})] + 1/2\text{Var}^*[\ln(S_{t,T})] \\
 &= e^{-\kappa(T-t)}\chi_t^* + \xi_t^* + A(t, T)
 \end{aligned} \tag{4.25}$$

onde

$$\begin{aligned}
 A(t, T) &= \mu_\xi^*(T-t) - (1 - e^{-\kappa(T-t)})\frac{\lambda_\chi}{\kappa} + \\
 &+ \frac{1}{2} \left((1 - e^{-2\kappa(T-t)})\frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2(T-t) + 2(1 - e^{-\kappa(T-t)})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} \right)
 \end{aligned} \tag{4.26}$$

5 Resultados

Este Capítulo visa mostrar os principais resultados dos métodos implementados no trabalho. Ele está dividido da seguinte forma: Primeiramente falaremos do pré-processamento dos dados, logo depois da simulação dos preços dos futuros de gás natural (*henry-hub*), entraremos então no processo de filtragem com o filtro de Kalman e por fim fecharemos com a calibração do modelo.

5.1 Pré-Processamento dos Dados

Os dados utilizados neste trabalho correspondem a 244 séries de futuros de gás natural *Henry-Hub* com vencimentos mensais começando em Junho de 1990 e terminando em Setembro de 2010. Todos os dados foram extraídos da Bloomberg.

Os dados reais são compostos de séries de preços com um determinado vencimento, onde as séries mais longas chegam a 6 anos. Para evitar problemas com o vencimento de um determinado contrato futuro foi feito o que chamamos de "rolagem" do contrato. Assim, teremos séries temporais no qual cada uma corresponde a um vencimento.

O algoritmo para "rolagem" dos contratos é bastante simples. Basta concatenar o final da série do futuro anterior com o final da série do futuro atual. As figuras 2 e 3 mostram o procedimento.

Nota-se que as partes pontilhadas antes do pré-processamento são os valores a um mês do vencimento. Estes "pedaços" são concatenados de modo a formar uma série única que não expira. Após o pré-processamento escolhemos o maior bloco retangular possível com 24 vencimentos. Ao final do processo teremos um total de 24 séries que correspondem a algum vencimento em meses.

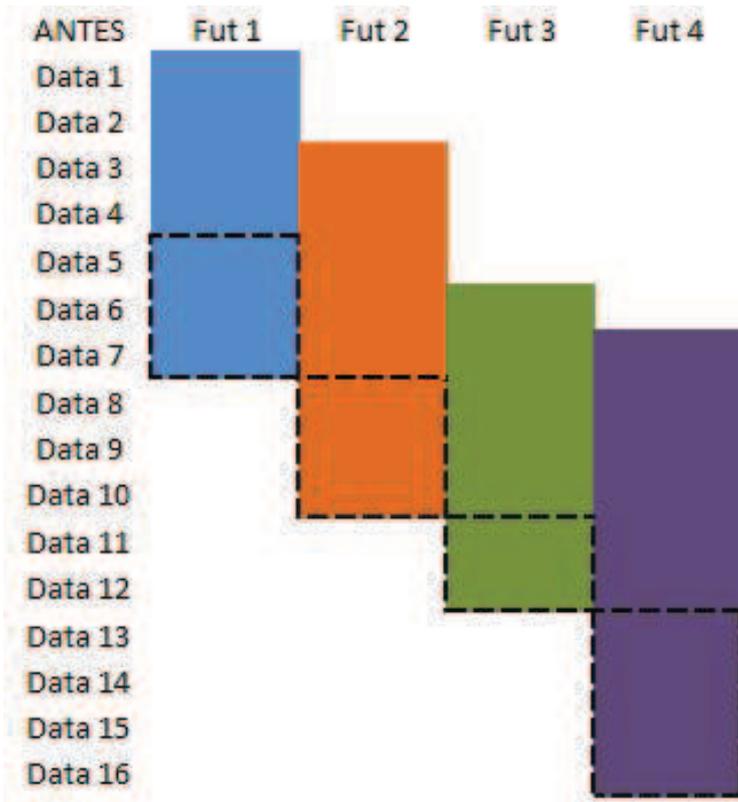


Figura 2: Formato dos dados antes da rolagem



Figura 3: Formato dos dados depois da rolagem

A Figura 4 mostra os dados reais após o pré-processamento.

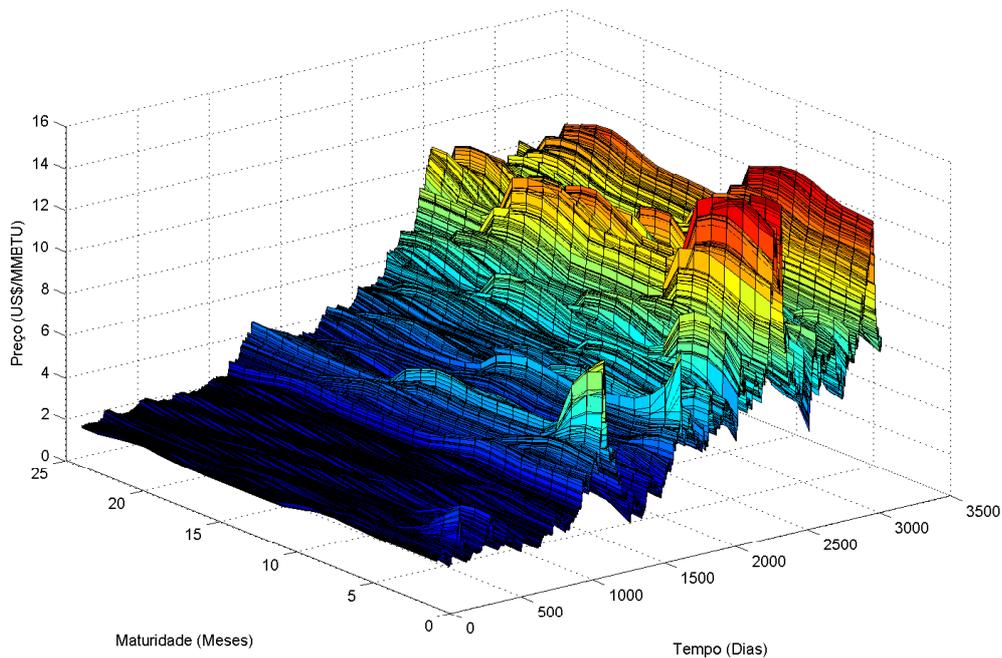


Figura 4: Dados do Henry Hub após o pré-processamento

Após o pré-processamento das séries de futuros obtivemos uma série que corresponde ao tempo ao vencimento (τ) de um determinado futuro (Figura 5). Nota-se claramente que ela acompanha um padrão bem definido para as novas séries após o pré-processamento. Se observarmos a série de tempo ao vencimento do contrato mais próximo veremos que o vencimento do contrato varia de um dia útil até vinte um ou vinte e dois dias úteis. O mesmo raciocínio se aplica as demais séries de tempo ao vencimento. Nota-se que o padrão de cores se repete, mas não há nenhuma relação entre as séries de tempo ao vencimento de cores iguais.

5.2 Simulação Henry-Hub

Nesta seção será simulada uma possível realização dos preços futuros de gás natural seguindo o modelo de Schwartz-Smith.

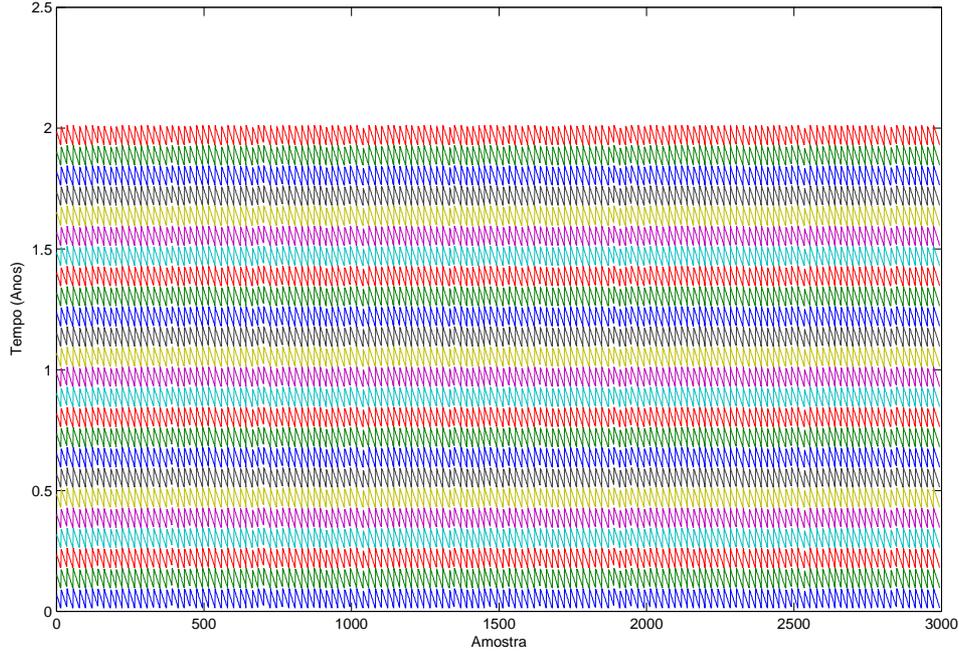


Figura 5: Tempo ao Vencimento (τ)

5.2.1 Simulação dos Processos χ_t e ξ_t

Conforme dito no Capítulo 4 temos que o processo χ_t governa a dinâmica dos desvios de preço de curto prazo e o processo ξ_t governa a dinâmica do preço de equilíbrio. As equações que devem ser discretizadas segundo o método de Euler-Maruyama dos processos citados estão descritas em (5.1) e (5.2). Ambas já se encontram na medida neutra ao risco.

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dZ_\chi^* \quad (5.1)$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dZ_\xi^* \quad (5.2)$$

Após a discretização pelo Método de Euler-Maruyama chegamos as equações (5.3) e (5.4).

$$\begin{aligned} \chi_{n+1} &= \chi_n + (-\kappa\chi_n - \lambda_\chi)\Delta t + \sigma_\chi \Delta W_n \\ &= -\lambda_\chi + (1 - \kappa\Delta t)\chi_n + \sigma_\chi \Delta W_n \end{aligned} \quad (5.3)$$

κ	λ_χ	σ_χ	μ_ξ	λ_ξ	σ_ξ	$\rho_{\chi\xi}$
1,49	15,7%	28,6%	-1,25%	-2,5%	14,5%	0,3

Tabela 1: Parâmetros para Simulação

$$\xi_{n+1} = \xi_n + (\mu_\xi - \lambda_\xi)\Delta t + \sigma_\xi \Delta W_n \quad (5.4)$$

Os parâmetros utilizados na simulação encontram-se na Tabela 1. Estes parâmetros são os mesmos encontrados no artigo [3].

Nas Figuras 6 e 7 podemos observar a realização dos processos χ_t e ξ_t para 1000 observações e 10000 observações respectivamente.

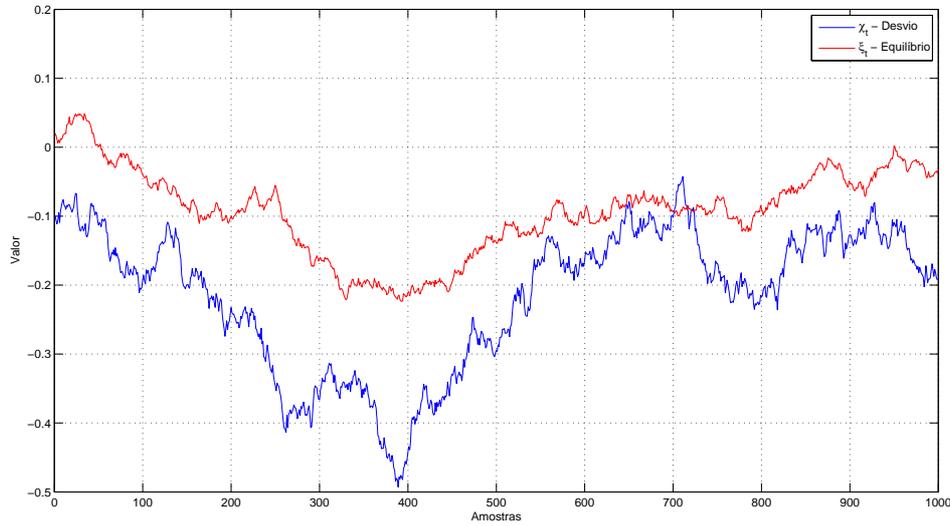


Figura 6: Realização dos processos χ_t e ξ_t com 1000 amostras

Para aferir se as simulações estão coerentes com os resultados teóricos do modelo foram feitas diversas simulações dos processos χ_t e ξ_t para então comparar os resultados teóricos de média, variância e covariância dos processos com os simulados. Estes resultados podem ser encontrados no Apêndice C nas figuras 13, 14, 15, 16 e 17.

5.2.2 Simulação dos Futuros

Para simular os preços dos futuros de Gás Natural vamos utilizar os mesmos parâmetros da Tabela 1, assim como os processos χ_t e ξ_t simulados na seção anterior.

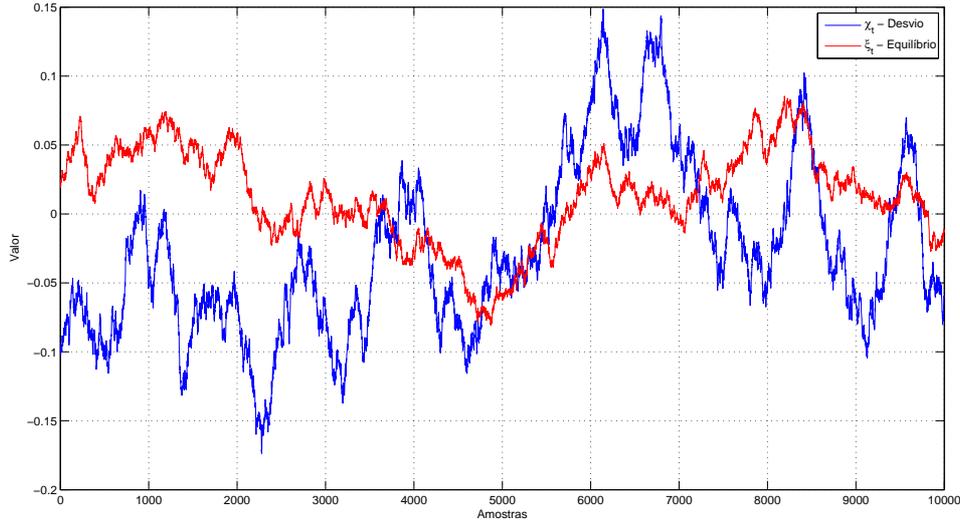


Figura 7: Realização dos processos χ_t e ξ_t com 10000 amostras

Sendo assim basta aplicar as equações (4.25) e (4.26).

De posse do tempo ao vencimento $\tau = T - t$ podemos calcular a função determinística do preço dos contratos futuros $A(\tau)$, como pode ser visto na figura 8.

A Figura 9 mostra uma possível realização dos preços futuros do Henry Hub simulados.

5.3 Filtragem de Kalman

Nesta seção será implementado o Filtro de Kalman conforme a Seção 3.4.

Como descrito em [3], temos que as variáveis de estado são o desvio de curto-prazo (χ_t), o preço de equilíbrio (ξ_t) e as observações são o logaritmo dos preços dos contratos futuros. Sendo assim podemos definir a equação de transição (*transition equation*) que rege a evolução das variáveis de estado e a equação de medição (*measurement equation*) que relaciona as variáveis de estado e os preços observados.

Para testar o filtro de Kalman implementado foram simulados os preços futuros como na seção 5.2. Lembrando que os parâmetros utilizados constam na Tabela 1.

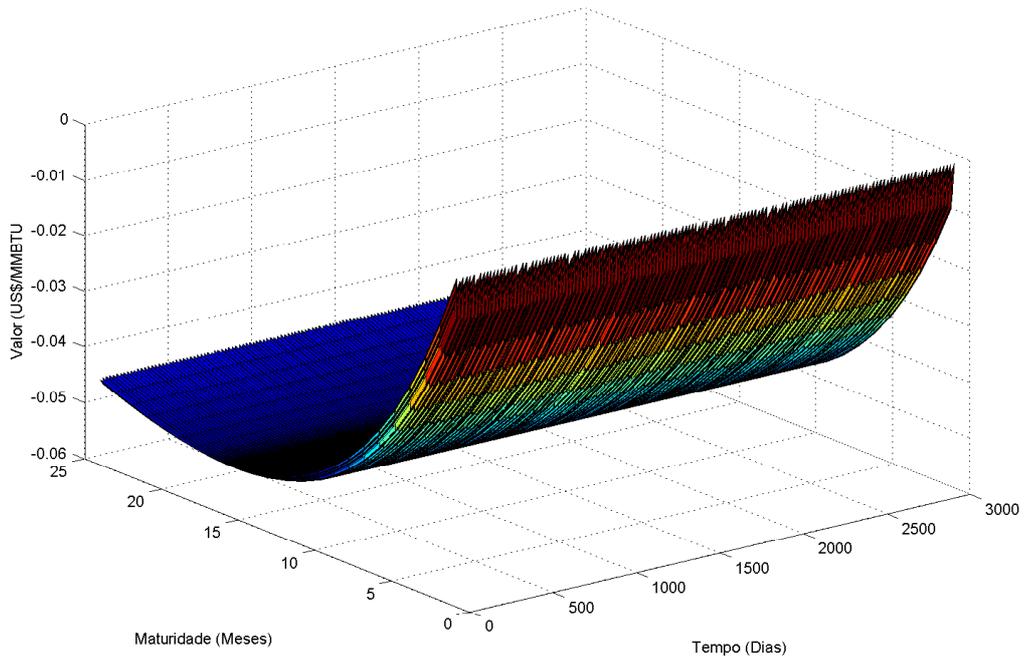


Figura 8: Superfície de $A(\tau)$

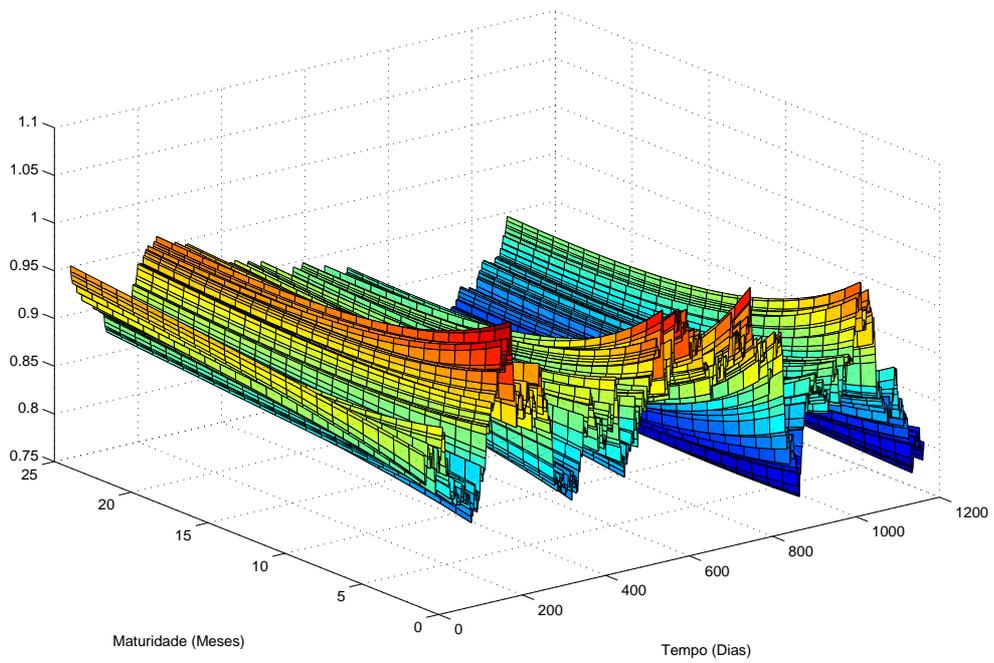


Figura 9: Preços Simulados para 24 maturidades e 1200 amostras

Uma vez terminada a simulação foi introduzido um ruído Gaussiano aos log-preços simulados o qual chamamos de ϵ_t . Este ruído possui média zero e variância arbitrária de 0,0001 em todas as componentes com maturidades diferentes. Chamamos de H_t a matriz de covariância do ruído gaussiano ϵ_t . Na prática, as variâncias certamente serão diferentes para cada componente. Por questão de simplicidade consideramos a matriz mostrada na equação (5.5) ou de forma simplificada $H_t = \text{diag}(s_1^2, s_2^2, \dots, s_N^2)$. Esta premissa é muito razoável pois supomos no modelo que os ruídos são descorrelacionados entre si e também porque no processo de otimização teremos que estimar apenas N variâncias dos ruídos, onde N é o número de séries de futuros gerados. Se tivéssemos uma matriz completa teríamos que estimar N^2 parâmetros o que tornaria o processo inviável. No trabalho em questão $N = 24$.

$$H_t = \begin{bmatrix} s_1^2 & 0 & \dots & 0 \\ 0 & s_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_N^2 \end{bmatrix} \quad (5.5)$$

Voltando ao equacionamento teremos ainda as matrizes Z_t e d_t para a equação de observação, conforme mostrado em (5.6).

$$Z_t = \begin{bmatrix} e^{-\kappa\tau_{1t}} & 1 \\ e^{-\kappa\tau_{2t}} & 1 \\ \vdots & \vdots \\ e^{-\kappa\tau_{Nt}} & 1 \end{bmatrix}, \quad d_t = \begin{bmatrix} A(\tau_{1t}) \\ A(\tau_{2t}) \\ \vdots \\ A(\tau_{Nt}) \end{bmatrix} \quad (5.6)$$

Com essas matrizes determinamos os coeficientes da Equação (3.10).

Agora vamos definir as matrizes para a Equação de Transição (3.11) como pode ser visto em (5.7).

$$x_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix}, \quad T_t = \begin{bmatrix} e^{-\kappa\Delta t} & 0 \\ 0 & 1 \end{bmatrix}, \quad c_t = \begin{bmatrix} 0 \\ \mu_\xi \Delta t \end{bmatrix}, \quad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (5.7)$$

O ruído η_t possui média zero em ambas as componentes e matriz de covari-

ância dado por:

$$\text{Cov}[\eta_t] = \begin{bmatrix} (1 - e^{-2\kappa\Delta t})\frac{\sigma_\chi^2}{2\kappa} & (1 - e^{-\kappa\Delta t})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} \\ (1 - e^{-\kappa\Delta t})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} & \sigma_\xi^2\Delta t \end{bmatrix} \quad (5.8)$$

Para este teste já conhecemos todos os parâmetros e os dados iniciais, $x_1 = [\chi_1, \xi_1]$ e P_1 . Após rodar o filtro Kalman com os parâmetros conhecidos obtivemos 24 séries de futuros estimados. Para ilustrar podemos ver na Figura 10 que o gráfico da curva simulada do futuro com vencimento de um mês e da curva estimada estão praticamente sobrepostos o que indica que o filtro funcionou bem.

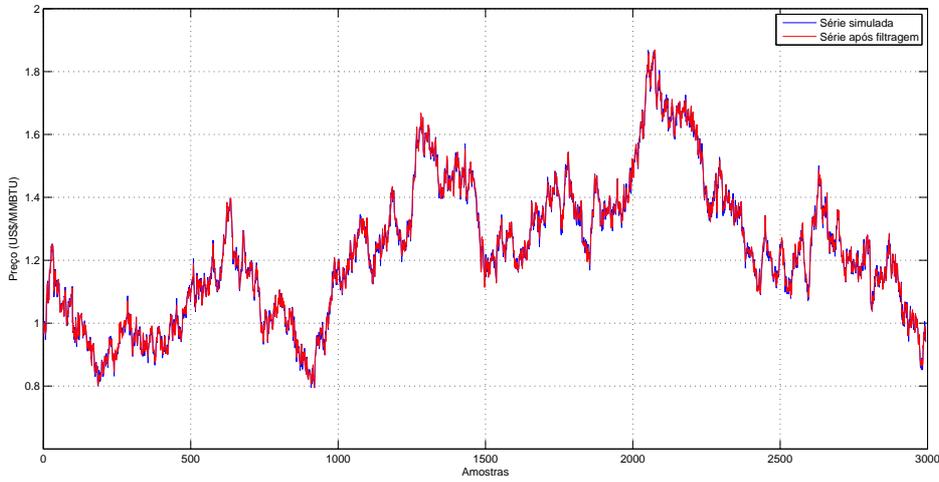


Figura 10: Preços Simulados vs Preços pós Filtragem de Kalman

Para garantir que o filtro funcionou de forma adequada foram feitos alguns testes. O primeiro deles foi calcular o coeficiente de correlação de Pearson para cada uma das séries temporais com maturidades diferentes, logo após foi feito o teste Ljung-Box nos resíduos das séries para checar se aceitam a hipótese de ruído branco, e por último foi feito o teste de Jarque-Bera nos resíduos para checar se aceitam a hipótese de normalidade. A Tabela 2 mostra os resultados. Para maiores detalhes sobre os teste econométricos, consultar a referência [14].

Como pode-se notar a correlação entre as séries simuladas e as séries filtradas apresentam coeficiente bastante alto, além disso em todos os testes de Ljung-Box e Jarque-Bera foi aceita a hipótese nula de ruído branco e normalidade respecti-

T (meses)	ρ	H (Ljung-Box)	H (Jarque-Bera)
1	0,9992	0	0
2	0,9991	0	0
3	0,9989	0	0
4	0,9987	0	0
5	0,9985	0	0
6	0,9984	0	0
7	0,9984	0	0
8	0,9983	0	0
9	0,9982	0	0
10	0,9983	0	0
11	0,9981	0	0
12	0,9980	0	0
13	0,9981	0	0
14	0,9979	0	0
15	0,9980	0	0
16	0,9981	0	0
17	0,9980	0	0
18	0,9980	0	0
19	0,9981	0	0
20	0,9980	0	0
21	0,9981	0	0
22	0,9981	0	0
23	0,9981	0	0
24	0,9982	0	0

Tabela 2: Testes para o Filtro de Kalman

vamente. Destes resultados podemos concluir que o filtro foi implementado com sucesso.

5.4 Calibragem do Modelo

O processo de calibração do modelo resume-se no seguinte problema de otimização: Encontrar o vetor de hiperparâmetros $\Theta = [\kappa, \lambda_\chi, \sigma_\chi, \mu_\xi, \lambda_\xi, \sigma_\xi, \rho_{\chi\xi}, \chi_0, \xi_0, s_1, \dots, s_N]$ que minimize a função de verossimilhança dada por (3.17) dadas as restrições em (5.9).

$$\min_{\Theta} \ln(L[y|\Theta]) \quad s.a. \quad \left\{ \begin{array}{l} -\infty < \chi_0 < \infty \\ -\infty < \xi_0 < \infty \\ 0 \leq \kappa < \infty \\ -\infty < \lambda_\chi < \infty \\ 0 \leq \sigma_\chi < \infty \\ -\infty < \mu_\xi < \infty \\ -\infty < \lambda_\xi < \infty \\ 0 \leq \sigma_\xi < \infty \\ -1 \leq \rho_{\chi\xi} \leq 1 \\ 0 \leq s_i < \infty \end{array} \right. \quad (5.9)$$

O primeiro passo foi simular os preços futuros exatamente como na seção 3.4. Após a simulação escolhe-se arbitrariamente um vetor de hiperparâmetros respeitando as condições expostas em (5.9). Após isso, implementa-se o filtro de kalman para calcular a função de verossimilhança ($L[y|\Theta]$) e seu gradiente com uma aproximação numérica no ponto escolhido arbitrariamente. Feito isto foi utilizada a função do Matlab *fmincon* para obter o vetor de hiperparâmetros ótimo.

Neste problema há um total de trinta e três parâmetros. Dois de condição inicial (χ_0 e ξ_0), sete dos processos χ_t e ξ_t e vinte e quatro variâncias dos ruídos de observação de cada série de futuro (s_i). De modo a tornar o problema mais palpável foi adotada uma hipótese simplificadora onde a variância do ruído de observação de

Parâmetro	Valor Real	Valor Estimado	Erro
χ_0	0.0000	0.0013	0.0005
ξ_0	1.0000	0.9874	0.0001
κ	1.4900	1.4895	0.0048
λ_χ	0.1570	0.1357	0.0294
σ_χ	0.2860	0.2897	0.0011
μ_ξ	-0.0125	-0.0189	0.0087
λ_ξ	-0.0240	-0.0115	0.0099
σ_ξ	0.1450	0.1478	0.0043
$\rho_{\chi\xi}$	0.3000	0.2594	0.0356
s_i	0.0100	0.0054	0.0023

Tabela 3: Resultado da Otimização para Dados Simulados

cada série de futuros foi considerado igual. Esta premissa é muito razoável pois em [3] estes ruídos são interpretados como erros na medida de preços, que por sua vez são iguais para todas as maturidades. Deste modo, o problema resume-se a calibrar o modelo com dez parâmetros.

O código implementado no Matlab leva cerca de 15 minutos para dar a resposta do vetor de hiperparâmetros ótimo. O resultado da otimização pode ser visto na Tabela 3.

O erro das estimativas foi calculado executando-se a rotina de otimização diversas vezes e ao final calculando-se o desvio-padrão dos parâmetros.

Para garantir a robustez da rotina de otimização implementada foram iniciadas diversas vezes o algoritmo com perturbações aleatórias no vetor de hiperparâmetros ótimos com magnitudes diferentes (0.1%, 1.0%, 10.0%, 100.0% e 1000.0%). Os resultados obtidos foram muito semelhantes, o que nos leva a crer que o ponto atingido é próximo do máximo global.

O próximo passo foi calibrar o modelo para dados reais. O procedimento adotado foi exatamente igual ao realizado com dados simulados. O resultado pode ser visto na Tabela 4.

Parâmetro	Valor Estimado	Erro
χ_0	-0.7757	0.0078
ξ_0	0.5438	0.0034
κ	3.3501	0.0295
λ_χ	-0.3678	0.1876
σ_χ	0.5734	0.0235
μ_ξ	0.0012	0.0013
λ_ξ	-0.0150	0.0034
σ_ξ	0.3267	0.0135
$\rho_{\chi\xi}$	-0.2649	0.0781
s_i	0.0456	0.0008

Tabela 4: Resultado da Otimização para Dados Reais

5.5 Aplicação para Opções Reais

Para ilustrar uma das possibilidades de uso do modelo de Schwartz-Smith será aplicada a teoria de opções reais para valorar uma opção de espera para entrada em um investimento num poço de exploração de Gás Natural. A teoria de opções reais configura uma alternativa moderna para a análise econômica de projetos e decisões de investimentos sob incerteza.

O método clássico da análise de projetos por Valor Presente Líquido (VPL) muitas vezes não leva em conta diversos fatores, tais como: Incertezas no preço de commodities que estão relacionadas com a receita do projeto, valor da opção de adiar o começo do projeto, tempo ótimo para iniciar o projeto em questão. Podemos citar algumas referências que tratam do assunto. Desde o trabalho embrionário de Tourinho [15] que trata da valoração de uma mina de cobre pela abordagem de opções reais até trabalhos mais recentes como o livro do Dixit e Pindyck [16] que trata da teoria de investimentos sob incerteza.

O principal resultado utilizado nesta aplicação será o valor da opção de espera para começar um investimento em um determinado projeto. Podemos interpreta-lo como uma opção de compra americana, onde o ativo subjacente é o valor presente do projeto e o *strike* é o investimento total para o projeto. A Equação (5.10) mostra este resultado.

	Opção Financeira	Opção Real
S_t	Preço Subjacente	Valor Presente do Projeto
K	<i>Strike</i>	Investimento
T	Tempo de Maturação	Tempo para tomada de decisão
r	Taxa Livre de Risco	Taxa de Desconto do Projeto
σ^2	Variância do Ativo	Variância do Valor do Projeto

Tabela 5: Quadro comparativo entre opções reais e financeiras

$$P(t, V_t; T) = \sup_{t \leq \tau \leq T} \mathbb{E}_t^* [e^{r(t-\tau)} (V_\tau - I)^+] \quad (5.10)$$

onde: V_t é o valor do projeto, I é o investimento para viabilizar o projeto, r taxa de juros livre de risco e T é o tempo de maturação do projeto.

Podemos observar na Tabela 5 um paralelo entre opções reais e opções financeiras.

O modelo de Schwartz-Smith após sua calibração com dados reais é capaz de extrair as componentes não observáveis que compõe o preço *Spot* da commodity em estudo, neste caso o Gás Natural (*Henry-Hub*). De posse do preço *Spot* (S_t) saberemos com exatidão qual é o valor esperado em termos de receita deste projeto de exploração de Gás Natural fictício, pois a receita do mesmo está diretamente relacionada à quantidade de volume produzido pelo campo de produção multiplicado pelo seu preço *Spot*.

Agora para calcular o valor da opção de espera vamos utilizar o modelo de Longstaff-Schwartz [17] que consiste em simular diversas vezes o valor do ativo subjacente, calcular o payoff no vencimento da opção para cada simulação e em cada uma das simulações aplicar o algoritmo *Least-Squares Monte Carlo* (LSM) para aproximar a esperança condicional dos fluxos de caixa descontados do valor da opção subsequente dado o estado atual.

Por questões de simplificação vamos considerar um investimento fixo $I = 1,510^9$ e $V_t = 1,010^9 \times S_t$, sendo que S_t é uma simulação do preço *Spot* do modelo de Schwartz-Smith calibrado com dados reais segundo a Seção 5.4 e $r = 6\%$ a taxa de juros livre de risco.

Após a aplicação do modelo de Longstaff-Schwartz o resultado do valor da

opção de espera foi igual a $1,010^9$. Além disso, foram feitas outras simulações para verificar a sensibilidade do valor da opção em função do investimento e da taxa de juros.

A Figura 11 mostra a evolução do preço da opção em função da taxa de juros (r) com o investimento fixo $I = 1,510^9$, onde podemos notar um comportamento decrescente. Este resultado já é esperado pois ao final de todas as simulações é feita uma média dos fluxos de caixa descontados pela taxa de juros que resultam no valor da opção. Avaliando de forma qualitativa podemos dizer que quanto maior for a taxa de juros, maior será seu custo de oportunidade o que resulta em um menor incentivo para realizar o projeto neste momento.

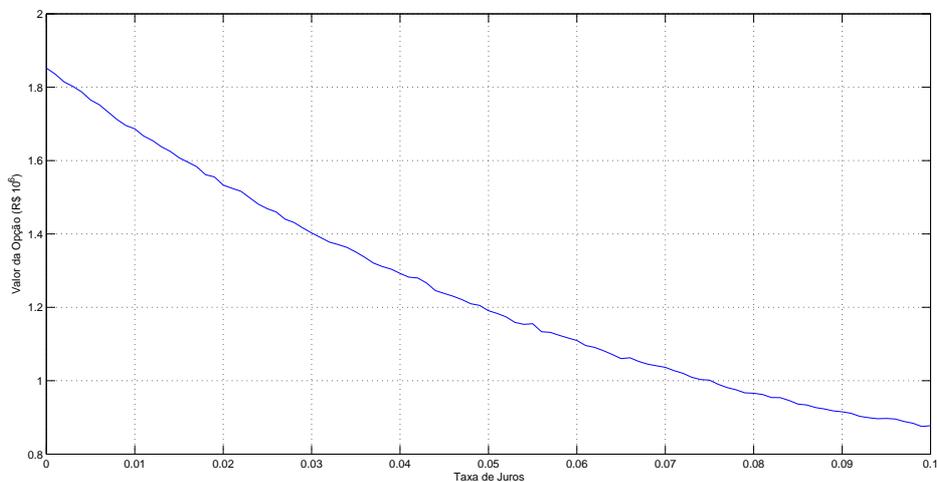


Figura 11: Valor da Opção vs Taxa de Juros

A Figura 12 agora mostra o quanto o preço da opção varia em função do valor do investimento (I) com taxa de juros $r = 6\%$, onde notamos um comportamento decrescente. Este resultado também é esperado pois quanto maior o investimento I a ser feito no projeto torna mais difícil o payoff da opção em qualquer tempo t ser maior que zero e por sua vez ser maior do que o fluxo de caixa futuro, que por se tratar de uma opção de compra americana não seria exercida.

Um dado muito importante quando falamos em opções reais é o tempo de exercício ótimo da opção, que neste exemplo foi de 9,2 anos, ie, o agente deve esperar um pouco mais de nove anos para começar o projeto.

Podemos concluir então que no presente exemplo, o agente deve esperar para

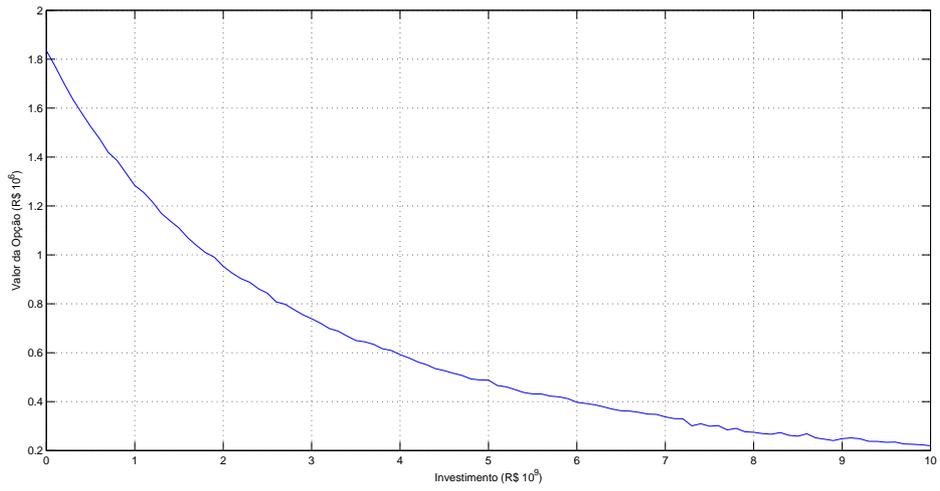


Figura 12: Valor da Opção vs Valor do Investimento

exercer a opção e seguir em frente com o projeto.

6 Conclusão

Este trabalho teve como principal objetivo a calibração do modelo de Schwartz-Smith com a utilização do Filtro de Kalman. Para atingir tal objetivo foram estudados diversos tópicos de finanças que envolviam probabilidade, cálculo estocástico e apreçamento neutro ao risco.

Todo o processo começa com a simulação de preços futuros, seguido de alguns testes para checar se a mesma foi bem sucedida. Logo em seguida o filtro de Kalman é utilizado para extrair as componentes não observáveis. Após isso a coleta e pré-processamento dos dados reais de preços de Gás Natural (*Henry-Hub*) foi essencial para obter bons resultados.

Passada a parte de coleta de dados foi calculada a função de verossimilhança do filtro de Kalman para dados simulados de modo a testar o algoritmo otimizador que de fato calibra o modelo de Schwartz-Smith. Uma vez tendo sucesso nesta parte os dados reais foram utilizados para extraírmos os parâmetros ótimos que governam a dinâmica dos futuros do *Henry-Hub*.

A implementação do modelo de Schwartz-Smith foi bem sucedida na medida em que os resultados observados na simulação dos preços futuros e os resultados teóricos foram coerentes. A implementação do filtro de Kalman mostrou resultados condizentes com as hipóteses de normalidade, ruído branco e correlação entre as séries simuladas e as filtradas. O processo de calibração se mostrou robusto tanto para os dados simulados quanto para os dados reais, no entanto mostrou convergência lenta o que inviabilizaria a utilização do mesmo para aplicações em tempo real.

Por fim, neste trabalho também foi feita uma aplicação em opções reais para dados fictícios o que resultou no preço de uma opção de adiar o início do investimento

em um projeto de exploração de Gás Natural.

Como trabalhos futuros podemos citar algumas melhorias, tais como simular e calibrar o modelo de Schwartz-Smith para diferentes parâmetros e para outras commodities diferentes do Gás Natural. Assegurando assim a robustez do método tanto para parâmetros diferentes quanto para commodities de natureza distintas.

Referências

- [1] E. Schwartz K. Miltersen. Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis*, 1998.
- [2] P. Seguin e M. Smoller H. Bessembinder, J. Coughenour. Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *The Journal of Finance*, 50:361–375, 1995.
- [3] E. Schwartz e J. Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46:893–911, 2000.
- [4] L. Müller. Mathematical methods in finance: Modeling and numerical analysis. Tese de doutorado, IMPA, Julho 2009.
- [5] E. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52:923–973, 1997.
- [6] R. Korn e E. Korn. *Option Pricing and Portfolio Optimization*. American Mathematical Society, 2001.
- [7] M. Manoliu e S. Tompaidis. Energy futures prices: Term structure models with kalman filter estimation. Report, Universidade do Texas, Novembro 2000.
- [8] A. Carvalho. Calibration of the schwartz-smith model for commodity prices. Dissertação de mestrado, IMPA, Março 2010.
- [9] F. Aiube. Modelagem dos preços futuros de commodities: Abordagem por filtro de partículas. Tese de doutorado, PUC Rio, Julho 2005.
- [10] S. Shreve. *Stochastic Calculus for Finance II. Continuous-Time Models*. Springer, 8 edition, 2008.
- [11] D. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *Society for Industrial and Applied Mathematics*, 43:525–546, 2001.
- [12] P. Kloeden e E. Platen. *Numerical Solution of Stochastic Differential Equations (Stochastic Modelling and Applied Probability)*. Springer, 1992.
- [13] A. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1994.
- [14] R. Tsay. *Analysis of Financial Time Series*. Wiley-Interscience, 2nd edition, 2005.

- [15] O. Tourinho. The option value of reserves of natural resources. Working paper, University of California Berkeley, 1979.
- [16] R. Pindyck A. Dixit. *Investment under Uncertainty*. Princeton University Press, 1994.
- [17] E. Schwartz F. Longstaff. Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14:113–147, 2001.

APÊNDICE A – Distribuição Log-Normal

Seja $Y = \ln(X)$ uma variável aleatória que possui distribuição normal de média μ e variância σ^2 . Então pela definição de log-normalidade X será uma variável aleatória com distribuição log-normal.

Sabemos que se Y é normal então sua densidade de probabilidade é dada por A.1.

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad (\text{A.1})$$

Pela definição da densidade de probabilidade temos que $\int_{-\infty}^{\infty} f(y)dy = 1$. Para achar a densidade de uma v.a. log-normal basta usar a relação $y = \ln(x)$.

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = \frac{1}{\sqrt{2\pi}\sigma} \int_0^{\infty} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \frac{dx}{x} = 1 \quad (\text{A.2})$$

De posse da densidade de X conforme A.2, podemos calcular $\mathbb{E}[X]$.

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{\sqrt{2\pi}\sigma} \int_0^{\infty} x e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} \frac{dx}{x} \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_0^{\infty} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^y e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ &= e^{\mu + \frac{\sigma^2}{2}} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(y-(\mu+\sigma^2))^2}{2\sigma^2}} dy \\ &= e^{\mu + \frac{\sigma^2}{2}} \end{aligned} \quad (\text{A.3})$$

Por A.3 concluímos que $\ln(\mathbb{E}[X]) = \mathbb{E}[Y] + 1/2\text{Var}[Y]$.

APÊNDICE B – Distribuição conjunta de χ_t e ξ_t

Neste Apêndice desenvolveremos as Equações de Transição do Filtro de Kalman passo a passo para demonstrar que a distribuição conjunta dos processos χ_t e ξ_t é normal. A demonstração segue a referência [3].

A idéia principal é achar o vetor de média e covariância para a aproximação em tempo discreto para os processos χ_t e ξ_t . Após isso basta aplicar o limite para passos no tempo cada vez menores. A aproximação discreta dos processos com passo de tamanho $\Delta t = t/n$ pode ser escrita como $\mathbf{x}_t = \mathbf{c} + \mathbf{Q}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t$, onde $\mathbf{x}_t \equiv [\chi_t, \xi_t]$, $\mathbf{c} \equiv [0, \mu_\xi \Delta t]$,

$$\mathbf{Q} \equiv \begin{bmatrix} \phi & 0 \\ 0 & 1 \end{bmatrix},$$

onde $\phi \equiv 1 - \kappa \Delta t$, e $\boldsymbol{\eta}_t$ é um vetor 2×1 de perturbações normalmente distribuídas e serialmente descorrelacionadas com $\mathbb{E}[\boldsymbol{\eta}_t] = 0$, e

$$\text{Var}[\boldsymbol{\eta}_t] = \mathbf{W} \equiv \begin{bmatrix} \sigma_\chi^2 \Delta t & \rho_{\chi\xi} \sigma_\chi \sigma_\xi \Delta t \\ \rho_{\chi\xi} \sigma_\chi \sigma_\xi \Delta t & \sigma_\xi^2 \Delta t \end{bmatrix}.$$

Com esse processo, o vetor de média n-passos a frente (\mathbf{m}_n) e a matriz de covariância (\mathbf{V}_n) são dados recursivamente por $\mathbf{m}_n = \mathbf{c} + \mathbf{Q}\mathbf{m}_{n-1}$ e $\mathbf{V}_n = \mathbf{Q}\mathbf{V}_{n-1}\mathbf{Q}' + \mathbf{W}$, com $\mathbf{m}_0 = \mathbf{x}_0 \equiv [\chi_0, \xi_0]$ e $\mathbf{V}_0 = 0$. Aplicando esta recursão chegamos em

$$\mathbf{m}_n = \begin{bmatrix} \phi^n \chi_0 & \xi_0 + \mu_\xi n \Delta t \end{bmatrix}',$$

$$\mathbf{V}_n = \begin{bmatrix} \sigma_\chi^2 \Delta t \sum_{i=0}^{n-1} \phi^{2i} & \rho_{\chi\xi} \sigma_\chi \sigma_\xi \Delta t \sum_{i=0}^{n-1} \phi^i \\ \rho_{\chi\xi} \sigma_\chi \sigma_\xi \Delta t \sum_{i=0}^{n-1} \phi^i & \sigma_\xi^2 n \Delta t \end{bmatrix}.$$

Podemos então re-escrever a série geométrica em \mathbf{m}_n e \mathbf{V}_n , usando

$$\sum_{i=0}^{n-1} \phi^i = \frac{1 - \phi^n}{1 - \phi} \quad e \quad \sum_{i=0}^{n-1} \phi^{2i} = \frac{1 - \phi^{2n}}{1 - \phi^2}$$

Os erros nesta aproximação discreta são de ordem menor que Δt . Assim, se tomarmos o limite com $n \rightarrow \infty$ e $\Delta t = t/n \rightarrow 0$, então $\phi^n = (1 - \kappa t/n)^n$ converge para $e^{-\kappa t}$, ϕ^{2n} converge para $e^{-2\kappa t}$, e

$$\frac{1 - \phi^n}{1 - \phi} \rightarrow \frac{1 - e^{-\kappa t}}{\kappa} \quad e \quad \frac{1 - \phi^{2n}}{1 - \phi^2} \rightarrow \frac{1 - e^{-2\kappa t}}{2\kappa}$$

Substituindo estas formas limites nas expressões de \mathbf{m}_n e \mathbf{V}_n chegamos no vetor de média e matriz de covariâncias das Equações (4.13) e (4.14).

APÊNDICE C – Figuras

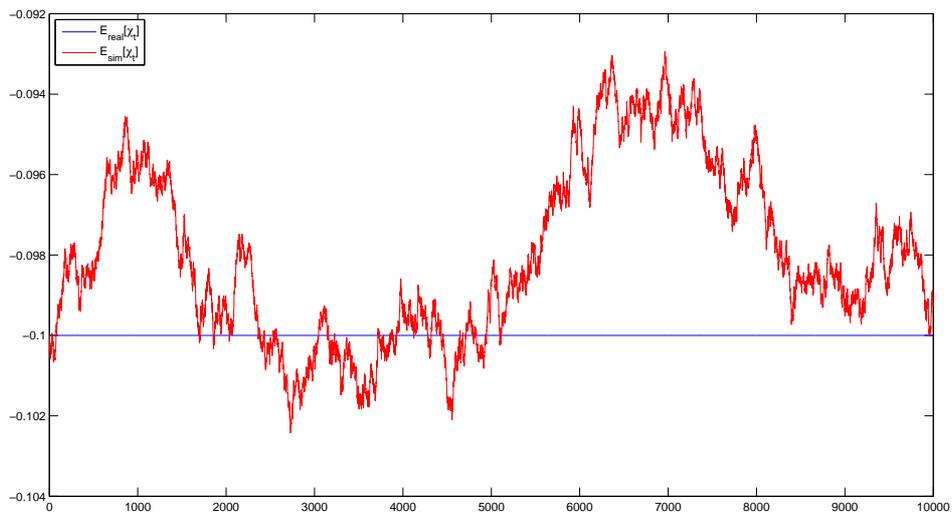


Figura 13: Média do Processo χ_t com 1000 simulações e 10000 amostras

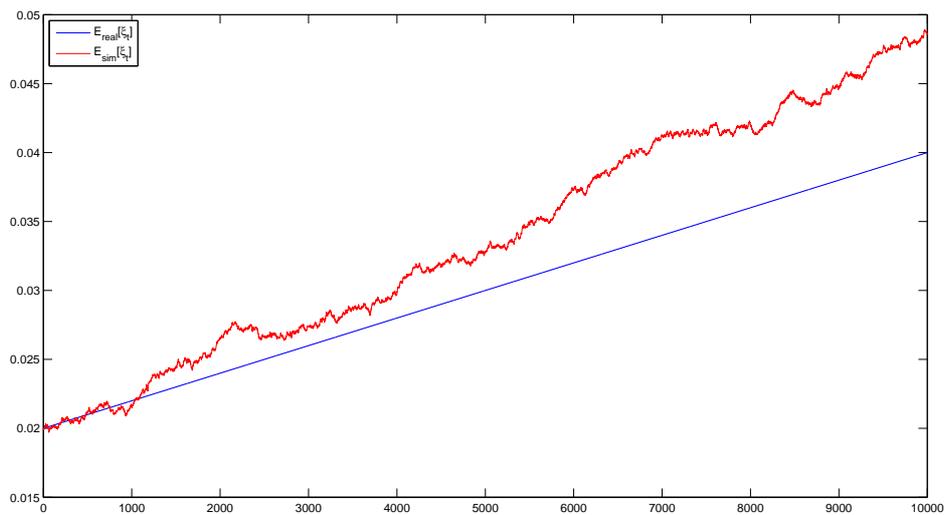


Figura 14: Média do Processo ξ_t com 1000 simulações e 10000 amostras

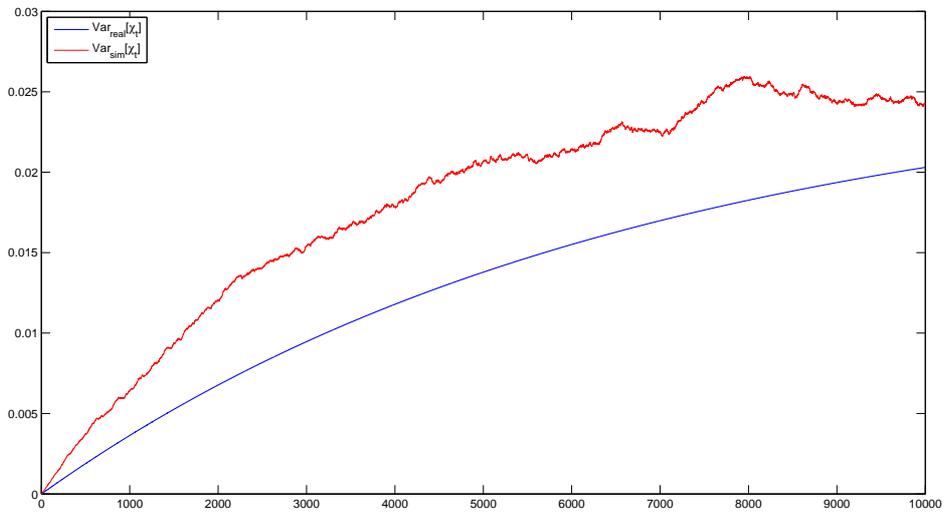


Figura 15: Variância do Processo χ_t com 1000 simulações e 10000 amostras

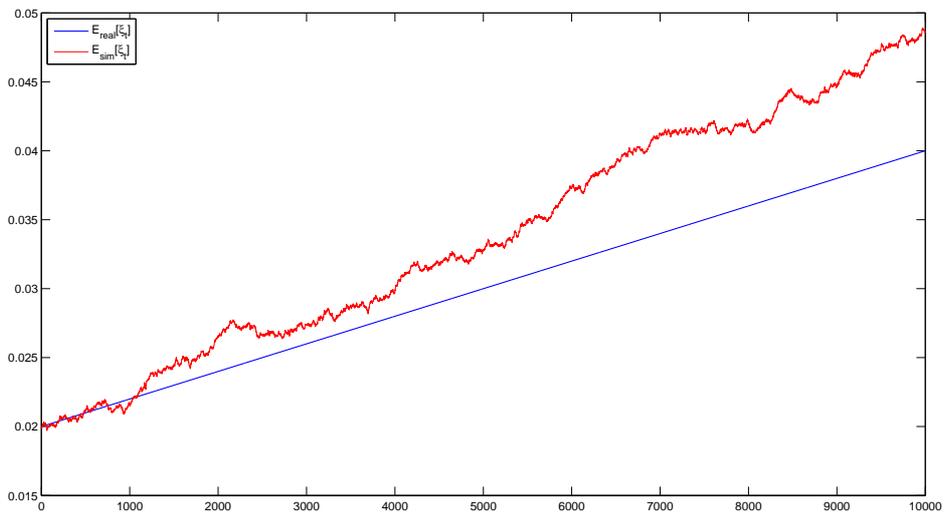


Figura 16: Variância do Processo ξ_t com 1000 simulações e 10000 amostras

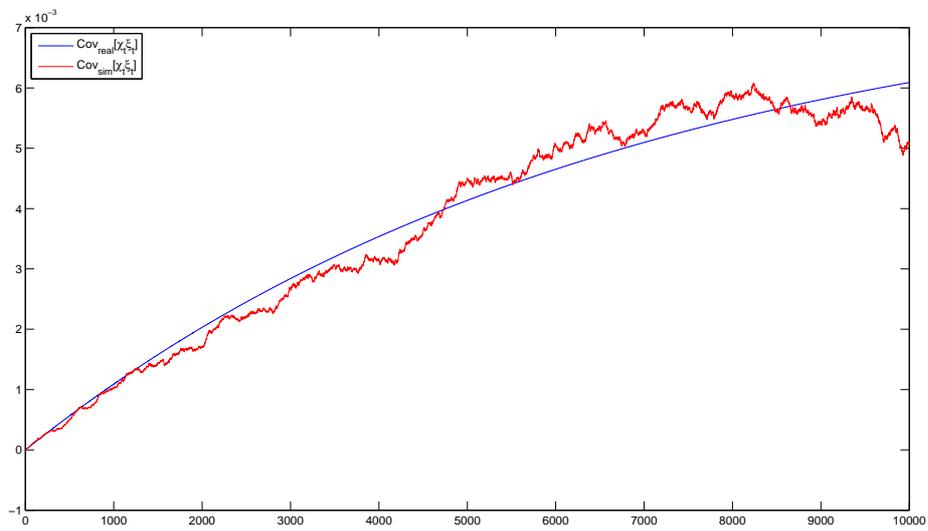


Figura 17: Covariância dos Processos χ_t e ξ_t com 1000 simulações e 10000 amostras

ANEXO A – Códigos

A.1 Arquivo simChiXi2.m

```
function [chi,xi,rho] = simChiXi2(N, dt, chi_0, xi_0, kappa, lambda_chi,
sigma_chi, mu_xi, lambda_xi, sigma_xi, chart)

% ----- %
% Esta funcao simula os processos chi e xi do modelo de
% Schwartz-Smith
% ----- %

% variancia do processo chi
var_chi = sigma_chi^2/(2*kappa)*(1-exp(-2*kappa*dt));

% variancia do processo xi
var_xi = sigma_xi^2*dt;

% gerando os ruidos correlacionados
Z1 = sqrt(var_chi)*randn(N,1);
Z2 = sqrt(var_xi)*randn(N,1);

rho_dt = corrccoef(Z1,Z2);
rho = rho_dt(2,1)/dt;

% criando chi
chi = zeros(1,N);
chi(1) = chi_0;

for i = 1:1:N-1
    chi(i+1) = (1-kappa*dt)*chi(i) + Z1(i);
end

% criando xi
xi = zeros(1,N);
xi(1) = xi_0;

for i = 1:1:N-1
    xi(i+1) = mu_xi*dt + xi(i) + Z2(i);
end

if chart == 1
    figure
    plot(chi, 'LineWidth',1)
    hold on
    plot(xi, 'r', 'LineWidth',1)
    legend('\chi_{t} - Desvio', '\xi_{t} - Equilírio')
    xlabel('Amostras')
end

end
```

A.2 Arquivo simSSfutPrices.m

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Este código visa simular o modelo de Schwartz–Smith para uma commodity %
% qualquer, além de calcular o preço do futuro gerado pelo modelo de SS. %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Fut] = simSSfutPrices(tau, kappa, lambda_chi, sigma_chi, mu_xi,
lambda_xi, sigma_xi, rho, chi, xi)

% pegando o tamanho de observações do processo chi
[M,L] = size(chi);

% adequando o tamanho do tau
tau = tau(end-(L-1):end,:);

% inicializando a matriz dos preços futuros
Fut = zeros(24,L);

for j = 1:1:24

    % calculando A em partes

    A = (mu_xi - lambda_xi)*tau(:,j);
    A = A - (1-exp(-kappa*tau(:,j)))*(lambda_chi/kappa);
    A = A + 1/2*((1-exp(-2*kappa*tau(:,j)))*(sigma_xi^2/(2*kappa)));
    A = A + 1/2*(sigma_xi^2*tau(:,j));
    A = A + (1-exp(-kappa*tau(:,j)))*rho*sigma_chi*sigma_xi/kappa;

    % calculo do log Futuro
    logFut = exp(-kappa*tau(:,j)).*chi' + xi' + A;
    %logFut = A;
    % calculo Futuro
    Fut(j,:) = exp(logFut);

end

end

```

A.3 Arquivo inicialsimSSfutPrices2.m

```

% apagando as variaveis
clear all

% carregando o tau
load 'C:\Users\Marotta\Documents\IMPA\Projeto Final\Dados\tau'

% parametros da simulação
kappa = 1.49;
lambda_chi = 0.157;
sigma_chi = 0.286;
mu_xi_star = 0.0115;
mu_xi = -0.0125;
lambda_xi = mu_xi - mu_xi_star;
sigma_xi = 0.145;
%rho = 0.3;

% tamanho das amostras
N = max(size(tau));

% discretizacao do tempo em dias
dt = 1/252;

% valores iniciais
chi_0 = 0;
xi_0 = 1;

% gera a simulacao de chi e xi
[chi,xi,rho] = simChiXi2(N, dt, chi_0, xi_0, kappa, lambda_chi, sigma_chi,
mu_xi, lambda_xi, sigma_xi, 0);

% gera a curva de futuros
[Fut] = simSSfutPrices(tau, kappa, lambda_chi, sigma_chi, mu_xi, lambda_xi,

```

```
sigma_xi, rho, chi, xi);
```

```
figure
surf(Fut)
ylabel('Maturidade (Meses)')
xlabel('Tempo (Dias)')
```

A.4 Arquivo testKalman.m

```
% script para a filtragem de kalman e otimização

% simula o preço dos futuros
inicialsimSSfutPrices2

Var_y = 0.0001;
u_t = sqrt(Var_y)*randn(max(size(Fut)),min(size(Fut)));

% equacao das observacoes
% observation equation
y = log(Fut)' + u_t;

% matriz de covariancia de u_t
H = cov(u_t);

% criando a matriz Z, sem os 1s na segunda coluna
Z = exp(-kappa*tau)';

% criando d
A = (mu_xi - lambda_xi)*tau;
A = A - (1-exp(-kappa*tau))*(lambda_chi/kappa);
A = A + 1/2*((1-exp(-2*kappa*tau))*(sigma_xi^2/(2*kappa)));
A = A + 1/2*(sigma_xi^2*tau);
A = A + (1-exp(-kappa*tau))*rho*sigma_chi*sigma_xi/kappa;
d = A'; clear A;

figure
surf(y')
ylabel('Maturidade (Meses)')
xlabel('Tempo (Dias)')

% equacao de estado
% state equation
T = [exp(-kappa*dt) 0 ; 0 1];
c = [0 ; mu_xi*dt];
R = [1 0 ; 0 1];

% gerando os ruidos correlacionados novamente
var_chi = sigma_chi^2/(2*kappa)*(1-exp(-2*kappa*dt));
var_xi = sigma_xi^2*dt;
neta1 = sqrt(var_chi)*randn(N,1);
neta2 = sqrt(var_xi)*randn(N,1);

Q = cov(neta1, neta2);

x_0 = [chi_0 ; xi_0];
P_0 = [0 0 ; 0 0];

% inicializando o filtro de kalman
x_e = zeros(2,N);
y_e = zeros(24,N);

x_e(:,1) = x_0;
y_e(:,1) = [Z(:,1) ones(min(size(y)),1)]*x_e(:,1) + d(:,1);

P = T*P_0*T' + R*Q*R';

F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;

lnL = zeros(max(size(tau)),1);
% ln da funcao de verossimilhança
lnL(1) = -min(size(tau))*max(size(tau))/2*log(2*pi)-1/2*log(det(F))
```

```

-1/2*(y(1,:)'-y_e(:,1))'/F*(y(1,:)'-y_e(:,1));

for i = 2:1:N
    x_e(:,i) = T*x_e(:,i-1) + c;
    P = T*P*T' + R*Q*R';
    K = P*[Z(:,i) ones(min(size(y)),1)]'/([Z(:,i) ones(min(size(y)),1)]*P
        *[Z(:,i) ones(min(size(y)),1)]' + H);
    x_e(:,i) = x_e(:,i) + K*(y(i,:) - [Z(:,i) ones(min(size(y)),1)]
        *x_e(:,i) - d(:,i));
    P = (eye(2,2) - K*[Z(:,i) ones(min(size(y)),1)])*P;
    y_e(:,i) = [Z(:,i) ones(min(size(y)),1)]*x_e(:,i) + d(:,i);
    F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;
    lnL(i) = lnL(end)-1/2*log(det(F))-1/2*(y(i,:)'-y_e(:,i))'/F*(y(i,:) -
        -y_e(:,i));
end

% fazendo testes para checar se o filtro está adequado
y = y'; % modificando a dimensão do y para ficar igual ao do y_e

H_lbq = [];
% teste de Ljung-Box para medir se o resíduo é um ruído branco
for i = 1:1:min(size(y))
    H_lbq = [H_lbq lbqtest(y(i,:)-y_e(i,:))];
end

H_jb = [];
% teste de Jarque-Bera para medir se o resíduo é normal
for i = 1:1:min(size(y))
    H_jb = [H_jb jbstest(y(i,:)-y_e(i,:))];
end

corr = [];
for i = 1:1:min(size(y))
    aux = corrcoef(y(i,:),y_e(i,:));
    corr = [corr aux(2,1)];
end

chart = 0;
if chart == 1
    for i = 1:1:min(size(y))
        figure
        plot(y(i,:))
        hold on
        grid on
        plot(y_e(i,:), 'r')
        legend('Série simulada','Série após filtragem')
        ylabel('Preço (US$/MMBTU)')
        xlabel('Amostras')
    end
end
end

```

A.5 Arquivo logLikeGrad.m

```

function [f,g] = logLikeGrad(theta,tau,y,dt,N)

% argumentos da função
% tau: é argumento de entrada fixo
% theta são os parâmetros iniciais de entrada (total de xxx)
% theta(1) = kappa;
% theta(2) = lambda_chi;
% theta(3) = sigma_chi
% theta(4) = mu_xi;
% theta(5) = lambda_xi;
% theta(6) = sigma_xi;
% theta(7) = rho;
% -----
% theta(8) = s_0;
% -----
% theta(9) = chi_0;
% theta(10) = xi_0;
% ...

```

```

% premissas:
% vou supor que P_0 = 0;
% vou supor que o ruído na série de observações para maturidades diferentes
% é igual

kappa = theta(1);
lambda_chi = theta(2);
sigma_chi = theta(3);
mu_xi = theta(4);
lambda_xi = theta(5);
sigma_xi = theta(6);
rho = theta(7);
s_0 = theta(8);
chi_0 = theta(9);
xi_0 = theta(10);

% ruído das observações
u_t = s_0*randn(max(size(tau)),min(size(tau)));

% matriz de covariância do ruído das observações
H = cov(u_t);

% criando a matriz Z, sem os 1s na segunda coluna
Z = exp(-kappa*tau)';

% criando d
A = (mu_xi - lambda_xi)*tau;
A = A - (1-exp(-kappa*tau))*(lambda_chi/kappa);
A = A + 1/2*((1-exp(-2*kappa*tau))*(sigma_xi^2/(2*kappa)));
A = A + 1/2*(sigma_xi^2*tau);
A = A + (1-exp(-kappa*tau))*rho*sigma_chi*sigma_xi/kappa;
d = A'; clear A;

% equacao de estado
% state equation
T = [exp(-kappa*dt) 0 ; 0 1];
c = [0 ; mu_xi*dt];
R = [1 0 ; 0 1];

% matriz de covariância de neta
var_chi = sigma_chi^2/(2*kappa)*(1-exp(-2*kappa*dt));
var_xi = sigma_xi^2*dt;
cov_chi_xi = (1-exp(-kappa*dt))*rho*sigma_chi*sigma_xi/kappa;

Q = [var_chi cov_chi_xi ; cov_chi_xi var_xi];

% inicialização
x_0 = [chi_0 ; xi_0];
P_0 = [0 0 ; 0 0];

% inicializando o filtro de kalman
x_e = zeros(2,N);
y_e = zeros(24,N);

x_e(:,1) = x_0;
y_e(:,1) = [Z(:,1) ones(min(size(y)),1)]*x_e(:,1) + d(:,1);

P = T*P_0*T' + R*Q*R';

F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;

lnL = zeros(max(size(tau)),1);
% ln da funcao de verossimilhança
lnL(1) = -min(size(tau))*max(size(tau))/2*log(2*pi)-1/2*log(det(F))-
1/2*(y(1,:)'-y_e(:,1))'/F*(y(1,:)'-y_e(:,1));

for i = 2:1:N
    x_e(:,i) = T*x_e(:,i-1) + c;
    P = T*P*T' + R*Q*R';
    K = P*[Z(:,i) ones(min(size(y)),1)]'/([Z(:,i) ones(min(size(y)),1)]*
    P*[Z(:,i) ones(min(size(y)),1)]' + H);
    x_e(:,i) = x_e(:,i) + K*(y(i,:)'- [Z(:,i) ones(min(size(y)),1)]*
    x_e(:,i) - d(:,i));

```

```

P = (eye(2,2) - K*[Z(:,i) ones(min(size(y)),1)])*P;
y_e(:,i) = [Z(:,i) ones(min(size(y)),1)]*x_e(:,i) + d(:,i);
F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;
lnL(i) = lnL(i-1)-1/2*log(det(F))-1/2*(y(i,:)'-y_e(:,i))'/F*(y(i,:)'-
-y_e(:,i));
end

% calculando o gradiente em cada componente
epsilon = 1e-6;
theta_now = [kappa, lambda_chi, sigma_chi, mu_xi, lambda_xi, sigma_xi
, rho];
num_param = 7;
ctrl = eye(num_param,num_param);
g = zeros(1,num_param);

for i = 1:1:num_param
g(i) = (logLike(theta_now + epsilon*ctrl(:,i)',tau,y,dt,N,chi_0,
xi_0,u_t) - logLike(theta_now,tau,y,dt,N,chi_0,xi_0,u_t))/epsilon;
end

f = -lnL(end);

```

A.6 Arquivo optimizeParamGrad.m

```

% script para a filtragem de kalman e otimização
tic
% simula o preço dos futuros
inicialsimSSFutPrices2

Var_y = 0.000001;
%Var_y = 0;
u_t = sqrt(Var_y)*randn(max(size(Fut)),min(size(Fut)));

% equacao das observacoes
% observation equation
y = log(Fut)' + u_t;

% matriz de covariancia de u_t
H = cov(u_t);
%H = zeros(24,24);

% criando a matriz Z, sem os 1s na segunda coluna
Z = exp(-kappa*tau)';

% criando d
A = (mu_xi - lambda_xi)*tau;
A = A - (1-exp(-kappa*tau))*(lambda_chi/kappa);
A = A + 1/2*((1-exp(-2*kappa*tau))*(sigma_xi^2/(2*kappa)));
A = A + 1/2*(sigma_xi^2*tau);
A = A + (1-exp(-kappa*tau))*rho*sigma_chi*sigma_xi/kappa;
d = A'; clear A;

figure
surf(y')
ylabel('Maturidade (Meses)')
xlabel('Tempo (Dias)')

% equacao de estado
% state equation
T = [exp(-kappa*dt) 0 ; 0 1];
c = [0 ; mu_xi*dt];
R = [1 0 ; 0 1];

var_chi = sigma_chi^2/(2*kappa)*(1-exp(-2*kappa*dt));
var_xi = sigma_xi^2*dt;
cov_chi_xi = (1-exp(-kappa*dt))*rho*sigma_chi*sigma_xi/kappa;

Q = [var_chi cov_chi_xi ; cov_chi_xi var_xi];

x_0 = [chi_0 ; xi_0];
P_0 = [0 0 ; 0 0];

```

```

% inicializando o filtro de kalman
x_e = zeros(2,N);
y_e = zeros(24,N);

x_e(:,1) = x_0;
y_e(:,1) = [Z(:,1) ones(min(size(y)),1)]*x_e(:,1) + d(:,1);

P = T*P_0*T' + R*Q*R';

F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;

lnL = zeros(max(size(tau)),1);
% ln da funcao de verossimilhança
lnL(1) = -min(size(tau))*max(size(tau))/2*log(2*pi)-1/2*log(det(F))
        -1/2*(y(1,:)'-y_e(:,1))'/F*(y(1,:)'-y_e(:,1));

for i = 2:1:N
    x_e(:,i) = T*x_e(:,i-1) + c;
    P = T*P*T' + R*Q*R';
    K = P*[Z(:,i) ones(min(size(y)),1)]'/([Z(:,i) ones(min(size(y)),1)]
        *P*[Z(:,i) ones(min(size(y)),1)]' + H);
    x_e(:,i) = x_e(:,i) + K*(y(i,:) - [Z(:,i) ones(min(size(y)),1)]
        *x_e(:,i) - d(:,i));
    P = (eye(2,2) - K*[Z(:,i) ones(min(size(y)),1)])*P;
    y_e(:,i) = [Z(:,i) ones(min(size(y)),1)]*x_e(:,i) + d(:,i);
    F = [Z(:,1) ones(min(size(y)),1)]*P*[Z(:,1) ones(min(size(y)),1)]' + H;
    lnL(i) = lnL(i-1)-1/2*log(det(F))-1/2*(y(i,:)'-y_e(:,i))'/F*(y(i,:) -
        -y_e(:,i));
end

lnL(end)

% teste da function handle para lnL
theta_otimo = [kappa, lambda_chi, sigma_chi, mu_xi, lambda_xi, sigma_xi,
    rho, sqrt(Var_y), chi_0, xi_0];
[f,g] = logLikeGrad(theta_otimo,tau,y,dt,N,chi_0,xi_0,u_t)

% vetor inicial de parâmetros
theta_0 = 2*theta_otimo;

options = optimset('Display','iter','GradObj','on','TolX',1e-12,'TolFun',
    1e-12,'MaxFunEvals',5000,'MaxIter',10000);

[opt fval] = fmincon(@(theta) logLikeGrad(theta,tau,y,dt,N,chi_0,xi_0,u_t),
    theta_0,[],[],[],[-Inf,-Inf,0,-Inf,-Inf,0,-1],[Inf,Inf,Inf
    ,Inf,Inf,Inf,1],[],options);

% fazendo testes para checar se o filtro está adequado
y = y'; % modificando a dimensão do y para ficar igual ao do y_e

H_lbq = [];
% teste de Ljung-Box para medir se o resíduo é um ruído branco
for i = 1:1:min(size(y))
    H_lbq = [H_lbq lbqtest(y(i,:)-y_e(i,:))];
end

H_jb = [];
% teste de Jarque-Bera para medir se o resíduo é normal
for i = 1:1:min(size(y))
    H_jb = [H_jb jbstest(y(i,:)-y_e(i,:))];
end

corr = [];
for i = 1:1:min(size(y))
    aux = corrcoef(y(i,:),y_e(i,:));
    corr = [corr aux(2,1)];
end

chart = 0;
if chart == 1
    for i = 1:1:min(size(y))
        figure

```

```

        plot(y(i,:))
        hold on
        grid on
        plot(y_e(i,:), 'r')
        legend('Série simulada', 'Série após filtragem')
        ylabel('Preço (US$/MMBTU)')
        xlabel('Amostras')
    end
end
toc

```

A.7 Arquivo SSSpotSim.m

```

function S = SSSpotSim(N, dt, chi_0, xi_0, kappa, lambda_chi, sigma_chi,
    mu_xi, lambda_xi, sigma_xi, rho)

[chi, xi] = simChiXi2(N, dt, chi_0, xi_0, kappa, lambda_chi, sigma_chi,
    mu_xi, lambda_xi, sigma_xi, rho, 0);

S = exp(chi+xi);

```

A.8 Arquivo AmOptLSM.m

```

function [Price, CF, S, t] = AmOptLSM(K, r, T, N, M, type, chi_0, xi_0, kappa,
    lambda_chi, sigma_chi, mu_xi, lambda_xi, sigma_xi, rho)

dt = T/N;

t = 0:dt:T;
t = repmat(t', 1, M);

S = zeros(N, M+1);

for j = 1:1:M+1
    R = SSSpotSim(N, dt, chi_0, xi_0, kappa, lambda_chi, sigma_chi,
        mu_xi, lambda_xi, sigma_xi, rho);
    S(:, j) = R';
end

S0 = S(1, 1);

ExTime = (M+1)*ones(N, 1);

CF = zeros(size(S));

CF(end, :) = max(K-S(end, :), 0);

for ii = size(S)-1:-1:2
    if type
        Idx = find(S(ii, :) < K);
    else
        Idx = find(S(ii, :) > K);
    end
    X = S(ii, Idx)'; X1 = X/S0;
    Y = CF(ii+1, Idx)'*exp(-r*dt);
    R = [ ones(size(X1)) (1-X1) 1/2*(2-4*X1-X1.^2)];
    a = R\Y;
    C = R*a;
    if type
        Jdx = max(K-X, 0) > C;
    else
        Jdx = max(X-K, 0) > C;
    end
    nIdx = setdiff((1:M), Idx(Jdx));
    CF(ii, Idx(Jdx)) = max(K-X(Jdx), 0);
    ExTime(Idx(Jdx)) = ii;
    CF(ii, nIdx) = exp(-r*dt)*CF(ii+1, nIdx);
end

```

```
end
```

```
Price = mean(CF(2,:))*exp(-r*dt);  
end
```