

Calibração do Modelo Gibson-Schwartz para dados de *commodities* no Brasil

Leon Kacowicz

Alan De Genaro (Orientador)

Jorge P. Zubelli (Co-Orientador)

Instituto Nacional de Matemática Pura e Aplicada (IMPA)

Dezembro 2012

Resumo

Neste trabalho ajustamos o modelo Gibson-Schwartz, que descreve o comportamento estocástico de preços de *commodities*, para o mercado de boi-gordo negociado na BM&FBOVESPA. A estimação dos parâmetros deste modelo é feita por máxima-verossimilhança, calculada com Filtro de Kalman. Os dados utilizados são séries de preços de contratos futuros desta *commodity* de janeiro de 2003 a fevereiro de 2012, assim como a curva de juros de mercado referente ao mesmo período, extraída dos futuros de DI. Os fatores latentes - preço *spot* e *convenience-yield* - encontrados têm comportamento esperado e os parâmetros estimados são estatisticamente significantes. Em particular, encontramos alto grau de significância estatística para a taxa de reversão à média do *convenience-yield*. Em seguida é feita uma comparação *out-of-sample* dos preços de vencimentos menos líquidos com os preços arbitrados pela bolsa. Propomos, então, este modelo como uma alternativa livre de arbitragem para atribuir preço a tais contratos.

Palavras-chave: Gibson-Schwartz, Filtro de Kalman, *Commodities*

Abstract

In the present work we fit the Gibson-Schwartz model, which describes the stochastic behavior of commodity prices, to live cattle market prices traded at BM&FBOVESPA. Parameter estimation is performed by mean of optimizing the model's likelihood function where the latent factors are extracted using Kalman Filter. The dataset comprises the live cattle future price series ranging from January 2003 to February 2012 and the interest rate term structure for the same period, extracted from DI future contract prices. Our findings can be summarized as follows: estimated parameters are statistically significant in all levels and latent factors - spot price and convenience yield - present the expected behavior. In particular, we found strong statistical significance supporting the mean reversion behavior for the convenience yield. Next, we performed an out-of-sample comparison between model's price and prices assigned by the exchange for less liquid maturities. Therefore our contribution in this literature is to propose an arbitrage-free pricing of live cattle futures contracts.

Keywords: Gibson-Schwartz, Kalman Filter, Commodities

Agradecimentos

Gostaria de agradecer à minha família pelo constante apoio e amor incondicional, à Carol que sempre me apoia e acompanha, ao professor Jorge Zubelli pela preciosa orientação e ajuda necessárias para a conclusão do mestrado, ao meu orientador Alan De Genaro, cuja contribuição foi imprescindível para a produção deste trabalho, e aos meus amigos e colegas de turma Andrew e Bernardo, com quem compartilhei a experiência vivida ao longo deste curso.

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Revisão Bibliográfica	2
1.2.1	Modelo de um fator (Schwartz-97)	2
1.2.2	Modelo de dois fatores (Schwartz-97)	3
1.2.3	Modelo de três fatores (Schwartz-97)	5
1.2.4	Modelo Schwartz-Smith	5
1.2.5	Calibração do Modelo Schwartz-Smith com mínimos quadrados não lineares [Car10]	6
1.2.6	Calibração do Modelo Schwartz-Smith com Filtro de Kalman [Mar11]	7
2	Revisão Teórica - Métodos de Estimação	8
2.1	Espaço de Estado	8
2.2	Filtro de Kalman	8
2.3	Quasi-Máxima-Verossimilhança	10
3	Desenvolvimento da solução	12
3.1	Implementação	12
3.2	Dados utilizados	14
3.3	Teste com dados aleatórios	17
4	Resultados empíricos	21
5	Conclusão	26
A	Códigos implementados	27
A.1	Filtro de Kalman	27
A.2	Função gibson-schwartz	29
A.3	Função otimiza-gibson-schwartz	30

Lista de Figuras

3.1	Preço diário de ajuste dos contratos BGI (boi-gordo)	14
3.2	Volume diário (em R\$) dos contratos BGI (boi-gordo)	15
3.3	Prazos ao vencimento dos contratos utilizados em cada dia	16
3.4	Taxa de juros livre de risco continuamente capitalizada ao ano	17
3.5	Processo X_t teórico e estimado no exercício com dados simulados	19
3.6	Processo δ_t teórico e estimado no exercício com dados simulados	20
4.1	Preço <i>spot</i> estimado e preço do primeiro futuro BGI (boi-gordo)	22
4.2	<i>Convenience-yield</i> estimado para os futuros BGI (boi-gordo)	23
4.3	Preço calculado e preço divulgado do quarto futuro BGI (boi-gordo)	24
4.4	Preço calculado e preço divulgado do quinto futuro BGI (boi-gordo)	25

Lista de Tabelas

3.1	Parâmetros reais e estimados utilizados no exercício com dados simulados . . .	18
4.1	Parâmetros encontrados para o conjunto completo de dados	21
4.2	Parâmetros encontrados para os dados <i>in sample</i>	24

Capítulo 1

Introdução

1.1 Motivação

A modelagem estocástica para a dinâmica de preços de *commodities* tem sido recorrentemente objeto de estudo acadêmico devido a sua enorme importância para o apreçamento dos respectivos derivativos, assim como no auxílio à avaliação de investimentos para sua produção ou extração.

Um dos motivos que fazem com que esse assunto não seja trivial é o fato de que, embora as *commodities* também sejam negociadas à vista, a formação de preços ocorre no mercado futuro. Portanto o preço *spot* não é diretamente observável, o que torna mais complexa a estimação dos parâmetros dos modelos.

Os primeiros trabalhos publicados sobre este assunto - [BS85], por exemplo - costumavam tomar como hipótese que o *convenience yield* e taxas de juros fossem constantes. Sob estas hipóteses, o preço *spot* das *commodities* se comporta como um passeio aleatório descrito por um movimento geométrico browniano, tal como na modelagem de Black-Scholes para opções. Embora essas hipóteses simplifiquem a modelagem, e tornem o processo de cálculo de contratos futuros e opções mais simples e intuitivo, elas claramente não são satisfatórias para modelar o comportamento real dos preços e não reproduzem os fatos estilizados sobre séries de preços de *commodities*, pois, sob esta abordagem, espera-se que os preços das *commodities* cresça a uma taxa constante e que a variância do preço *spot* em instantes de tempo futuros cresça proporcionalmente ao tempo.

Pode-se justificar, do ponto de vista econômico, um comportamento de reversão à média dos preços de algumas *commodities*. Caso o preço esteja muito acima de um patamar de equilíbrio, a produção ou extração desta mercadoria seria mais lucrativa e geraria lucro até para produtores que incorrem em maiores custos. Isso justificaria a entrada destes produtores no mercado. Com este aumento na oferta, o preço tenderia a diminuir, tudo o mais constante. Por outro lado, se os preços estiverem muito abaixo daquele patamar de equilíbrio, estes produtores menos eficientes passarão a sofrer prejuízos. Isso os fará suspender a produção, diminuindo assim a oferta, e conseqüentemente aumentando os preços. Como estas realocações de capital levam algum tempo para ocorrer, é esperado que os preços fiquem desviados do valor de equilíbrio por algum tempo até retornarem a este patamar.

Trabalhos mais recentes têm levado em consideração comportamentos como o caráter estocástico e reversão à média do *convenience yield* ([GS90]), dinâmica estocástica da taxa de juro sem-risco ([Sch97]), e reversão à média no curto prazo e incerteza do preço de equilíbrio no longo-prazo ([SS00]).

Frequentemente os trabalhos acadêmicos sobre este tema utilizam os modelos propostos no apoio à decisão de investimentos na produção ou extração da *commodity* utilizando a abordagem de opções-reais.

Neste trabalho, tentaremos ajustar o modelo Gibson-Schwartz utilizando a abordagem proposta em [Sch97] ao mercado de boi-gordo da BM&FBOVESPA. Utilizaremos o modelo ajustado para os vencimentos com mais liquidez e tentaremos extrapolar os resultados com o objetivo de apreçar vencimentos com menor liquidez. Em seguida proporemos como outra aplicação uma estratégia de arbitragem estatística.

Embora o modelo Schwartz-Smith seja matematicamente equivalente ao Gibson-Schwartz e considerado melhor sob vários aspectos, que serão discutidos mais adiante, utilizaremos o modelo Gibson-Schwartz. A razão para isso reside no fato de que este permite que utilizemos explicitamente a curva de juros de mercado na estimação dos outros parâmetros, em vez de utilizar uma taxa de juros igual para todos os vencimentos e invariante no tempo. Podemos fazer isso, pois os contratos futuros de DI na BM&FBovespa são suficientemente líquidos para considerarmos não apenas conhecida a taxa de juros para cada vencimento, como também considerá-la "hedgeável". Isso nos poupa de ter que utilizar um modelo com taxa de juro estocástica, como, por exemplo, o modelo de três fatores explorado em [Sch97], que seria muito mais complexo. Schwartz argumenta que utilizar uma taxa de juros fixa não prejudica significativamente o modelo pois quaisquer oscilações na taxa de juros serão capturadas pelo processo estocástico do *convenience yield*. Por outro lado, utilizando a curva de juros de mercado, podemos garantir que esse efeito não ocorrerá, e que o processo estimado do *convenience yield* não estará contaminado por variações dos juros.

1.2 Revisão Bibliográfica

Nesta seção apresentamos os principais trabalhos e modelos a respeito do tema e relevantes para o trabalho.

Em [Sch97], Eduardo Schwartz faz um estudo expondo três modelos de preços para *commodities* utilizando a seguinte abordagem: é proposto, primeiramente, cada modelo na medida real. A seguir é apresentado o modelo na medida neutra ao risco equivalente. Depois, é proposto o uso do filtro de Kalman, para se estimar o valor das variáveis latentes. Finalmente, otimizando-se a verossimilhança do modelo, também calculada através do uso do filtro de Kalman, determinam-se os parâmetros ótimos. Discutiremos o filtro de Kalman em maiores detalhes no capítulo 2.

Neste artigo, Schwartz tenta ajustar os modelos aos preços de futuros de petróleo, cobre e ouro.

1.2.1 Modelo de um fator (Schwartz-97)

O primeiro modelo apresentado, de um fator, define uma dinâmica temporal apenas para o preço spot da *commodity* conforme o seguinte processo estocástico:

$$dS_t = \kappa(\mu - \ln(S_t))S_t dt + \sigma S_t dW_t$$

Onde $t \in \mathbb{R}$, $t \geq 0$, S_t é o preço *spot* do ativo no instante t , σ é W_t é o movimento *browniano*.

Tomando-se $X_t = \ln(S_t)$, e utilizando-se o lema de Itô, obtemos

$$dX_t = \kappa \left(\mu - \frac{\sigma^2}{\kappa} - X_t \right) dt + \sigma dW_t,$$

donde concluímos que S_t é o exponencial de um processo Ornstein-Uhlenbeck com média $\alpha = \mu - \frac{\sigma^2}{\kappa}$ e taxa de reversão à média κ .

O modelo equivalente na medida neutra ao risco pode ser escrito como:

$$dX_t = \kappa(\alpha^* - X_t) dt + \sigma dW_t^*$$

Nesta equação, $\alpha^* = \alpha - \lambda$, em que λ é o prêmio de risco de mercado, considerado constante, e dW_t^* é o incremento do movimento browniano na medida neutra ao risco.

O apreamento de um contrato futuro F_T utilizando-se este modelo é feito da seguinte forma:

$$F_T = \mathbb{E}^*[S_T]$$

Como X_t é um processo Ornstein-Uhlenbeck, sabe-se que, na medida neutra ao risco,

$$X_T \sim N\left(e^{-\kappa T} X_0 + (1 - e^{-\kappa T}) \alpha^*, \frac{\sigma^2}{2\kappa} (1 - e^{-2\kappa T})\right)$$

e portanto

$$\mathbb{E}^*[S_T] = \mathbb{E}^*[\exp(X(T))] = \exp\left(\mathbb{E}^*[X(T)] + \frac{1}{2}\mathbb{V}^*[X(T)]\right)$$

Assim chegamos à seguinte expressão para o preço do contrato futuro:

$$F(S, T) = \exp\left(e^{-\kappa T} \ln(S) + (1 - e^{-\kappa T}) \alpha^* + \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa T})\right)$$

Com este modelo, embora ele não tenha mostrado boa aderência aos dados, Schwartz conseguiu obter evidência estatística de reversão à média tanto para os preços de petróleo quanto para os de cobre. Schwartz destaca ainda que, utilizando-se futuros de petróleo de maturidades mais longas, este modelo mostrou menor taxa de reversão à média. Ele ressalta também que não foi possível ajustar este modelo aos preços dos futuros de ouro, o que indicaria que não há qualquer reversão à média detectável para o período analisado.

1.2.2 Modelo de dois fatores (Schwartz-97)

O segundo modelo apresentado, de dois fatores, é o modelo Gibson-Schwartz, originalmente proposto em [GS90]. Schwartz provavelmente revisitou este modelo, pois no artigo original a metodologia utilizada para estimar os parâmetros era bastante rudimentar se comparada com a metodologia usada neste artigo, e ao mesmo tempo mais trabalhosa, utilizando-se aproximações para os processos latentes. Em [Sch97] esse modelo também foi ajustado para outros ativos além do petróleo.

Neste modelo, assume-se que o preço *spot* e o *convenience yield* seguem conjuntamente o seguinte processo estocástico:

$$dS_t = (\mu - \delta_t)S_t dt + \sigma_1 S_t dW_{1,t} \quad (1.1)$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dW_{2,t} \quad (1.2)$$

$$dW_{1,t}dW_{2,t} = \rho dt$$

Aqui, o *convenience yield* instantâneo (δ_t) segue um processo estocástico do tipo Ornstein-Uhlenbeck de reversão à média, e seu efeito no preço *spot* ocorre no *drift* (termo que multiplica

dt). Nesta modelagem, o *convenience yield* faz um papel análogo ao de um dividendo contínuo numa modelagem de preços de ações. Cabe ressaltar também que, caso o *convenience yield* fosse determinístico da forma $\delta(S) = \kappa \ln S$, este modelo se reduziria ao modelo de um fator.

Novamente, se tomarmos $X_t = \ln S_t$, e aplicarmos o lema de Itô, a primeira equação se reduz a:

$$dX_t = \left(\mu - \delta_t - \frac{\sigma_1^2}{2} \right) dt + \sigma_1 dW_{1,t}$$

Este modelo na medida neutra ao risco pode ser escrito como:

$$\begin{aligned} dS_t &= (r - \delta_t)S_t dt + \sigma_1 S_t dW_{1,t}^* \\ d\delta_t &= [\kappa(\alpha - \delta_t) - \lambda]dt + \sigma_2 dW_{2,t}^* \\ dW_{1,t}^* dW_{2,t}^* &= \rho dt \end{aligned}$$

Na medida neutra ao risco, é introduzido o parâmetro λ que é o prêmio de risco de mercado do *convenience yield*, uma vez que este não pode ser "hedgado". Neste modelo, λ é considerado constante no tempo.

Tomando-se como hipótese a inexistência de oportunidades de arbitragem, o preço de futuros apreçados por este modelo deve obedecer à seguinte equação diferencial parcial:

$$\frac{\sigma_1^2 S^2}{2} \frac{\partial^2 F}{\partial S^2} + \sigma_1 \sigma_2 \rho S \frac{\partial^2 F}{\partial S \partial \delta} + \frac{\sigma_2^2}{2} \frac{\partial^2 F}{\partial \delta^2} + (r - \delta)S \frac{\partial F}{\partial S} + (\kappa(\alpha - \delta) - \lambda) \frac{\partial F}{\partial \delta} - \frac{\partial F}{\partial T} = 0$$

Com condição final $F(S, \delta, 0) = S$.

A solução para essa equação é:

$$F(S, \delta, T) = S \exp \left(-\delta \frac{1 - e^{-\kappa T}}{\kappa} + A(T) \right) \quad (1.3)$$

Onde

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

e

$$\hat{\alpha} = \alpha - \frac{\lambda}{\kappa}$$

Esta demonstração é melhor detalhada em [Bje91].

Utilizando-se este modelo, a aderência aos dados fica claramente melhor. Schwartz destaca que tanto para o petróleo quanto para o cobre, há forte indício de reversão à média, e alta correlação entre os processos estocásticos do preço *spot* e do *convenience yield*. Neste modelo também é obtido menor grau de reversão à média quando são utilizadas maturidades mais longas para os futuros de petróleo.

Ao contrário do anterior, este modelo foi capaz de se ajustar aos preços dos futuros de ouro. No entanto, o grau de reversão à média obtido foi substancialmente menor do que aqueles encontrados para petróleo e cobre, embora mantendo-se estatisticamente significativa. Também foi obtida menor correlação entre os processos estocásticos. Ao contrário dos outros mercados, para o ouro foi encontrado maior grau de reversão à média utilizando-se maturidades mais longas. Schwartz sugere que este pequeno grau de reversão à média pode ser espúrio e esteja, na verdade, capturando uma reversão à média das taxas de juros.

1.2.3 Modelo de três fatores (Schwartz-97)

O terceiro modelo apresentado, de três fatores, é uma generalização do anterior. Neste modelo a taxa de juro de curto prazo deixa de ser constante e passa a ser estocástica, de forma que o preço *spot*, o *convenience yield* e a taxa de juro de curto prazo são modelados pelas seguintes equações na medida neutra ao risco:

$$dS_t = (r_t - \delta_t)S_t dt + \sigma_1 S_t dW_{1,t}^*$$

$$d\delta_t = \kappa(\hat{\alpha} - \delta_t)dt + \sigma_2 dW_{2,t}^*$$

$$dr_t = a(m^* - r_t)dt + \sigma_3 dW_{3,t}^*$$

Onde $dW_{1,t}^* dW_{2,t}^* = \rho_1 dt$, $dW_{2,t}^* dW_{3,t}^* = \rho_2 dt$, $dW_{1,t}^* dW_{3,t}^* = \rho_3 dt$. Neste modelo, a equação para apreamento de futuros é:

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

A expressão para $C(T)$ é extensa e é apresentada em [Sch97]. Este modelo é substancialmente mais complexo do que o modelo Gibson-Schwartz. Neste artigo, o autor estima separadamente os parâmetros para o processo de juros, justificando que o processo que rege a dinâmica da taxa de juros é pouco influenciada pelos demais processos estimados.

Os resultados obtidos para este modelo são muito similares àqueles obtidos para o modelo Gibson-Schwartz para os mercados de petróleo e cobre. Apenas em um dos conjuntos de dados este modelo mostrou aderência substancialmente melhor que o modelo Gibson-Schwartz.

No caso do mercado de ouro, a taxa de reversão à média do *convenience yield* se mostra estatisticamente insignificante, confirmando a hipótese de que este parâmetro no modelo Gibson-Schwartz estava na verdade capturando a reversão à média da taxa de juros. Schwartz conclui, então, que nenhum dos 3 modelos é adequado para a modelagem do mercado de ouro.

1.2.4 Modelo Schwartz-Smith

Em [SS00], Eduardo Schwartz e James Smith propõem um novo modelo, que, embora seja matematicamente equivalente ao modelo Gibson-Schwartz, possui algumas vantagens, que serão discutidas em seguida. Neste modelo, assume-se que:

$$\ln S_t = \chi_t + \xi_t$$

Onde χ_t é o desvio de curto prazo do preço de equilíbrio, e ξ_t é o preço de equilíbrio de longo prazo, que seguem, conjuntamente, o seguinte processo estocástico:

$$d\chi_t = -\kappa\chi_t dt + \sigma_\chi dW_{\chi,t}$$

$$d\xi_t = \mu_\xi dt + \sigma_\xi dW_{\xi,t}$$

Assume-se ainda que $dW_{\chi,t} dW_{\xi,t} = \rho_{\chi\xi} dt$. Sob este modelo, o preço de equilíbrio e o desvio de curto prazo possuem a seguinte distribuição conjunta:

$$\begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} \sim N \left(\begin{bmatrix} e^{\kappa t} \chi_0 \\ \xi_0 + \mu_\xi t \end{bmatrix}, \begin{bmatrix} (1 - e^{-2\kappa t}) \frac{\sigma_\chi^2}{2\kappa} & (1 - e^{-\kappa t}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \\ (1 - e^{-\kappa t}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} & \sigma_\xi^2 t \end{bmatrix} \right)$$

O processo estocástico conjunto na medida neutra ao risco equivalente é dado por:

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dW_{\chi,t}^*$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dW_{\xi,t}^*$$

Onde $dW_{\chi,t}^* dW_{\xi,t}^* = \rho_{\chi\xi} dt$, e λ_ξ e λ_χ são os prêmios de risco de mercado do preço de equilíbrio e do desvio de curto prazo. Na medida neutra ao risco, a distribuição conjunta de χ_t e ξ_t também é normal com mesma matriz de covariâncias da medida verdadeira, mas com as seguintes esperanças condicionais:

$$\mathbb{E}^*[(\chi_t, \xi_t) | (\chi_0, \xi_0)] = \left[e^{-\kappa t} \chi_0 - (1 - e^{-\kappa t}) \frac{\lambda_\chi}{\kappa}, \xi_0 + \mu_\xi^* t \right]$$

Onde $\mu_\xi^* = \mu_\xi - \lambda_\xi$. O apreçamento de futuros, então, é dado por:

$$\ln(F_T) = \ln(\mathbb{E}^*[S_T]) = \mathbb{E}^*[\ln(S_T)] + \frac{1}{2} \mathbb{V}^*[\ln(S_T)]$$

$$F_T = \exp(e^{-\kappa T} \chi_0 + \xi_0 + A(T))$$

$$A(T) = \mu_\xi^* - (1 - e^{-\kappa T}) \frac{\lambda_\chi}{\kappa} + \frac{1}{2} \left((1 - e^{-2\kappa T}) \frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2 T + 2(1 - e^{-\kappa T}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \right)$$

Neste artigo, os autores mostram a equivalência matemática entre este modelo e o modelo Gibson-Schwartz através de uma mudança de variáveis: $\chi_t = \frac{1}{\kappa}(\delta_t - \alpha)$. Com esta mudança de variáveis, é possível reduzir o modelo Gibson-Schwartz ao modelo Schwartz-Smith. Os autores apresentam uma tabela com as expressões que levam os parâmetros do modelo Gibson-Schwartz nos parâmetros equivalentes do modelo Schwartz-Smith.

Como os modelos são formalmente equivalentes e o modelo Schwartz-Smith exige um parâmetro a menos - a taxa de juros -, os autores argumentam que este parâmetro exigido pelo modelo Gibson-Schwartz é redundante e indistinguível. De fato, se a taxa de juro, r , for alterada para $r + \Delta$, o *convenience yield* instantâneo, δ_t , para $\delta_t + \Delta$, o *convenience yield* médio, α , para $\alpha + \Delta$ e o *drift* do preço *spot*, μ , por $\mu + \Delta$, serão encontrados os mesmos valores de S_t e δ_t (descontando-se o deslocamento Δ), e o mesmo valor para a função de verossimilhança. Provavelmente por esse motivo, em [Sch97], o autor utiliza um valor exógeno para r e argumenta que qualquer discrepância entre o valor real de r e o valor escolhido seria absorvido pelo *convenience yield* instantâneo.

As vantagens deste modelo residem no fato de que não há redundâncias entre os parâmetros, e, segundo argumentam os autores, os fatores latentes são mais "ortogonais" em suas dinâmicas, o que levaria a resultados analíticos mais transparentes permitindo simplificar-se a análise de investimentos de longo prazo.

1.2.5 Calibração do Modelo Schwartz-Smith com mínimos quadrados não lineares [Car10]

Em [Car10], Ana Luiza Roriz apresenta uma metodologia alternativa para estimação do Modelo Schwartz-Smith. Este procedimento é feito sem utilização do filtro de Kalman. Em seu lugar é utilizado um algoritmo de mínimos-quadrados não lineares para extração dos fatores latentes do modelo. Em seguida é calculada a verossimilhança do modelo e, a partir disso, são calculadas

novas aproximações para os parâmetros. O procedimento é repetido até que a convergência seja alcançada. As principais vantagens apresentadas para este procedimento são o número reduzido de iterações necessárias para atingir-se a convergência, e a robustez do algoritmo em relação à aproximação inicial fornecida.

Neste trabalho além de dados sintéticos, também são utilizados preços de gás natural (Henry-Hub). O algoritmo apresentado foi capaz de estimar satisfatoriamente os parâmetros para este mercado.

1.2.6 Calibração do Modelo Schwartz-Smith com Filtro de Kalman [Mar11]

Em [Mar11], Leonardo Marotta, faz um estudo relevante sobre a utilização do filtro de Kalman para estimação dos parâmetros do modelo Schwartz-Smith para o mercado de gás natural (Henry-Hub). Além do uso de dados reais, também são feitas simulações de Monte-Carlo para avaliar a acurácia do procedimento para extração dos parâmetros do modelo. Tanto a simulação quanto a extração dos parâmetros para o mercado de gás natural são executados com sucesso e finalmente é apresentada uma aplicação dos parâmetros ao apoio a decisão de investimentos utilizando-se para abordagem de opções reais.

Capítulo 2

Revisão Teórica - Métodos de Estimação

2.1 Espaço de Estado

Para facilitar a implementação do filtro de Kalman, que será explicado em mais detalhes na próxima seção, é conveniente escrevermos as equações que governam a dinâmica dos processos estocásticos na forma de Espaço de Estado. A representação de sistemas em Espaço de Estado permite que uma rica variedade de modelos sejam descritos em função de variáveis não observáveis. Curiosamente, é relativamente difícil encontrar na literatura um padrão recorrente no uso de letras para representar as matrizes e vetores. Aqui utilizaremos uma notação parecida com a utilizada em [Ham94].

Seja um sistema modelado pelas seguintes equações:

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t \quad (2.1)$$

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t \quad (2.2)$$

Em que \mathbf{v}_t e \mathbf{w}_t são vetores aleatórios com alguma distribuição conjunta conhecida. Assumimos ainda que \mathbf{v}_t e \mathbf{w}_t são independentes entre si, independentes de seus valores passados e independentes de quaisquer valores de \mathbf{x}_t e \mathbf{y}_t .

Nesta representação, \mathbf{y}_t é um vetor $m \times 1$, \mathbf{H}_t é uma matriz $m \times n$, \mathbf{x}_t é um vetor $n \times 1$, \mathbf{v}_t é um vetor $m \times 1$, \mathbf{F}_t é uma matriz $n \times n$, \mathbf{w}_t é um vetor $n \times 1$, $t \in \{1, \dots, T\}$, sendo T o número de observações de \mathbf{y}_t .

A equação (2.1) é chamada de equação de medida, e através dela é modelada a relação entre as variáveis latentes (\mathbf{x}_t) e as variáveis observáveis (\mathbf{y}_t). A equação (2.2) é chamada de equação de transição, e através dela é modelada a dinâmica temporal das variáveis latentes.

2.2 Filtro de Kalman

O Filtro de Kalman (FK) é um algoritmo recursivo que permite estimar sequencialmente o valor de variáveis latentes em sistemas dinâmicos representados sob a forma de Espaço de Estado, conforme definida na seção anterior.

A cada iteração, o FK toma como entrada uma estimativa para o vetor \mathbf{x}_t condicional à informação conhecida até o tempo $t - 1$ e o valor de \mathbf{y}_t e calcula a esperança para \mathbf{x}_{t+1} .

Utilizaremos como notação:

$$\hat{\mathbf{x}}_{t|\tau} = \mathbb{E}[\mathbf{x}_t | \mathbf{y}_1, \dots, \mathbf{y}_\tau]$$

e

$$\hat{\mathbf{P}}_{t|\tau} = \mathbb{E} [(\mathbf{x}_t - \hat{\mathbf{x}}_{t|\tau})(\mathbf{x}_t - \hat{\mathbf{x}}_{t|\tau})']$$

Assumimos ainda que os vetores \mathbf{v}_t e \mathbf{w}_t são normais com a seguinte distribuição conjunta:

$$\begin{bmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \right)$$

Onde \mathbf{R} é uma matriz $m \times m$, \mathbf{Q} é uma matriz $n \times n$.

A recursão do FK começa utilizando-se estimações *a priori* $\hat{\mathbf{x}}_{1|0}$ e $\hat{\mathbf{P}}_{1|0}$ que não são baseadas em nenhuma informação sobre \mathbf{y} .

Com estes valores iniciais, desejamos calcular os valores análogos de $\hat{\mathbf{x}}_{2|1}$ e $\hat{\mathbf{P}}_{2|1}$ utilizando a informação de \mathbf{y}_1 . Este passo é feito genericamente da mesma forma para qualquer valor de t . Portanto, para valores de $\hat{\mathbf{x}}_{t|t-1}$ e $\hat{\mathbf{P}}_{t|t-1}$ calcularemos valores de $\hat{\mathbf{x}}_{t+1|t}$ e $\hat{\mathbf{P}}_{t+1|t}$. Para isso, calcularemos

$$\begin{aligned} \hat{\mathbf{y}}_{t|t-1} &= \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1} \\ \hat{\mathbf{S}}_{t|t-1} &= \mathbb{E} [(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'] \\ &= \mathbb{E} [(\mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1})(\mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1})'] \\ &= \mathbf{H}_t \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t' + \mathbf{R} \\ \hat{\mathbf{v}}_t &= \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} \end{aligned}$$

Em seguida, atualizamos a estimativa $\hat{\mathbf{x}}_{t|t-1}$ para $\hat{\mathbf{x}}_{t|t}$ utilizando a informação de \mathbf{y}_t :

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \hat{\mathbf{v}}_t$$

O valor de \mathbf{K}_t pode ser obtido de pelo menos duas formas diferentes (com o mesmo resultado): A primeira é o valor encontrado como solução do problema de minimização

$$\min_{\mathbf{K}_t} \mathbb{E}((\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})'(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}))$$

A segunda forma é utilizando o resultado da projeção linear

$$\mathbb{E} [\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1} | \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}] = \mathbb{E} [(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'] (E [(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'])^{-1} \hat{\mathbf{v}}_t,$$

Em ambas as formas o resultado obtido para \mathbf{K}_t é

$$\mathbf{K}_t = \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t' \hat{\mathbf{S}}_t^{-1}$$

Substituindo as equações acima, obtemos:

$$\hat{\mathbf{P}}_{t|t} = \mathbb{E} [(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})'] = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \hat{\mathbf{P}}_{t|t-1}$$

Agora, o último passo é calcular as esperanças $\hat{\mathbf{x}}_{t+1|t}$ e $\hat{\mathbf{P}}_{t+1|t}$, a partir dos valores obtidos de $\hat{\mathbf{x}}_{t|t}$ e $\hat{\mathbf{P}}_{t|t}$. Para isso, basta calcularmos a esperança condicional da equação de transição:

$$\begin{aligned} \hat{\mathbf{x}}_{t+1|t} &= \mathbf{F}_{t+1} \hat{\mathbf{x}}_{t|t} \\ \hat{\mathbf{P}}_{t+1|t} &= \mathbf{F}_{t+1} \hat{\mathbf{P}}_{t|t} \mathbf{F}_{t+1}' + \mathbf{Q} \end{aligned}$$

2.3 Quasi-Máxima-Verossimilhança

Em estatística, a máxima-verossimilhança é um método utilizado para estimar os parâmetros "verdadeiros" de um modelo. A função de verossimilhança é definida da seguinte forma:

$$L(\boldsymbol{\theta} | (\mathbf{y}_1, \dots, \mathbf{y}_T)) = \mathbb{P}((\mathbf{y}_1, \dots, \mathbf{y}_T) | \boldsymbol{\theta})$$

Onde $(\mathbf{y}_1, \dots, \mathbf{y}_T)$ são os dados observados, e $\boldsymbol{\theta}$ é o vetor de parâmetros do modelo. Para distribuições de probabilidade contínuas, utiliza-se a função de densidade de probabilidade no lugar da probabilidade propriamente dita.

No caso de um modelo qualquer escrito na forma de espaço de estado, o filtro de Kalman permite calcular a função de verossimilhança, uma vez que conhecemos a distribuição de probabilidade condicional das observações:

$$\mathbf{y}_t | ((\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta}) \sim N(\mathbf{H}_t \mathbf{x}_{t|t-1}, \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + \mathbf{R})$$

Portanto, para calcularmos a verossimilhança dos parâmetros para um dado modelo escrito na forma de espaço de estado, e utilizando-se o filtro de Kalman para se determinar o valor das variáveis de estado em cada instante de tempo, utilizamos a seguinte expressão:

$$L(\boldsymbol{\theta} | (\mathbf{y}_1, \dots, \mathbf{y}_T)) = \prod_{t=1}^T f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta})$$

Onde f é a densidade de probabilidade da distribuição normal multivariada:

$$f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta}) = \frac{\exp\left(-\frac{1}{2} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_{t|t-1})' (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + \mathbf{R})^{-1} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_{t|t-1})\right)}{\sqrt{(2\pi)^m \det(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + \mathbf{R})}}$$

Na prática, utiliza-se o logaritmo da função de verossimilhança, de forma que obtemos a seguinte expressão:

$$\ln L(\boldsymbol{\theta} | (\mathbf{y}_1, \dots, \mathbf{y}_T)) = \sum_{t=1}^T \ln f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta})$$

O estimador dos parâmetros por máxima-verossimilhança então é:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \ln L(\boldsymbol{\theta} | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}))$$

Segundo [Ham94], sob algumas condições de regularidade, a estimação $\hat{\boldsymbol{\theta}}$ por máxima-verossimilhança a partir de uma amostra de tamanho T é assintoticamente normal e consistente.

Essas condições são:

- 1) O modelo deve ser identificável, ou seja, o máximo valor da função de verossimilhança deve ocorrer em apenas um ponto.
- 2) Os auto-valores de \mathbf{F}_t devem estar contidos no círculo unitário.
- 3) O máximo da função de verossimilhança não deve ocorrer na fronteira do conjunto admissível.

Segundo [Ham94], quando $T \rightarrow \infty$:

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, (\mathcal{I}_{2D}\mathcal{I}_{OP}^{-1}\mathcal{I}_{2D})^{-1})$$

Onde $\boldsymbol{\theta}_0$ é o valor "verdadeiro" dos parâmetros, e as matrizes \mathcal{I}_{2D} e \mathcal{I}_{OP} são:

$$\mathcal{I}_{2D} = -\frac{1}{T}\mathbb{E}\left[\sum_{t=1}^T \frac{\partial^2 \ln f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right]$$

$$\mathcal{I}_{OP} = \text{plim} \frac{1}{T} \sum_{t=1}^T \left(\left(\frac{\partial \ln f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \ln f(\mathbf{y}_t | (\mathbf{y}_1, \dots, \mathbf{y}_{t-1}), \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)' \right)$$

Ambas calculadas no ponto $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

Desta forma, podemos também calcular a matriz de covariâncias dos parâmetros, e em particular o erro-padrão associado a cada parâmetro, com o objetivo de determinar se estes são estatisticamente significantes.

Deve-se salientar ainda que a estimação dos parâmetros por maximização da função de verossimilhança calculada com o filtro de Kalman é ótima caso os resíduos de estimação do modelo possuam distribuição normal. Para isso é executado um teste de hipótese nos resíduos do modelo para avaliar a hipótese de normalidade com relação à sua distribuição. Nestes casos a estimação por máxima-verossimilhança é ótima e consistente. Em casos de não normalidade, este método pode ainda assim ser utilizado, mas cabe destacar que, nestas situações, ele não será ótimo, mas continua consistente e assintoticamente normal.

Capítulo 3

Desenvolvimento da solução

3.1 Implementação

O primeiro passo para se implementar a estimação do modelo Gibson-Schwartz utilizando máxima-verossimilhança calculada com o filtro de Kalman, é escrever o modelo na forma de Espaço de Estado discreta.

Discretizando as equações do modelo, obtemos:

$$X_{t+\Delta t} = X_t + \left(\mu - \delta_t - \frac{\sigma_1^2}{2} \right) \Delta t + \varepsilon_{1,t}$$

$$\delta_{t+\Delta t} = \delta_t + \kappa(\alpha - \delta_t)\Delta t + \varepsilon_{2,t}$$

Onde $[\varepsilon_{1,t}, \varepsilon_{2,t}]$ são serialmente descorrelacionados e com distribuição conjunta:

$$\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Delta t \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

Agora, precisamos escrever estas equações em forma matricial. Nota-se que na definição apresentada na seção 2.1 tanto a equação de transição quanto a de medida não apresentam um termo afim, necessário para adaptar este modelo para a forma de Espaço de Estado. Para resolver isso definimos nosso vetor de estados como:

$$\mathbf{x}_t = \begin{bmatrix} X_t \\ \delta_t \\ 1 \end{bmatrix}$$

Desta forma, aumentamos o nosso vetor de estados com um estado *dummy* de valor 1 que não se altera no tempo. Agora podemos colocar os termos afins na última coluna da matriz \mathbf{F}_t . Então definimos a matriz de transição \mathbf{F}_t como:

$$\mathbf{F}_t = \begin{bmatrix} 1 & -\Delta t & \left(\mu - \frac{\sigma_1^2}{2} \right) \Delta t \\ 0 & 1 - \kappa\Delta t & \kappa\alpha\Delta t \\ 0 & 0 & 1 \end{bmatrix}$$

e a matriz \mathbf{Q} como:

$$\mathbf{Q} = \begin{bmatrix} \sigma_1^2\Delta t & \rho\sigma_1\sigma_2\Delta t & 0 \\ \rho\sigma_1\sigma_2\Delta t & \sigma_2^2\Delta t & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

O próximo passo é escrever a equação de medida. Como o preço *spot* não é observável, utiliza-se, então, o preço dos m futuros mais líquidos negociados como variáveis observáveis e a equação (1.3) como equação de medida, utilizada para apreçar contratos futuros. Para isso tomamos o logaritmo de ambos os lados da equação obtendo:

$$\ln F_t(T_i) = X_t - \delta_t \left(\frac{1 - e^{-\kappa T_i}}{\kappa} \right) + A(T_i), \quad 1 \leq i \leq m, \quad m \geq 2$$

Onde T_i indica o tempo ao vencimento do i -ésimo contrato no instante t . Esta abordagem exige que utilizemos no mínimo dois contratos futuros a cada instante de tempo para que o modelo seja identificável. Se utilizarmos apenas um contrato, o prêmio de risco de mercado do *convenience yield*, λ , deixa de ser identificável.

Escrevendo esta equação de forma matricial, então, obtemos:

$$\underbrace{\begin{bmatrix} \ln F_t(T_1) \\ \vdots \\ \ln F_t(T_m) \end{bmatrix}}_{\mathbf{y}_t} = \underbrace{\begin{bmatrix} 1 & -\frac{1 - e^{-\kappa T_1}}{\kappa} & A(T_1) \\ \vdots & \vdots & \vdots \\ 1 & -\frac{1 - e^{-\kappa T_m}}{\kappa} & A(T_m) \end{bmatrix}}_{\mathbf{H}_t} \mathbf{x}_t + \mathbf{v}_t$$

Sendo \mathbf{y}_t e \mathbf{H}_t definidos conforme indicado acima. Para simplificar a estimação, a matriz \mathbf{R} é definida como:

$$\mathbf{R} = \begin{bmatrix} \xi_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \xi_m^2 \end{bmatrix}$$

Essa simplificação se faz necessária pois quando se usa uma quantidade não tão pequena de contratos, a quantidade de parâmetros a serem otimizados cresce quadraticamente com o número de contratos, caso queiramos estimar as covariâncias de medida de cada par de vencimentos.

O desenvolvimento da solução, portanto, foi feito escrevendo-se rotinas na linguagem Matlab, que desempenham as seguintes funções:

1) Filtro de Kalman: a partir de matrizes \mathbf{F} , \mathbf{H}_t , \mathbf{Q} , \mathbf{R} , e valores de \mathbf{y}_t dados, esta rotina deve executar a estimação sequencial descrita na seção (2.2) e calcular o valor da verossimilhança para a configuração fornecida.

2) Modelo Gibson-Schwartz: a partir de valores de μ , κ , α , σ_1 , σ_2 , ρ , λ , S_0 , δ_0 , $[\xi_1 \dots \xi_m]$, e dos valores fornecidos para os preços dos futuros, os respectivos prazos e as taxas de juros de mercado para os respectivos prazos, esta rotina deve calcular as matrizes \mathbf{F} , \mathbf{H}_t , \mathbf{Q} , \mathbf{R} e os valores de \mathbf{y}_t e chamar a rotina acima para calcular a verossimilhança do conjunto de parâmetros fornecido.

3) Otimização dos parâmetros: a partir dos valores dos preços dos futuros, dos prazos e das taxas de juros, calcular uma aproximação inicial para os parâmetros e utilizar a função 'fmincon', fornecida pelo pacote de otimização do Matlab, para encontrar o conjunto ótimo de parâmetros para a função 'gibson-schwartz' (descrita acima). A função 'fmincon' também retorna a matriz Hessiana calculada neste ponto, e com ela calculamos o erro padrão associado à configuração ótima obtida.

O códigos desenvolvidos são apresentados no apêndice A.

3.2 Dados utilizados

Nesta seção apresentamos em detalhes os dados históricos utilizados neste trabalho.

Os dados utilizados neste trabalho são séries de preços de ajuste diário de contratos futuros de boi-gordo (BGI) e séries de futuros de juros (DI futuro) ambos negociados na BM&FBOVESPA.

A maturidade dos contratos de Boi-Gordo utilizados variam de um a cinco meses. A partir dos dados disponibilizados pela bolsa, foram montadas sinteticamente cinco séries de preços roladas. As séries foram montadas da seguinte forma: o preço da primeira série no dia t é o preço de ajuste do dia t do contrato que, naquele dia, era o mais próximo a vencer. As demais séries foram montadas analogamente utilizando-se do segundo ao quinto contratos mais próximos a vencerem em cada data. Desta forma, obtemos séries sintéticas do primeiro ao quinto futuro.

O gráfico abaixo mostra o valor dos preços de cada uma das séries sintéticas:



Figura 3.1: Preço diário de ajuste dos contratos BGI (boi-gordo)

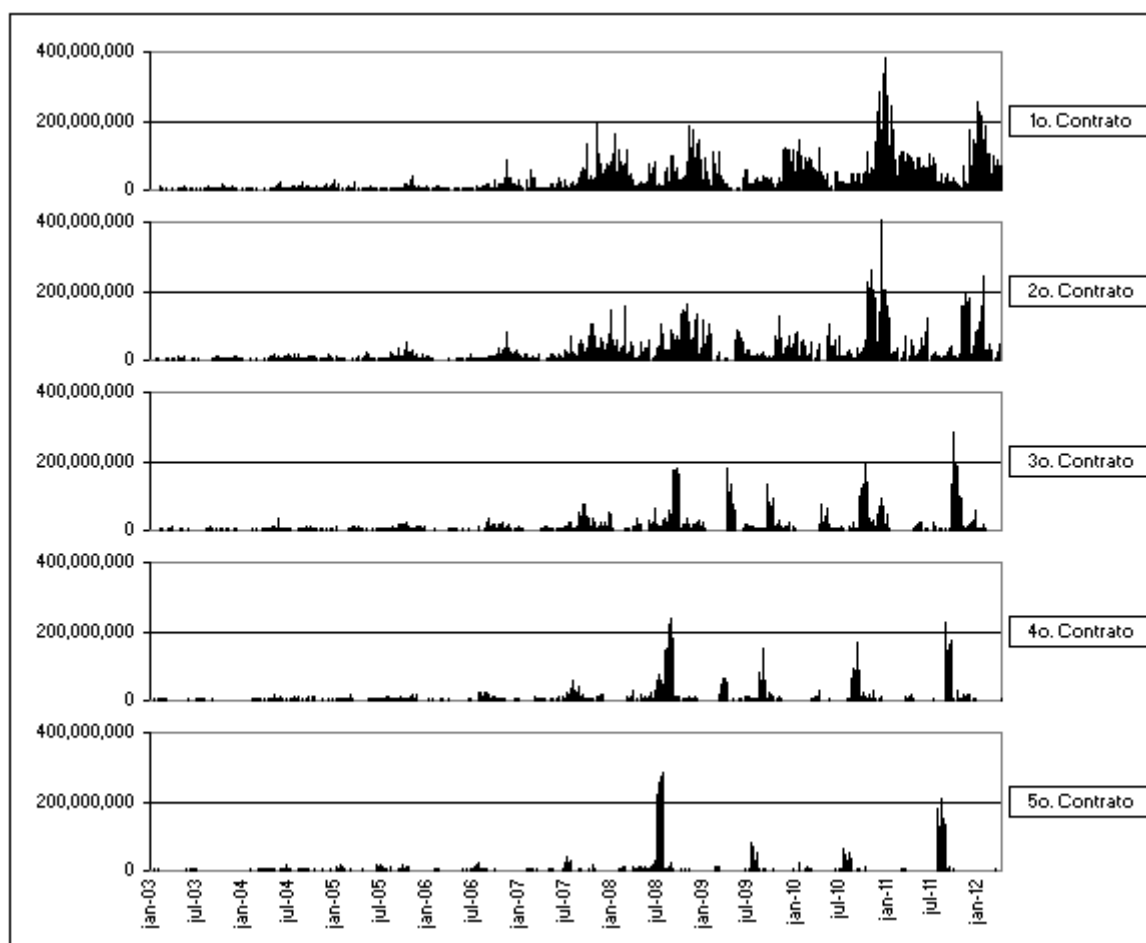


Figura 3.2: Volume diário (em R\$) dos contratos BGI (boi-gordo)

Para podermos utilizar estas séries, também geramos cinco séries de prazos de vencimento das séries sintéticas de preço. Essas séries foram geradas exatamente como as séries sintéticas de preços com a diferença de que foi utilizado o prazo em anos de cada contrato, no lugar de seu preço.

O gráfico abaixo mostra os tempos ao vencimento em dias úteis das séries sintéticas:

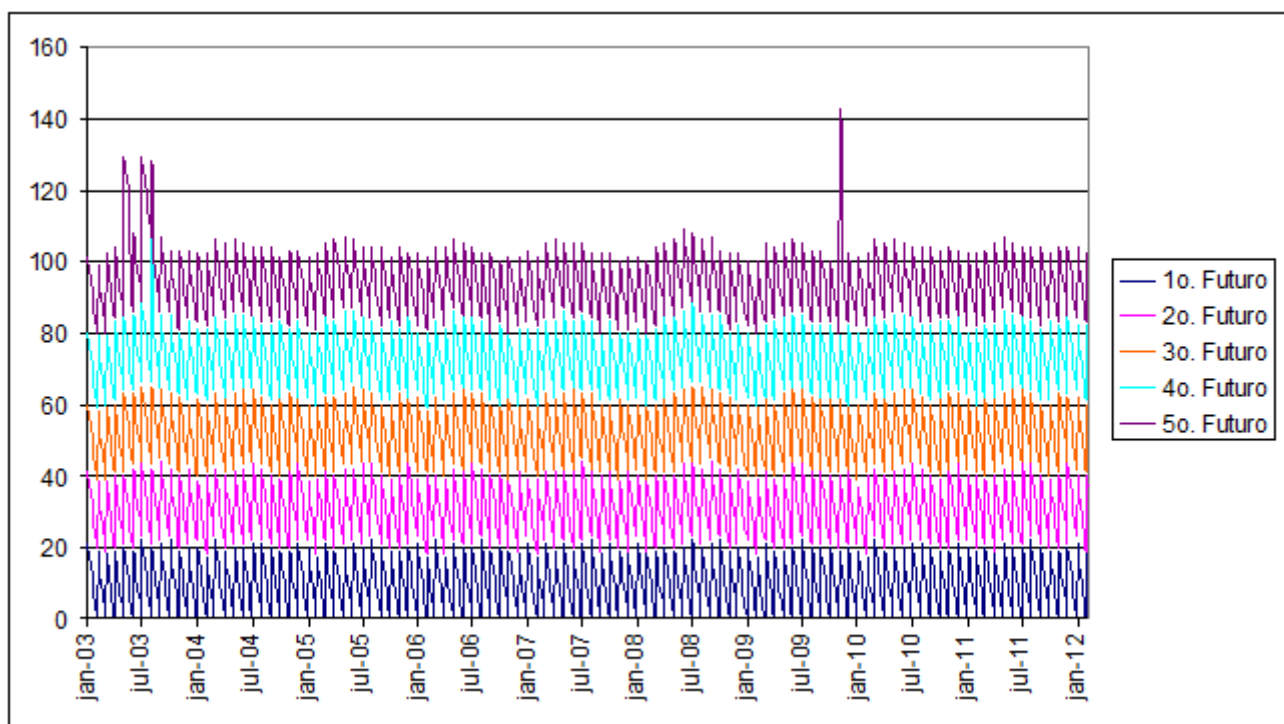


Figura 3.3: Prazos ao vencimento dos contratos utilizados em cada dia

Nota-se que em alguns dias com menor liquidez, o contrato que até então era o quinto mais próximo a vencer não foi negociado e não teve seu preço de ajuste determinado pela bolsa. Por isso, nestes dias o valor do quinto contrato foi substituído pelo valor do contrato seguinte, o qual foi negociado e teve preço de ajuste determinado pela bolsa.

Neste trabalho utilizaremos o modelo Gibson-Schwartz pois ele, ao contrário do modelo Schwartz-Smith, nos permite utilizar explicitamente o valor da taxa de juros que desejarmos. Isto é particularmente útil neste caso, já que podemos utilizar o valor de mercado do juro para cada vencimento em cada dia, pois podemos extrair a curva de juros diária a partir dos contratos futuros de DI, negociados na mesma bolsa. Com isso podemos considerar a curva de juros conhecida, e garantir que qualquer oscilação no valor do *convenience yield* não ocorre por interferência espúria de oscilações na taxa de juros.

Em seguida, utilizamos os valores da curva de juros de mercado, obtida dos preços de ajuste dos futuros de DI, e extraímos o valor do juro continuamente capitalizado para cada vencimento dos contratos de Boi-Gordo em cada dia, através de uma interpolação linear da curva de juros diária. Como utilizamos o valor das taxas de juros continuamente capitalizadas, e as datas de vencimento dos futuros de DI diferem dos vencimentos dos futuros de Boi-Gordo em apenas um dia útil, não julgamos necessária nenhuma técnica mais complexa para interpolar os valores necessários.

Obtivemos assim, cinco séries sintéticas de juros, com vencimentos iguais aos vencimentos das séries sintéticas da *commodity*.

O gráfico abaixo mostra o valor das séries sintéticas de juros que foram utilizadas.

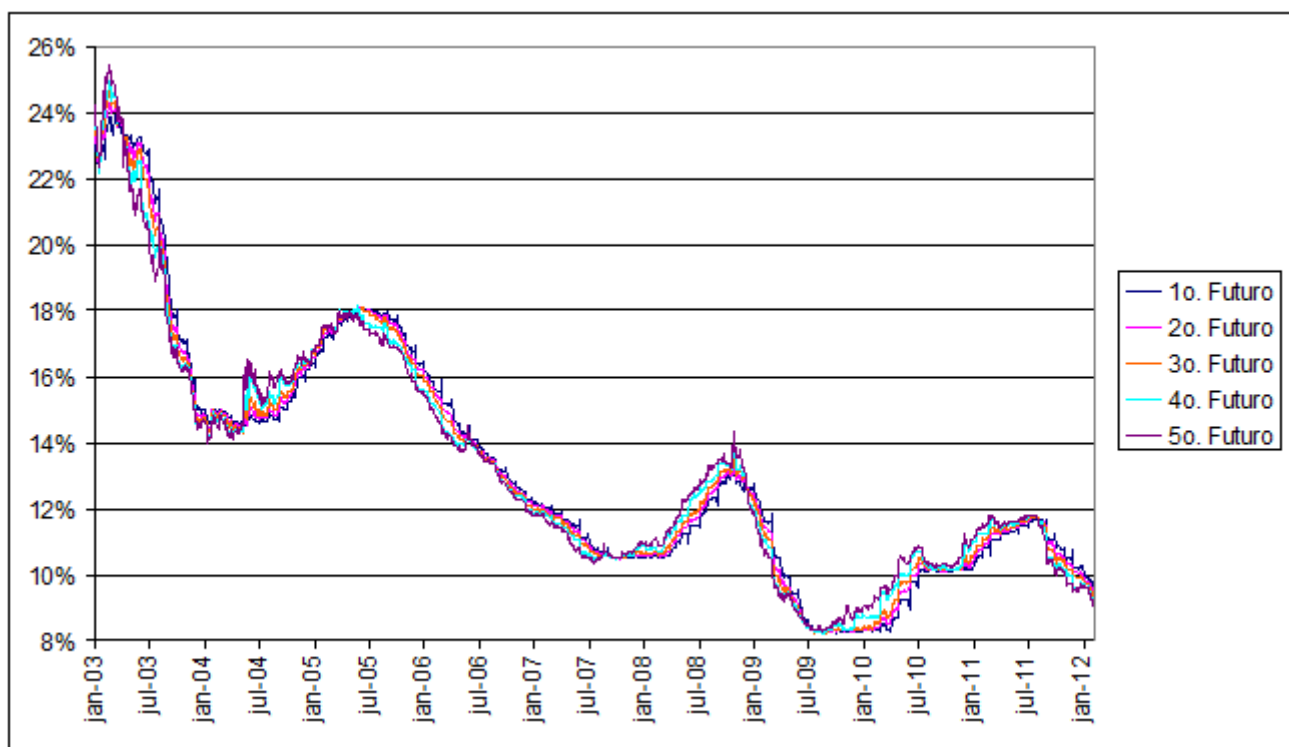


Figura 3.4: Taxa de juros livre de risco continuamente capitalizada ao ano

Nota-se que no período analisado, o valor da taxa de juros oscilou consideravelmente. Caso utilizássemos um valor médio constante para a taxa de juros, assim como em [Sch97], estas oscilações iriam "contaminar" a estimação do *convenience-yield*. Como o juro neste período apresentou clara tendência de queda, a estimação dos parâmetros do modelo ficaria prejudicada, já que esta tendência iria diminuir a significância estatística da taxa de reversão à média κ do modelo Gibson-Schwartz.

3.3 Teste com dados aleatórios

Com o objetivo de testar as rotinas de estimação, descritas na seção (3.1), foi feita uma simulação com dados aleatórios, antes de ajustar o modelo aos dados reais. Deve-se ressaltar que o procedimento descrito a seguir não é uma simulação de Monte-Carlo, mas sim uma ilustração de como deveria funcionar na prática a rotina de estimação desenvolvida.

Geramos aleatoriamente uma realização para os processos latentes S_t e δ_t , conforme as equações do modelo Gibson-Schwartz, utilizando intervalo de tempo de um dia útil a cada observação. A partir destes valores foram calculados os preços teóricos para três contratos futuros fictícios com maturidades a cada 60 dias, sendo o mais próximo a vencer em no máximo 60 dias, e os demais com vencimento 60 dias depois do anterior. Neste teste foi utilizada uma taxa de juros contínua constante de 8% ao ano. O preço teórico de cada contrato em cada instante de tempo foi então "contaminado" com um ruído gaussiano de média zero e desvio padrão de 0.05% do preço.

Em seguida, foi utilizado o procedimento descrito anteriormente para estimar, a partir dos preços dos contratos fictícios, os parâmetros utilizados na simulação e os valores de S_t e δ_t .

A tabela a seguir relaciona os parâmetros utilizados na geração dos dados e os parâmetros encontrados na estimação deles por máxima-verossimilhança calculada com o filtro de Kalman.

Parâmetro	Valor Real	Estimado	Erro Padrão
μ	0.1300	0.1300	0.0043
κ	2.5200	2.6495	0.2885
α	0.0500	0.0499	0.0035
σ_1	0.1000	0.1004	0.0000
σ_2	0.0200	0.0219	0.0004
ρ	0.0230	0.0456	0.1224
λ	0.0100	0.0099	0.0236
S_0	20.0000	19.9608	0.5370
δ_0	0.0500	0.0517	0.0024
ξ_1	0.0005	0.0005	0.0000
ξ_2	0.0005	0.0005	0.0000
ξ_3	0.0005	0.0005	0.0000

Tabela 3.1: Parâmetros reais e estimados utilizados no exercício com dados simulados

O gráfico abaixo mostra o processo X_t teórico (logaritmo do processo S_t sorteado aleatoriamente) em azul e o processo $\hat{X}_{t|t}$ (estimado com o filtro de Kalman utilizando os parâmetros estimados por máxima-verossimilhança) em verde. As curvas estão praticamente sobrepostas dando a impressão de serem a mesma.

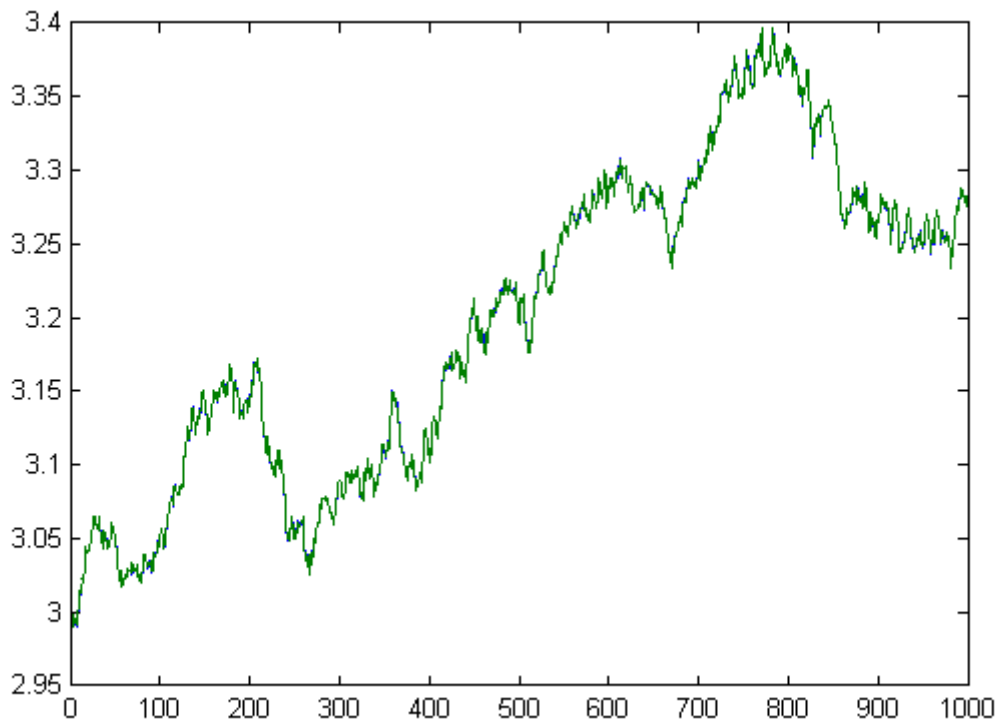


Figura 3.5: Processo X_t teórico e estimado no exercício com dados simulados

O gráfico abaixo mostra o processo δ_t teórico em azul e o processo $\hat{\delta}_{t|t}$ (estimado com o filtro de Kalman utilizando os parâmetros estimados por máxima-verossimilhança) em verde:

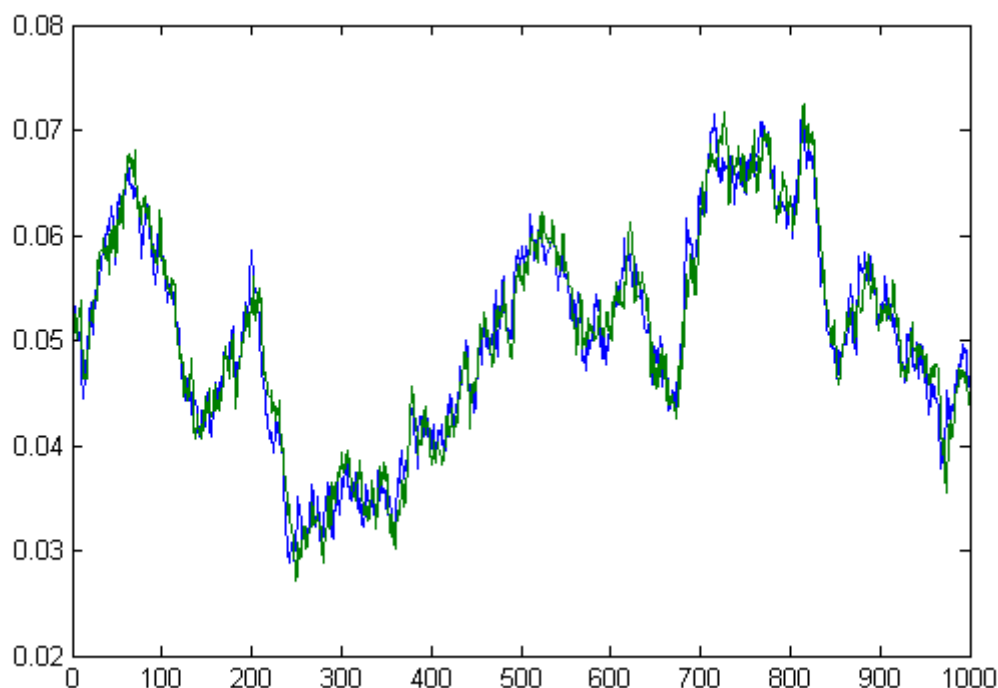


Figura 3.6: Processo δ_t teórico e estimado no exercício com dados simulados

Neste gráfico, ambas as séries δ_t (processo teórico) e $\hat{\delta}_{t|t}$ (valores estimados) estão sobrepostas, mas menos do que os processos X_t e $\hat{X}_{t|t}$. Isso ocorre pois a influência do valor de δ_t no preço dos futuros é proporcional ao tempo até a maturidade. Como neste teste foram utilizadas maturidades relativamente curtas, a estimação deste processo é menos precisa do que a do processo X_t . Além disso a estimação do processo X_t é naturalmente mais precisa pois o seu valor impacta diretamente o nível dos preços dos contratos, enquanto que o processo δ_t afeta as diferenças entre os preços para diferentes maturidades.

Quanto mais longas as maturidades dos contratos maior a acurácia da estimação do processo δ_t . Enquanto contratos futuros de petróleo ou gás natural são negociados com maturidades de até alguns anos, isso não acontece para os contratos de boi-gordo na BM&FBOVESPA, que são negociados com maturidades de até poucos meses. Por isso, foram escolhidos valores pequenos para o tempo de maturidade dos contratos fictícios utilizados neste teste.

Capítulo 4

Resultados empíricos

Nesta seção discutiremos os testes realizados e os resultados obtidos com os dados históricos descritos na seção (3.2).

Inicialmente ajustamos o modelo Gibson-Schwartz para o conjunto completo dos dados disponíveis, ou seja, utilizamos as séries dos cinco primeiros futuros disponíveis, assim como as séries das taxas de juros para os respectivos vencimentos, conforme descrito na seção (3.2).

A tabela a seguir relaciona os parâmetros encontrados na estimação por máxima-verossimilhança calculada com o filtro de Kalman.

Parâmetro	Valor Estimado	Erro Padrão
μ	0.1705	0.0181
κ	3.7670	0.0615
α	0.0855	0.0176
σ_1	0.1573	0.0028
σ_2	1.0561	0.0087
ρ	0.6440	0.0087
λ	0.2443	0.0179
S_0	58.8693	0.5335
δ_0	0.8198	0.0451
ξ_1	0.0020	0.0003
ξ_2	0.0149	0.0003
ξ_3	0.0131	0.0002
ξ_4	0.0000	0.0004
ξ_5	0.0175	0.0003
log(verossimilhança)	33749	-

Tabela 4.1: Parâmetros encontrados para o conjunto completo de dados

Observando-se os valores da tabela acima, podemos perceber que todos os parâmetros têm comportamento esperado e principalmente todos são estatisticamente significantes a qualquer nível razoável de confiança (por exemplo, 95%). A única exceção é o parâmetro ξ_4 , o que significa que a amplitude do ruído na série do quarto futuro não é estatisticamente significativa, indicando que o modelo se ajustou particularmente bem a essa série.

O gráfico a seguir mostra a evolução no tempo do preço *spot* obtido pelo filtro de Kalman, e o preço do primeiro futuro rolado.

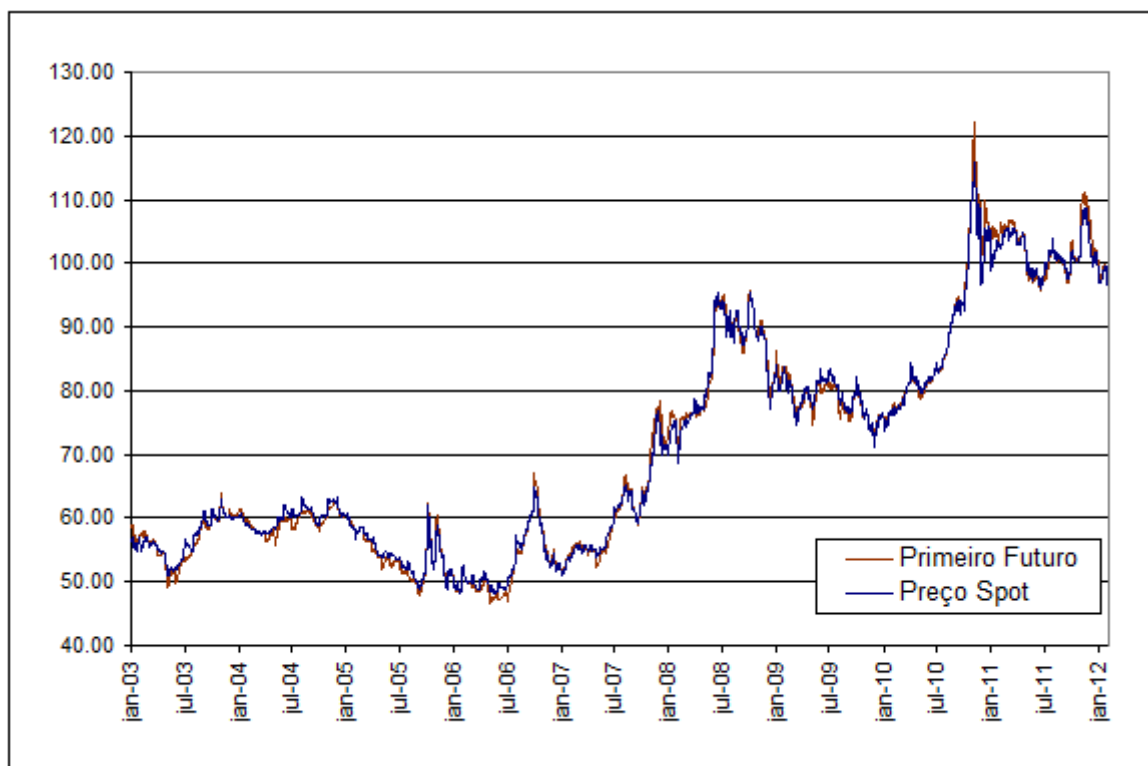


Figura 4.1: Preço *spot* estimado e preço do primeiro futuro BGI (boi-gordo)

Como esperado, as duas séries são muito parecidas, indicando bom ajuste do modelo aos dados.

O próximo gráfico apresenta a evolução no tempo do *convenience-yield* instantâneo estimado.

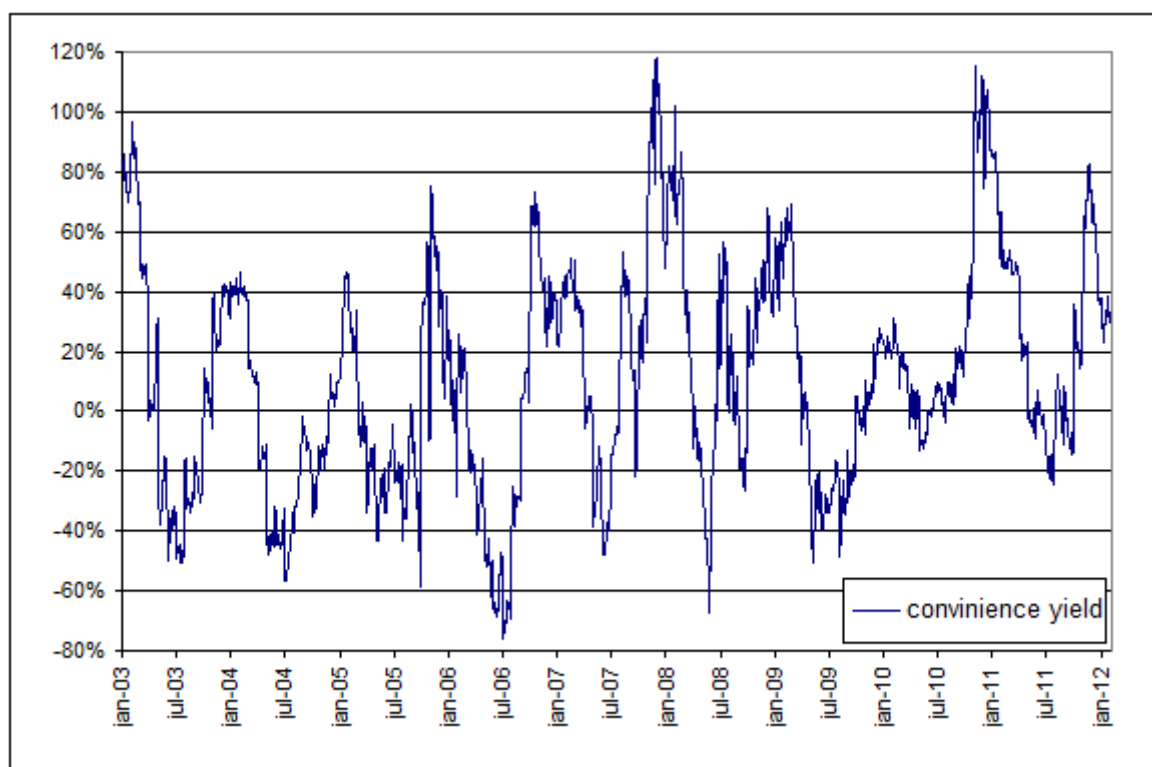


Figura 4.2: *Convenience-yield* estimado para os futuros BGI (boi-gordo)

Inspecionando-se visualmente o comportamento desta série, podemos confirmar seu caráter de reversão à média, já esperado pelo fato de o parâmetro κ ter se mostrado estatisticamente significativo.

Através desta estimação, podemos considerar que o modelo Gibson-Schwartz ajustado utilizando as taxas de juros de mercado, tal como descrito nas seções anteriores, é adequado para descrever os dados obtidos de boi-gordo.

A seguir, ajustamos o modelo novamente aos dados, mas desta vez retiramos os últimos 252 dias úteis (um ano) da amostra. Em seguida rodamos uma vez o filtro de Kalman para os dias retirados da amostra, e utilizando-se nesta execução do filtro apenas os três contratos mais curtos (com menores tempos ao vencimento). Feito isso, obtivemos uma estimativa para as séries S_t e δ_t para este último período de 252 dias úteis, e assim calculamos o valor dos contratos mais longos, deixados de fora da amostra para compararmos com os valores divulgados pela bolsa.

A tabela abaixo relaciona os parâmetros encontrados no ajuste do modelo ao conjunto de dados excluindo-se os últimos 252 dias úteis.

Parâmetro	Valor Estimado	Erro Padrão
μ	0.1470	0.0146
κ	1.9032	0.0223
α	0.0442	0.0134
σ_1	0.2067	0.0031
σ_2	0.8391	0.0095
ρ	0.7406	0.0056
λ	0.0800	0.0137
S_0	57.8805	0.7305
δ_0	0.5716	0.0480
ξ_1	0.0246	0.0004
ξ_2	0.0000	0.0009
ξ_3	0.0086	0.0001
ξ_4	0.0017	0.0002
ξ_5	0.0175	0.0003

Tabela 4.2: Parâmetros encontrados para os dados *in sample*.

O gráfico a seguir mostra a evolução no tempo do preço do quarto futuro e seu preço teórico calculado conforme descrito acima.

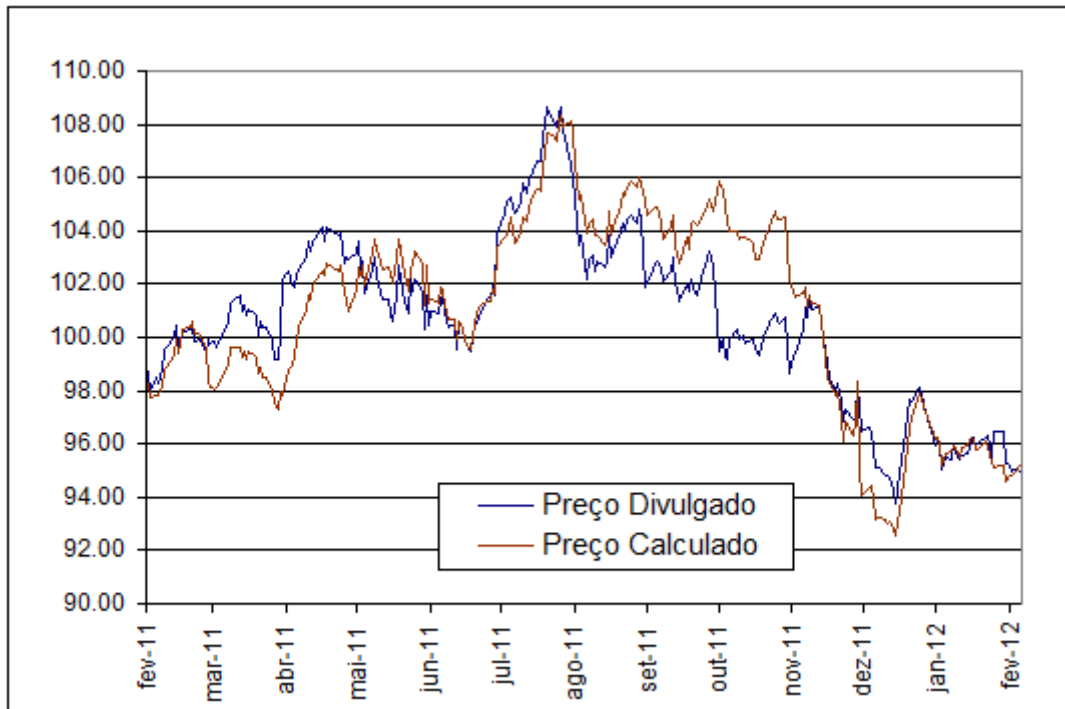


Figura 4.3: Preço calculado e preço divulgado do quarto futuro BGI (boi-gordo)

O gráfico a seguir é semelhante ao anterior mas mostra o preço do quinto futuro e seu preço teórico calculado.

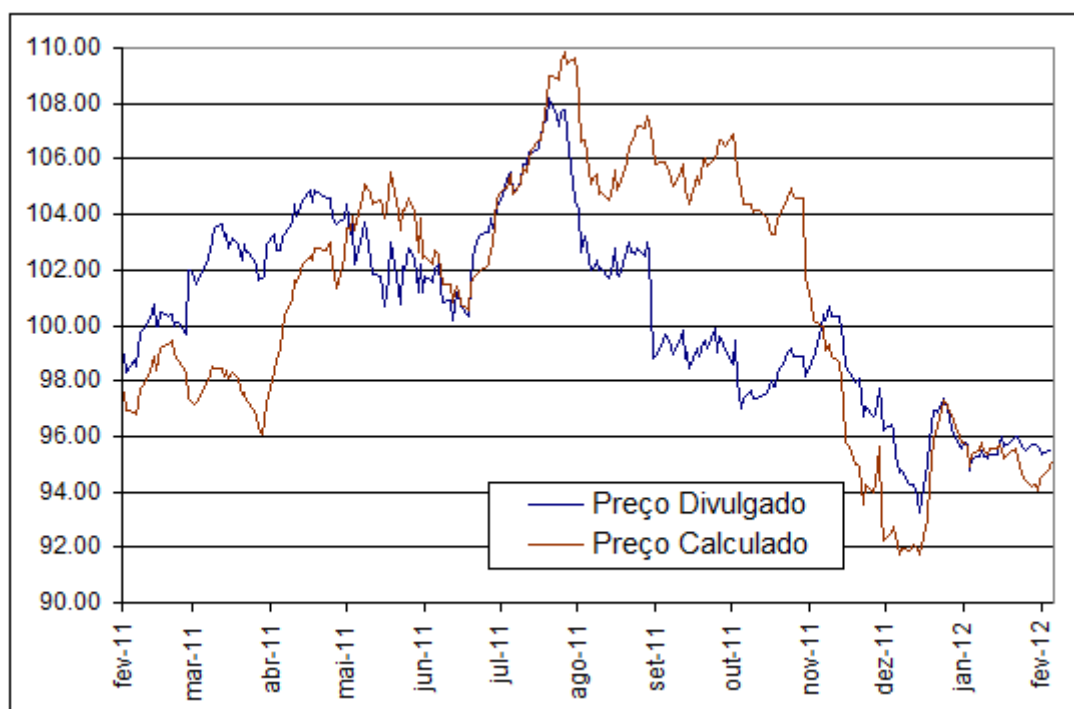


Figura 4.4: Preço calculado e preço divulgado do quinto futuro BGI (boi-gordo)

Observando-se os gráficos acima, pode-se perceber que há uma diferença significativa entre os valores calculados utilizando-se a metodologia desenvolvida e os valores divulgados pela bolsa, principalmente no gráfico do quinto contrato. Isto se deve ao fato de que, na maioria dos dias utilizados nesta comparação, não ocorreram negócios (embora o mercado estivesse aberto) por falta de liquidez. Nestes casos a bolsa é responsável por atribuir um preço aos contratos para poder fazer chamadas de margem e ajustes diários de contratos em aberto. A metodologia utilizada pela bolsa não é divulgada abertamente. No entanto, sabemos que esta metodologia se baseia em ajustes de sazonalidade e não é livre de arbitragem. Por isso consideramos que o procedimento apresentado neste trabalho é uma possível alternativa ao utilizado pela bolsa.

Propomos como sugestão para pesquisa futura uma estratégia de arbitragem para mostrar que a metodologia utilizada pela bolsa não é livre de arbitragem. Tal estratégia deve montar posições teóricas em dias em que não houve negociação em algum vencimento ao preço arbitrado pela bolsa. As posições devem ser zeradas no mercado quando houver negociações ou deve-se esperar o vencimento dos contratos. Consideramos este teste conclusivo caso a estratégia seja capaz de ganhar dinheiro sistematicamente.

Capítulo 5

Conclusão

Devido à grande importância do estudo da dinâmica de preços de *commodities*, neste trabalho tivemos como objetivo ajustar o modelo de apreçamento Gibson-Schwartz ao mercado brasileiro. Para isso foi escolhido o mercado de boi-gordo da BM&FBOVESPA, onde os respectivos contratos futuros são negociados diariamente, e cujos dados até então não haviam sido utilizados em um trabalho desta natureza.

Para isso, foram estudadas ferramentas matemáticas necessárias para resolvermos adequadamente o problema, em particular: apreçamento neutro ao risco e estimação por máxima verossimilhança com o Filtro de Kalman.

Utilizamos, então, uma abordagem semelhante àquela explorada em [Sch97] para ajustar o modelo aos dados de boi-gordo, e desta forma, estimar os valores dos parâmetros e os respectivos erros-padrão, para avaliarmos se o modelo é adequado para descrever os dados.

O ajuste do modelo foi feito com sucesso, pois o modelo se mostrou aderente aos dados e, em particular, a taxa de reversão à média do *convenience-yield* mostrou alto grau de significância estatística.

Em seguida realizamos uma comparação entre os preços dos contratos mais longos (e ilíquidos) divulgados pela bolsa, e os valores teóricos calculados utilizando o modelo.

Outra contribuição deste trabalho se deve ao uso da curva de juros de mercado em vez de um valor constante na estimação dos demais parâmetros do modelo. Consideramos que esta mudança representa uma melhora na abordagem, uma vez que as hipóteses utilizadas para justificar um valor constante para o juro em [Sch97] não são realistas para a realidade brasileira durante o período analisado.

Entende-se como um possível desenvolvimento futuro deste trabalho, o uso dos resultados obtidos para a realização de uma estratégia de arbitragem estatística. Esta estratégia consistiria em montar uma carteira que ficasse exposta apenas à variação do *convenience-yield*. Isto é possível, pois podemos "hedgear" a exposição da carteira ao preço *spot* comprando um vencimento e vendendo outro, e a exposição ao juro através dos futuros de DI, cujos vencimentos ocorrem apenas um dia após os vencimentos dos futuros de boi-gordo. Desta forma, utilizando a abordagem apresentada neste trabalho e determinando-se valores de para compra e venda do *convenience-yield*, podemos desenvolver uma estratégia de negociação que lucre com a reversão à média do *convenience-yield*.

Apêndice A

Códigos implementados

A.1 Filtro de Kalman

```
1 function [LL, results] = kalman_filter(Y, F, H, Q, R, X_init,
   P_init)
2
3     %  $Y_t = H_t * X_t + N(0, R)$ 
4     %  $X_t = F_t * X_{t-1} + N(0, Q)$ 
5
6     if (nargin < 8)
7         dosmooth = false;
8     end
9
10    T = size(Y, 3);
11    sizeF = size(F, 3);
12    if (sizeF < T)
13        F(:, :, sizeF + 1 : T) = repmat(F(:, :, end), [1 1 T -
   sizeF]);
14    end
15
16    sizeH = size(H, 3);
17    if (sizeH < T)
18        H(:, :, sizeH + 1 : T) = repmat(H(:, :, end), [1 1 T -
   sizeH]);
19    end
20
21    sizeX = length(X_init);
22    sizeY = length(Y(:, :, 1));
23
24    Xpred = zeros(sizeX, 1, T);
25    Xupdt = zeros(sizeX, 1, T);
26    Xsmth = zeros(sizeX, 1, T);
27
28    Ppred = zeros(sizeX, sizeX, T);
29    Pupdt = zeros(sizeX, sizeX, T);
```

```

30     Psmth = zeros(sizeX, sizeX, T);
31
32     Xpred(:, :, 1) = X_init;
33     Ppred(:, :, 1) = P_init;
34
35     K = zeros(sizeX, sizeY, T);
36     r = zeros(T, sizeY);
37     % Forward pass:
38     %LL = log(mvnpdf(Y(:, :, 1), H(:, :, 1) * Xpred(:, :, 1), H
39         (:, :, 1) * Ppred(:, :, 1) * H(:, :, 1)' + R));
40     klog2pi = sizeY*log(2*pi);
41     logLikelihood = @(x, u, v)(-.5 * (klog2pi + log(det(v)) + (
42         x-u)' * (v \ (x-u))));
43     u = H(:, :, 1) * Xpred(:, :, 1);
44     v = H(:, :, 1) * Ppred(:, :, 1) * H(:, :, 1)' + R;
45     r(1, :) = (Y(:, :, 1) - u);
46     if (v > 0)
47         LL = logLikelihood(Y(:, :, 1), u, v);
48     else
49         LL = 0;
50     end
51     for t = 1:(T - 1)
52         K(:, :, t) = Ppred(:, :, t) * H(:, :, t)' / v;
53
54         aux = (eye(sizeX) - K(:, :, t) * H(:, :, t));
55         Pupdt(:, :, t) = aux * Ppred(:, :, t) * aux' + K(:, :,
56             t) * R * K(:, :, t)';
57         Xupdt(:, :, t) = Xpred(:, :, t) + K(:, :, t) * (Y(:, :,
58             t) - u);
59
60         Xpred(:, :, t + 1) = F(:, :, t) * Xupdt(:, :, t);
61         Ppred(:, :, t + 1) = F(:, :, t) * Pupdt(:, :, t) * F(:,
62             :, t)' + Q;
63
64         u = H(:, :, t + 1) * Xpred(:, :, t + 1);
65         v = H(:, :, t + 1) * Ppred(:, :, t + 1) * H(:, :, t +
66             1)' + R;
67         r(t+1, :) = (Y(:, :, t + 1) - u);
68
69         LL = LL + logLikelihood(Y(:, :, t + 1), u, v);
70     end
71     %return;
72     t = T;
73     K(:, :, t) = Ppred(:, :, t) * H(:, :, t)' / v;
74     aux = (eye(sizeX) - K(:, :, t) * H(:, :, t));
75     Pupdt(:, :, t) = aux * Ppred(:, :, t) * aux' + K(:, :, t) *
76         R * K(:, :, t)';

```

```

70     Xupdt(:, :, t) = Xpred(:, :, t) + K(:, :, t) * (Y(:, :, t) -
        u);
71
72 end}

```

A.2 Função gibson-schwartz

```

1 function [loglik results] = gibson_schwartz(prazos, ...
2         precos, ...
3         juros, ...
4         dt, ...
5         mu, ...
6         kappa, ...
7         alpha, ...
8         sigma_S, ...
9         sigma_delta, ...
10        rho, ...
11        lambda, ...
12        S_init, ...
13        delta_init, xi)
14     kappa = abs(kappa);
15     A_ = A(prazos, juros, kappa, alpha, sigma_S, sigma_delta,
16           rho, lambda);
17     Y = log(precos);
18     Y = permute(Y, [2 3 1]);
19     H = -(1 - exp(-kappa * prazos))/kappa;
20     H = permute(H, [2 3 1]);
21     H = [ones(size(H)), H, permute(A_, [2 3 1])];
22
23     X_init = [log(S_init)
24              delta_init
25              1];
26
27     F = [1,          -dt, (mu - 0.5*sigma_S^2)*dt
28          0, 1-kappa*dt,          kappa*alpha*dt
29          0,          0,          1];
30
31     xi = [xi(:)' repmat(xi(end), 1, size(precos, 2) - length(xi)
32           )];
33     R = diag(xi .^ 2);
34
35     Q = dt * [          sigma_S^2 rho*sigma_S*sigma_delta 0
36               rho*sigma_S*sigma_delta          sigma_delta^2 0
37               0          0 0 0];
38     P_init = [0 0 0; 0 0 0; 0 0 0];

```

```

38
39     if nargout == 1
40         loglik = kalman_filter(Y, F, H, Q, R, X_init, P_init,
                               false);
41     else
42         [loglik results] = kalman_filter(Y, F, H, Q, R, X_init,
                                          P_init, true);
43     end
44
45 end
46
47 function a = A(T, r, kappa, alpha, sigma_S, sigma_delta, rho,
               lambda)
48     alpha_hat = alpha - lambda / kappa;
49     a = (r - alpha_hat + 0.5 * sigma_delta^2 / kappa^2 -
          sigma_S*sigma_delta*rho/kappa) .* T + ...
50         0.25 * (sigma_delta^2) * (1 - exp(-2*kappa .* T)) / (
          kappa^3) + ...
51         (alpha_hat * kappa + sigma_S * sigma_delta * rho - (
          sigma_delta^2)/kappa)* (1 - exp(-kappa * T)) / (
          kappa ^2);
52 end

```

A.3 Função otimiza-gibson-schwartz

```

1 function [logLik, mu, kappa, alpha, sigma_S, sigma_delta, rho,
           lambda, ...
2           initial_S, initial_delta, measurementStdev, logS_t,
           delta_t] ...
3           = otimiza_gibson_schwartz(prazos, precos, juros, dt, mu,
           kappa, alpha, sigma_S, sigma_delta, rho, lambda, ...
4           initial_S, initial_delta, measurementStdev)
5
6 prazos = prazos * dt;
7
8 theta = [mu, kappa, alpha, sigma_S, sigma_delta, rho, lambda,
           ...
9           initial_S, initial_delta, measurementStdev];
10
11 f = @(x)-gibson_schwartz(prazos, precos, juros, dt, x(1), x(2),
12                          x(3), x(4), ...
13                          x(5), x(6), x(7), x(8), x(9), x(10:end));
14
15 opt = optimset('fmincon');
16 opt = optimset(opt, 'Algorithm', 'interior-point', ...
17                'Display', 'iter', ...

```

```

17         'MaxFunEvals', 5000, ...
18         'MaxIter', 5000, ...
19         'TolX', 1e-10, ...
20         'TolFun', 1e-10, ...
21         'UseParallel', 'always');
22
23 LB = [-2, 0.001, -2, .0001, .0001, -.98, -3, 0, -3, 1e-12 +
      zeros(1, length(measurementStdev))];
24 UB = [ 2, 20, 2, 2, 3, .98, 3, 200, 3, ones(1,
      length(measurementStdev))];
25
26 [theta, logLik, ~, ~, ~, grad, hessian] = fmincon(f, theta, [], [],
      [], [], LB, UB, [], opt);
27
28 stderr_theta = sqrt(diag(inv(hessian)));
29
30 mu = theta(1);
31 kappa = theta(2);
32 alpha = theta(3);
33 sigma_S = theta(4);
34 sigma_delta = theta(5);
35 rho = theta(6);
36 lambda = theta(7);
37 initial_S = theta(8);
38 initial_delta = theta(9);
39 measurementStdev = theta(10:end);
40
41 [logLik, results] = gibson_schwartz(prazos, precos, juros, dt,
      ...
42     mu, kappa, alpha, sigma_S, sigma_delta, rho, lambda, ...
43     initial_S, initial_delta, measurementStdev);
44 logS_t = permute(results.Xsmth(1, 1, :), [3 1 2]);
45 delta_t = permute(results.Xsmth(2, 1, :), [3 1 2]);
46 plot(logS_t);
47
48 end

```

Referências Bibliográficas

- [Bje91] Petter Bjerksund, *Contingent claims evaluation when the convenience yield is stochastic: Analytical results*.
- [BS85] Michael J. Brennan and Eduardo S Schwartz, *Evaluating natural resource investments*, *The Journal of Business* **58** (1985), no. 2, 135–157.
- [Car10] Ana Luiza A. R. S. Carvalho, *Calibration of the schwartz-smith model for commodity prices*, Instituto de Matemática Pura e Aplicada (2010).
- [GS90] Rajna Gibson and Eduardo S Schwartz, *Stochastic convenience yield and the pricing of oil contingent claims*, *Journal of Finance* **45** (1990), no. 3, 959–76.
- [Ham94] James D. Hamilton, *Time series analysis*, Princeton, 1994.
- [Mar11] Leonardo L. S. Marotta, *Calibração do modelo de schwartz-smith com filtro de kalman*, Instituto de Matemática Pura e Aplicada (2011).
- [Sch97] Eduardo S Schwartz, *The stochastic behavior of commodity prices: Implications for valuation and hedging*, *Journal of Finance* **52** (1997), no. 3, 923–73.
- [SS00] Eduardo S. Schwartz and James E. Smith, *Short-term variations and long-term dynamics in commodity prices*, *Management Science* **46** (2000), no. 7, 893–911.