

Processos Estocásticos dos Preços das Commodities: uma abordagem através do filtro de partículas*

Fernando Antonio Lucena Aiube[†], Tara Keshar Nanda Baidya[‡], Edison Americo Huarsaya Tito[§]

Sumário: 1. Introdução; 2. Trabalhos relevantes sobre o tema; 3. Comparação entre o filtro de Kalman e o filtro de partículas; 4. Extensão do modelo de Schwartz e Smith (2000); 5. Conclusões; A. O filtro de partículas.

Palavras-chave: Processos Estocásticos, Filtro de Kalman, Filtro de partículas.

Códigos JEL: G13, C13, C15.

O preço à vista de uma commodity pode ser decomposto em dois fatores estocásticos, o preço de equilíbrio de longo prazo e o desvio (ou variação) de curto prazo dos preços que por sua vez representa mudanças temporárias. Estes dois fatores não são observáveis diretamente no mercado. Nosso objetivo é estimá-los assumindo um movimento geométrico Browniano para os preços de equilíbrio de longo prazo e um processo de Ornstein-Uhlenbeck adicionado a um processo de Poisson, modelando saltos, para o desvio de curto prazo. Entretanto, a adição do processo de Poisson torna o modelo não Gaussiano. O filtro de Kalman não pode ser usado. Existe outra metodologia para estimação das variáveis não observáveis denominada filtro de partículas que é apropriada em tais casos. Utilizamos esta metodologia no processo de filtragem. Os resultados evidenciaram que o modelo com saltos explica melhor o comportamento empírico que aquele sem saltos.

*Os autores são gratos aos dois pareceristas pelos comentários e sugestões. Os erros remanescentes são de nossa responsabilidade.

[†] Engenheiro Petrobras, Prof. Depto de Enga. Industrial - PUC - RJ, DSc em Enga. de Produção - PUC - RJ. aiube@ind.puc-rio.br

[‡] Professor do Departamento de Engenharia Industrial - PUC - RJ, PhD Universidade da Califórnia, Berkeley. baidya@ind.puc-rio.br

[§] Engenheiro Consultor da Petrobras, DSc em Engenharia Elétrica - PUC - RJ. edison@ele.puc-rio.br



The spot prices of a commodity can be decomposed into two stochastic factors, the long term equilibrium price levels and the short term deviation in prices, which represents temporary changes. These two factors are not observable in the market. We consider that the long term equilibrium price follows a geometric Brownian motion and the short term deviations in prices follow an Ornstein-Uhlenbeck process with Poisson jumps. It is our objective to test this hypothesis. The addition of a Poisson process makes the model non-Gaussian. The Kalman filter methodology can not be used here. But there exists another methodology known as Particle filter method, which is appropriate in our case. We have used it and found that the model with jump process explains better than the model without it.

1. INTRODUÇÃO

A valoração de ativos reais e o apreçamento de derivativos financeiros constituem um dos ramos mais importantes de estudo em Finanças. Basta observar o enorme número de publicações tanto em livros como em periódicos. Este ramo de estudo em Finanças constitui uma de suas maiores vertentes. O rápido desenvolvimento de métodos de apreçamento de derivativos e a sua aplicabilidade cotidiana nas empresas e no mercado financeiro propiciaram o surgimento e o crescimento vertiginoso de bolsas de valores e mercadorias onde são negociados títulos à vista e derivativos por todo o mundo. Por outro lado, a existência de vigorosos mercados com grandes volumes negociados impulsionou a pesquisa na busca de novos produtos financeiros.

O apreçamento de um derivativo depende fundamentalmente do processo estocástico do ativo objeto. Por esta razão o conhecimento preciso deste processo é tão relevante e muitos artigos e estudos têm sido dedicados à sua estimação acurada. Quando este ativo é uma *commodity* surgem complicações adicionais. Em geral, as *commodities* não são negociadas no mercado à vista. Portanto, o preço à vista, S , não é uma variável observável. Em geral são negociadas em mercados futuros e o preço futuro, F , é a variável observável. O preço à vista S pode ser estimado a partir do preço futuro F . Não é por outra razão que normalmente o contrato futuro mais próximo é considerado uma “proxy” para o preço à vista. O exemplo mais trivial é o caso do petróleo. Dada a sua relevância no mundo, o petróleo é a *commodity* de maior negociação nas bolsas de futuros. Os preços que normalmente são veiculados pela imprensa diariamente referem-se exatamente aos preços do primeiro contrato futuro, aquele que tem vencimento no período de um mês à frente. Em geral são divulgados os preços do petróleo do tipo Brent negociado na bolsa de Londres e do tipo WTI (*West Texas Intermediate*) negociado em Nova Iorque. Estes preços futuros são as referências para o preço à vista do petróleo.

Existe uma relação muito conhecida entre os preços nos mercados à vista (S) e futuro (F): $F = S \exp((r - \delta)(\tau - t))$. A variável que faz a ligação entre F e S é δ , denominada retorno de conveniência. As demais variáveis são a taxa livre de risco r , a data do vencimento do contrato τ e a data atual t . O retorno de conveniência é o prêmio que faz jus o proprietário do ativo físico, ao contrário daquele que detém apenas um contrato futuro sobre a *commodity*. Para uma maior compreensão do papel do retorno de conveniência na valoração de ativos contingentes, veja Brennan (1991). Quando o retorno de conveniência é considerado constante, um simples ajuste no processo estocástico (*drift*) permite que os modelos de apreçamento sejam meras extensões do modelo de Black e Scholes (1973) e Merton (1973). Entretanto se δ é considerado constante, a volatilidade dos contratos futuros torna-se idêntica à volatilidade do mercado à vista. Evidências empíricas mostram que a volatilidade dos contratos futuros é decrescente com a maturidade dos contratos. Além disso, sabe-se que o retorno de conveniência varia inversamente com o nível dos estoques. Portanto, considerá-lo constante é apenas

uma aproximação. Os primeiros trabalhos em Opções Reais¹ adotaram esta hipótese simplificadora. O mais correto é considerá-lo uma variável estocástica. Assim, para descrever adequadamente o preço futuro de uma *commodity* são necessárias, pelo menos, duas variáveis estocásticas: S e δ . Estas variáveis são denominadas variáveis de estado, não são observáveis e devem ser estimadas. A observação de F (preço futuro) permite esta estimação. A importância de definir o preço à vista está relacionada à necessidade de seu uso em modelos de valoração de ativos reais (Opções Reais), planejamento de curto e longo prazo, decisão de proteção ou *hedging* e decisões de investimento de empresas e/ou instituições financeiras.

O modelo de Schwartz e Smith (2000) tem sido um dos mais utilizados na literatura. A questão fundamental aqui tratada está relacionada ao processo de estimação das variáveis de estado. Via de regra a literatura adota o filtro de Kalman como metodologia desta estimação. O filtro de Kalman pressupõe linearidade e Gaussianidade dos modelos. Nem sempre estas condições são atendidas na prática. Este trabalho propõe um modelo não Gaussiano para o preço à vista e faz uso de uma metodologia alternativa que pode ser aplicada quando os pressupostos do filtro de Kalman não se verificam. Trata-se do filtro de partículas, que flexibiliza as suposições de linearidade e Gaussianidade dos modelos.

O artigo foi organizado da seguinte maneira: a seção 2 resume os artigos mais relevantes que tratam diretamente de processos estocásticos para *commodities*; a seção 3 trata dos procedimentos empíricos utilizados para a comparação do filtro de Kalman e do filtro de partículas; a seção 4 apresenta o modelo com saltos para os preços das *commodities* e a seção 5 apresenta as conclusões.

2. TRABALHOS RELEVANTES SOBRE O TEMA

Brennan e Schwartz (1985) apresentam um relevante trabalho dentro da disciplina que mais tarde foi denominada de Opções Reais. Neste artigo os autores investigam a valoração, o gerenciamento e o instante ótimo de investir em um projeto de desenvolvimento de uma mina de cobre. Os autores modelam o preço futuro do cobre com um único fator estocástico, o preço à vista.

Pode-se dizer que o artigo de Gibson e Schwartz (1990) no início da década de 90 foi o primeiro passo na direção de modelos guiados por mais de um fator para preços de *commodities*. Os autores desenvolveram um modelo de dois fatores para os preços do petróleo. Os fatores que carregam as incertezas (ou variáveis estocásticas) são o preço à vista e o retorno de conveniência. O processo estocástico do preço à vista é o movimento geométrico Browniano. Para o retorno de conveniência foi adotado o processo de reversão à média. Os resultados mostraram que o retorno de conveniência tem a tendência de reverter para a sua média de longo prazo sendo a variável de estado de maior volatilidade, enquanto o preço à vista é menos volátil e sua evolução assemelha-se a um passeio aleatório. O modelo também foi capaz de mostrar a diferença entre a volatilidade dos preços à vista e aquela dos contratos futuros, evidenciando que quanto maior a maturidade menor a volatilidade.

Schwartz (1997) analisou o comportamento estocástico dos preços das *commodities*. O autor investiga o desempenho de três modelos distintos. O modelo de um fator considera que o logaritmo do preço à vista segue um processo de reversão à média. O segundo modelo considera dois fatores: (i) o primeiro fator estocástico é o preço à vista e segue um movimento geométrico Browniano, (ii) o segundo fator estocástico é o retorno de conveniência que evolui segundo um processo de reversão à média. O terceiro modelo agrega ao segundo a taxa de juros como variável estocástica, e esta segue um processo de reversão à média tal qual o modelo de Vasicek (1977). Nos três casos, as soluções analíticas do modelo apresentam o logaritmo do preço futuro como uma função linear das variáveis de estado. Apesar da linearidade, uma dificuldade persiste no problema: as variáveis de estado não são observáveis diretamente nos mercados, apenas o preço futuro é observável. O modelo foi então colocado na forma espaço-estado. E aproveitando a sua propriedade linear, o autor usou o filtro de Kalman para estimar as variáveis de

¹Dixit e Pindyck (1994) apresentam os fundamentos da Teoria das Opções Reais e os modelos clássicos desta nova disciplina.



estado e os hiperparâmetros do modelo. Os resultados evidenciaram que o modelo de um fator tem um desempenho fraco. Os modelos de dois e três fatores apresentam desempenhos equivalentes e o autor observa que a taxa de juros como um fator estocástico agrega pouca informação.

Schwartz e Smith (2000) propuseram um modelo de dois fatores à semelhança de trabalhos anteriores do primeiro autor. Neste modelo o logaritmo do preço à vista é decomposto como a soma de duas parcelas que são as variações de curto prazo e o preço de equilíbrio. Ou seja, $\ln(S_t) = \chi_t + \xi_t$, onde χ_t representa as variações de curto prazo e ξ_t os preços de equilíbrio. Estas duas variáveis de estado são os dois fatores incertos. A primeira é modelada por um processo de reversão à média do tipo Ornstein-Uhlenbeck e a segunda é modelada por um processo geométrico Browniano. No modelo derivado, o logaritmo do preço futuro é linear com relação às duas variáveis de estado. Esta linearidade possibilita a estimação dos estados usando o filtro de Kalman. Os hiperparâmetros são estimados pela maximização da verossimilhança. Os autores implementaram o modelo com dados dos contratos futuros de petróleo. Os resultados mostraram aderência para os dados empíricos no curto e longo prazos. Este modelo é formalmente equivalente ao modelo de Gibson e Schwartz (1990) e a variável χ_t é linearmente relacionada ao retorno de conveniência. O modelo pode ser adaptado para outras situações e pode ser estendido de maneira a considerar mais variáveis estocásticas.

Manoliu e Tompaidis (2000) usaram o modelo de Schwartz e Smith (2000) para analisar o comportamento dos preços do gás natural. Foi incluída uma função determinística para descrever a sazonalidade. Os parâmetros de reversão encontrados são significantes e a volatilidade para a variável de curto prazo (reversão à média) é muito superior a volatilidade da variável de longo prazo. Foi realizada uma comparação entre modelos de um e dois fatores através o erro de previsão um passo à frente. Os resultados mostram que o modelo de dois fatores permite melhor ajuste que o modelo de um fator.

Lucia e Schwartz (2001) apresentaram um estudo sobre os derivativos de energia elétrica. Fizeram uso dos modelos de um e dois fatores. No modelo de um fator é usado o processo de Ornstein-Uhlenbeck. Para o modelo de dois fatores são usados os modelos de reversão à média e geométrico Browniano. Os modelos são implementados usando dados de preços de energia da bolsa de energia dos países nórdicos (*Nordic Power Exchange, Nord Pool ASA*). É utilizada uma componente sazonal com variáveis *dummies* e funções senoidais. Os resultados mostraram que os parâmetros de sazonalidade são significantes componentes dos preços de energia elétrica na região de onde foi extraída a amostra. Esta sazonalidade tem um papel fundamental na definição de como os preços dos derivativos de energia são formados. O estudo conclui que outros modelos também poderiam ser especificados tais como um processo de reversão para a volatilidade. Outra especificação relevante seria a inclusão de saltos.

Sørensen (2002) apresentou um estudo sobre a modelagem da sazonalidade para os preços futuros de *commodities* agrícolas. Usou o modelo de dois fatores de Schwartz e Smith (2000) introduzindo uma componente determinística para modelar a sazonalidade. A sazonalidade foi modelada por uma combinação linear de funções trigonométricas com frequências sazonais. Foi implementado usando dados empíricos para as *commodities* milho, soja e trigo. Como resultado observou-se que as componentes sazonais acarretam um pico de preços dois a três meses antes do período de colheita. Outra verificação, que foi possível fazer a partir dos resultados, está relacionada ao fato de que em períodos de elevados estoques o retorno de conveniência é baixo, e vice-versa.

Javaheri et alii (2003) apresentaram um artigo onde é analisado o problema de filtragem em Finanças. São apresentadas as metodologias do filtro de Kalman, do filtro de Kalman estendido e do filtro de partículas. São mostradas aplicações destas metodologias: o filtro de Kalman é utilizado nos modelos de Schwartz (1997) e o filtro de partículas é utilizado na estimação da volatilidade estocástica do modelo de Heston (1993). O modelo de Heston considera que o processo de evolução dos preços é semelhante ao geométrico Browniano, porém com a variância não constante. Além disso, propõe para a variância um processo de reversão à média. Os dois processos estocásticos são correlacionados. Desta forma, o modelo de Heston leva em consideração: (a) que a distribuição dos preços não é lognormal; (b) a presença do “efeito alavanca” e (c) a reversão à média da volatilidade. A dificuldade de se trabalhar com este modelo reside no fato de que os modelos de apreçamento baseados no conceito da Medida Martingal

Equivalente (MME) não são viáveis, pois não é possível montar um portfólio sem risco devido, principalmente, ao fato de que a volatilidade não é um ativo negociável. Não sendo uma variável observável diretamente, a volatilidade pode ser estimada, e o filtro de partículas é útil para esta finalidade.

Pode-se notar que o modelo de dois fatores de Schwartz e Smith (2000) foi usado em trabalhos subsequentes para a *commodity* gás natural em Manoliu e Tompaidis (2000), para *commodities* agrícolas em Sørensen (2002). Além disso, permitiu extensões como em Lucia e Schwartz (2001) que analisa os preços da energia elétrica.

Em síntese, podemos dizer que a grande parte dos trabalhos faz uso do filtro de Kalman para a estimação das variáveis não observáveis. Além disso, a estimação dos hiperparâmetros dos modelos é feita pela maximização da verossimilhança do erro de previsão (veja Harvey (1989) capítulo 3). Tal fato decorre da condição dos modelos serem lineares e Gaussianos. Entretanto, as condições de não linearidade ou não Gaussianidade têm se tornado cada vez freqüente em Finanças. À medida que os modelos avançam na tentativa de descrever os fatos estilizados observados nas séries de preços, os modelos tornam-se mais complexos agregando tais propriedades. Este trabalho apresenta os resultados do uso do filtro de partículas como metodologia alternativa para lidar com as perdas de linearidade e Gaussianidade de um modelo com saltos, usado para modelar os preços do petróleo.

3. COMPARAÇÃO ENTRE O FILTRO DE KALMAN E O FILTRO DE PARTÍCULAS

A comparação entre os filtros de Kalman e de partículas será feita através do modelo de Schwartz e Smith (2000). Este é um modelo linear e Gaussiano e, portanto as duas metodologias podem ser empregadas sem restrições. O objetivo é mostrar que o filtro de partículas leva a resultados tão bons quanto o filtro de Kalman quando são requeridas as condições para o uso deste último. Adiante, na seção 4, mostraremos o uso do filtro de partículas para as situações em que o filtro de Kalman apresenta limitações. Por questão de brevidade será mostrado o modelo de Schwartz e Smith (2000) no que se refere aos aspectos gerais. Para o detalhamento da derivação sugerimos que seja consultado o artigo original. Trata-se de um modelo de dois fatores em que o logaritmo do preço à vista é decomposto em duas parcelas: a primeira modela as variações de curto prazo e a segunda os preços de equilíbrio de longo prazo. As duas variáveis de estado serão estimadas já que não são observáveis. Então, pode-se escrever:

$$\ln(S_t) = \chi_t + \xi_t \quad (1)$$

onde S_t é o preço à vista da *commodity*, χ_t representa as variações de curto prazo e ξ_t os preços de equilíbrio de longo prazo. A parcela χ_t , variações de curto prazo, segue um processo de Ornstein-Uhlenbeck:

$$d\chi_t = -k_\chi \chi_t dt + \sigma_\chi dW_\chi \quad (2)$$

onde k_χ é a velocidade de reversão, σ_χ é o parâmetro de difusão do processo e dW_χ é o incremento do processo padrão de Wiener. A parcela ξ_t , preço de equilíbrio, segue um processo geométrico Browniano, dado por:

$$d\xi_t = \mu_\xi dt + \sigma_\xi dW_\xi \quad (3)$$

onde μ_ξ é a tendência (*drift*) do processo, σ_ξ é o parâmetro de difusão e dW_ξ é o incremento do processo padrão de Wiener. Ainda temos que $dW_\chi dW_\xi = \rho dt$, onde ρ representa a correlação entre as variáveis de estado. O modelo acima foi escrito sob a MME e procedida a derivação dos preços futuros.

Seja $F_{\tau,t}$ o valor em t de um contrato futuro com vencimento em τ . Sob a MME, o preço do contrato futuro é igual ao valor esperado do preço à vista em τ . Isto é, $F_{\tau,t} = E_t^Q[S_\tau]$, ou ainda $\ln[F_{\tau,t}] = \ln(E_t^Q(S_\tau))$. Onde $E_t^Q(\cdot)$ representa o valor esperado em t sob a MME. Pode-se demonstrar que

$$\ln(F_{\tau,t}) = e^{-k_\chi(\tau-t)} \chi_t + \xi_t + A(\tau - t) \quad (4)$$



onde $A(\tau - t)$ é dado por

$$A(\tau - t) = \mu_{\xi}^*(\tau - t) - \left(1 - e^{-k_{\chi}(\tau - t)}\right) \frac{\lambda_{\chi}}{k_{\chi}} + \frac{1}{2} \left[\left(1 - e^{-2k_{\chi}(\tau - t)}\right) \frac{\sigma_{\chi}^2}{2k_{\chi}} + \sigma_{\xi}^2(\tau - t) + 2 \left(1 - e^{-k_{\chi}(\tau - t)}\right) \frac{\rho\sigma_{\chi}\sigma_{\xi}}{k_{\chi}} \right]$$

onde $\mu_{\xi}^* = \mu_{\xi} - \lambda_{\xi}$; λ_{χ} e λ_{ξ} são os preços de risco de mercado de χ e ξ , respectivamente.

Agora serão descritos os procedimentos empíricos que foram seguidos para estimação das variáveis de estado. As duas metodologias usadas para tal finalidade foram os filtros de Kalman e o de partículas.

O filtro de Kalman é um conjunto de equações matemáticas que constitui um processo recursivo eficiente de estimação uma vez que o erro quadrático é minimizado. Através da observação da variável denominada “variável de observação” outra variável, não observável, denominada variável de estado pode ser estimada eficientemente. Podem ser estimados os estados passados, o estado presente e mesmo previsto os estados futuros. Inicialmente o modelo é escrito na forma espaço-estado, e a seguir as equações do filtro de Kalman são usadas. Além disso, é realizada a estimação dos hiperparâmetros do modelo através da maximização da verossimilhança via decomposição do erro de previsão. Veja em Harvey (1989) e em Durbin e Koopman (2002) exposições detalhadas sobre o filtro de Kalman.

O filtro de partículas é um procedimento recursivo para integração, dentro da classe dos métodos sequenciais de Monte-Carlo. Os métodos sequenciais de Monte-Carlo são particularmente interessantes para o cálculo de distribuições *a posteriori*, isto é, das distribuições das variáveis de estado dada a ocorrência da observação y_t . Os métodos sequenciais de Monte-Carlo dispensam as condições de Gaussianidade e linearidade do modelo e ainda possuem adequadas condições de convergência. Eles abrangem os métodos que apareceram na literatura sob as denominações de filtros *bootstrap*, filtro de partículas, filtro de Monte-Carlo, dentre outros. O filtro de partículas baseia-se na distribuição por importância a qual é definida como sendo a distribuição *a priori* (*Sequential Importance Sampling*). Infelizmente o processo degenera-se quando o tempo cresce e não se consegue a distribuição *a posteriori*. Uma etapa adicional é acrescida para viabilizar o procedimento. Trata-se da amostragem com reposição. As partículas amostradas que são pouco representativas dentro da distribuição são retiradas do processo e as partículas “sobreviventes” constituem a distribuição, que é usada como base para o prosseguimento do método em cada instante de tempo. Os fundamentos do filtro de partículas podem ser encontrados em Doucet et alii (2001). No anexo deste artigo é apresentada uma resenha do filtro de partículas bem como o seu algoritmo básico.

O modelo de Schwartz e Smith (2000) está sob as condições requeridas para o uso do filtro de Kalman pois a equação das observações (equação (4)) é linear em relação aos estados χ_t e ξ_t . Também é imediato que o logaritmo de dos preços futuros é uma distribuição condicionalmente normal. Isto é, dado que se conhece os valores das variáveis de estado em $t - 1$, então χ_t e ξ_t são condicionalmente Gaussianas e por conseguinte a equação de observações define que o $\ln(F_{\tau,t})$ é também Gaussiana.

A comparação entre as duas metodologias foi realizada em dois estágios distintos. O primeiro estágio tratou o problema como um caso ideal. A variável de observação foi criada artificialmente. Foi então atendido o requisito da Gaussianidade do modelo para uso do filtro de Kalman. No segundo estágio da comparação os dados de observação são reais: os preços observados no mercado futuro de petróleo.

Considerando os hiperparâmetros do modelo como aqueles em Schwartz e Smith (2000), os preços futuros (variável de observação) serão gerados artificialmente.² Primeiramente as equações (2) e (3) são discretizadas abaixo. Estas equações são denominadas equações de transição:

$$\chi_t = (1 - k_{\chi}\Delta t)\chi_{t-1} + \varepsilon_t \quad (5)$$

² A estimação dos hiperparâmetros no modelo de Schwartz e Smith (2000) é realizada através da maximização da verossimilhança do erro de previsão, conforme a metodologia do filtro de Kalman. Este mesmo procedimento está implementado na seção 4. Para maiores detalhes veja Harvey (1989) e Durbin e Koopman (2002).

$$\xi_t = \mu_\xi \Delta t + \xi_{t-1} + \nu_t \quad (6)$$

onde $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ e $\nu_t \sim N(0, \sigma_\nu^2)$ sendo $\rho \Delta t$ a correlação entre as duas distribuições padronizadas e Δt o intervalo de tempo usado na discretização (uma semana ou 1/52 anos). As distribuições de χ_t e ξ_t são condicionalmente normais. Ou seja, dados χ_{t-1} e ξ_{t-1} , as variâncias condicionais e a matriz de covariância são dadas por: $Var(\chi_t | \chi_{t-1}) = (1 - e^{-2k_\chi \Delta t}) \frac{\sigma_\chi^2}{2k_\chi} = \sigma_\varepsilon^2$ e $Var(\xi_t | \xi_{t-1}) = \sigma_\xi^2 \Delta t = \sigma_\nu^2$

$$Cov(\chi_t, \xi_t) = \begin{bmatrix} (1 - e^{-2k_\chi \Delta t}) \frac{\sigma_\chi^2}{2k_\chi} & (1 - e^{-k_\chi \Delta t}) \frac{\rho \sigma_\chi \sigma_\xi}{k_\chi} \\ (1 - e^{-k_\chi \Delta t}) \frac{\rho \sigma_\chi \sigma_\xi}{k_\chi} & \sigma_\xi^2 \Delta t \end{bmatrix}$$

A equação das observações é dada por:

$$\ln(F_{\tau,t}) = e^{-k_\chi(\tau-t)} \chi_t + \xi_t + A(\tau - t) + u_t \quad (7)$$

onde $u_t \sim N(0, \sigma_u^2)$ representa os erros nas observações que são descorrelacionados dos erros das equações de transição (5) e (6).

Foram gerados simultaneamente T ($T = 2000$) ruídos Gaussianos para ε_t , para ν_t e ainda para u_t ($t = 1, 2, \dots, T$) de tal modo que os dois primeiros ruídos são correlacionados, conforme mencionado acima, e ainda descorrelacionados do terceiro ruído.

Em cada uma das três séries geradas foram realizados os testes estatísticos para verificação da adequabilidade dos dados: (i) testes de normalidade de Jarque-Bera e Anderson-Darling; (ii) análise dos correlogramas com uso da FAC (Função de Auto-Correlação), FACP (Função de Auto-Correlação Parcial) e da estatística Ljung-Box (estatística Q); (iii) teste de independência BDS (Brock, Dechert e Sheinkman (1987)). Os resultados indicam que os ruídos gerados são normais, descorrelatados e independentes. Além disso, as séries dos resíduos das variáveis de estado estão correlacionadas, enquanto a correlação destas mesmas séries com a série dos ruídos da variável de observação é baixíssima. Portanto, os valores gerados artificialmente para os ruídos estão adequados às premissas requeridas.

A etapa seguinte foi calcular para cada intervalo de tempo as variáveis de estado a partir de valores iniciais de χ_0 e ξ_0 . Para tal, foram usadas as equações discretizadas (5) e (6), bem como os ruídos gerados acima. Então, de posse dos valores calculados para as variáveis de estado e do ruído das observações, foi calculado, em cada instante de tempo, o valor da variável de observação para o primeiro contrato futuro (equação (7)). Os dados artificiais, assim gerados, serão os dados de entrada para o processo de filtragem. Eles serão utilizados no filtro de Kalman e no filtro de partículas. Faremos somente a filtragem dos dados pelas duas metodologias. Por enquanto não haverá estimação de hiperparâmetros por nenhum critério. O objetivo, até o momento, é tão somente de comparar os dois processos de filtragem dos dados observados, para estimação das variáveis de estado.

O preço do primeiro contrato futuro, gerado artificialmente, serviu como dado de entrada para os dois filtros. Esta é a variável de observação que será filtrada. A filtragem estimará as duas variáveis de estados χ_t e ξ_t por cada metodologia. Também será obtido o sinal de observação filtrado. Serão avaliados os erros entre o sinal observado e aquele gerado em cada metodologia de filtragem. Os filtros e todos os procedimentos citados acima foram implementados como uso do *software* MATLAB.

A variável de observação, gerada artificialmente, foi introduzida no filtro de Kalman e, juntamente com os mesmos parâmetros utilizados na sua geração, foi rodado o filtro. A saída fornece as variáveis de estado estimadas juntamente com a variável de observação filtrada. Foram realizados testes estatísticos nas saídas resultantes do filtro de Kalman. O resíduo da variável de observação foi analisado para a verificação das propriedades de sua distribuição. Foram realizados os testes de normalidade Jarque-Bera e Anderson-Darling e em ambos aceita-se a hipótese nula de normalidade. A FAC e a FACP, junto com a estatística Q, mostram que os dados são descorrelatados. O teste BDS mostra fortemente a aceitação da hipótese nula de independência.



O mesmo procedimento foi realizado em relação ao filtro de partículas. Ou seja, usando a mesma variável de observação anterior e os mesmos parâmetros, fizemos a filtragem com o filtro de partículas. Os dois procedimentos de filtragem foram comparados em termos quantitativos usando-se as seguintes medidas para o erro entre as variáveis observada e filtrada:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (y_{ot} - y_{ft})^2}{T}} \quad MAE = \sum_{t=1}^T \frac{|y_{ot} - y_{ft}|}{T} \quad MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_{ot} - y_{ft}}{y_{ot}} \right|$$

onde y_{ot} significa o valor da variável de observação no instante t e y_{ft} o valor da variável de observação filtrada no mesmo instante. A primeira medida RMSE é a raiz do erro quadrático médio, a segunda medida, MAE, é o erro absoluto médio e a terceira medida, MAPE, é o erro absoluto percentual médio.

Os valores iniciais das variáveis de estados foram os mesmos para as duas metodologias e iguais a $\chi_0 = 0,2$ e $\xi_0 = 3,2$ (para fazer estas estimativas dos valores iniciais foi rodada uma regressão da equação (7) sendo as variáveis χ e ξ os parâmetros). No filtro de Kalman todas as distribuições envolvidas são normais e apenas a média $\hat{x} = (\chi, \xi)$ e a variância P são suficientes para descrever as distribuições. O filtro de partículas é um procedimento numérico em que as amostras (partículas) descrevem as distribuições. Foram realizados inicialmente 100 experimentos onde cada experimento representa uma amostra diferente da distribuição inicial. Posteriormente, foram realizados 500 experimentos. Os resultados obtidos estão mostrados na Tabela 1.

Tabela 1 – Comparação entre as duas metodologias de filtragem para a variável de observação.

	RMSE	MAE	MAPE
	(US\$/bbl)	(US\$/bbl)	(%)
Filtro de Kalman	0,9954	0,7639	2,5849
Filtro de partículas (100)	1,0406	0,7917	2,6803
Filtro de partículas (500)	1,0402	0,7912	2,6780

Na Tabela 1 o filtro de Kalman mostra o erro cometido na solução exata. O algoritmo do filtro de partículas fornece a solução numérica e cada número da tabela representa a média do erro para o caso de 100 experimentos e 500 experimentos. Os resultados mostram que a solução numérica está bem próxima da solução exata. O tempo computacional requerido para cada experimento da solução numérica (filtro de partículas) é de aproximadamente 26 segundos (com o uso de 200 partículas³). Na solução exata (filtro de Kalman) os resultados são imediatos.

Agora comparamos as duas metodologias usando dados reais de mercado. A variável de observação é o primeiro contrato futuro (F1) de petróleo. Desta forma, o modelo é univariado nas observações, pois contém apenas uma variável de observação. Os dados abrangem os preços dos contratos futuros de petróleo WTI no período de janeiro de 1985 a abril de 2004 negociados no NYMEX (*New York Mercantile Exchange*), tomados com frequência semanal. O tamanho da amostra é de 1007 dados. A Tabela 2 mostra os resultados.

A comparação com dados reais de mercado abrange o período que envolve a guerra do Golfo, quando se observou preços anormalmente elevados relativamente aos preços daquela época. Notamos que o resíduo entre o preço observado e o preço filtrado é elevado durante este período, se comparado aos demais resíduos da série.

Embora não se tenha o objetivo de comparar os resultados dos dados artificiais com os dados de mercado, cumpre ressaltar um aspecto com relação à magnitude dos erros obtidos nos dois casos.

³Não existe um critério definitivo para determinar o número de partículas a ser utilizado. Um procedimento empírico é efetuar simulações com diferentes números de partículas e observar o nível em que os resultados estabilizam-se.

Tabela 2 – Comparação entre as duas metodologias de filtragem para a variável de observação

	RMSE	MAE	MAPE
	(US\$/bbl)	(US\$/bbl)	(%)
Filtro de Kalman	0,5242	0,3675	1,6970
Filtro de partículas (100)	0,5810	0,3965	1,8282
Filtro de partículas (500)	0,5812	0,3963	1,8279

Observa-se que com dados artificiais o erro foi superior do que com dados de mercado. Deve-se isto simplesmente ao fato de que o desvio padrão do ruídos das observações que foi utilizado é superior àquele dos dados de mercado. Outras rodadas foram realizadas em que o desvio padrão do ruído das observações era menor que àquele com dados reais. Nesta situação, o erro apresentado pelos dados artificiais foi menor que os de mercado. Cumpre ainda lembrar que o modelo que está sendo usado é linear e Gaussiano. O menor erro do filtro de Kalman retrata o quão distante estão os dados da condição de Gaussianidade quando são tratados por tal modelo.

4. EXTENSÃO DO MODELO DE SCHWARTZ E SMITH (2000)

A seção anterior deteve-se em examinar tão somente o processo de filtragem pelas duas metodologias que estimam as variáveis não observáveis. Agora trataremos do processo de estimação dos hiperparâmetros. Nesta seção será proposto um modelo não Gaussiano e analisado através do filtro de partículas. Serão estimados as variáveis estado (filtro de partículas) e simultaneamente os hiperparâmetros do modelo (maximização da verossimilhança do erro de previsão).

Um dos fatos estilizados das séries de retornos financeiros é o excesso de curtose (caudas pesadas). A literatura tem modelado esta propriedade através de: (i) volatilidade variando no tempo (modelos da família ARCH/GARCH), (ii) volatilidade estocástica e (iii) processos de difusão com saltos. Este artigo usa esta última abordagem propondo uma extensão do modelo de Schwartz e Smith (2000). A utilização de saltos em processos estocásticos dos preços de *commodities* pode ser encontrado em Deng (1998) que analisou os derivativos dos preços da energia elétrica. Considere o modelo dois fatores com saltos, proposto abaixo:

$$\begin{aligned}\ln(S_t) &= f(t) + \chi_t + \xi_t \\ d\chi_t &= -k_\chi \chi_t dt + \sigma_\chi dW_\chi + \nu(\mu_\nu, \sigma_\nu^2) d\pi(\varpi_t) \\ d\xi_t &= \mu_\xi dt + \sigma_\xi dW_\xi\end{aligned}\tag{8}$$

A definição dos parâmetros das equações em (8) são os mesmos já apresentados no modelo de Schwartz e Smith (2000). As duas variáveis de estados são correlacionadas por $\rho dt = dW_\chi dW_\xi$. Foi adicionado um processo de saltos na variável χ por ser esta a variável que possui maior volatilidade. A variável ξ representa a componente de equilíbrio de longo prazo. Para representar um modelo de forma mais genérica incluímos uma componente sazonal representada pela função determinística $f(t)$. Intuitivamente, o salto acima possui tamanho aleatório dado pela distribuição ν e o processo de Poisson tem intensidade ϖ . Os processos de Poisson e de difusão são independentes um do outro e cada um deles é independente de ν . Note que os processos do modelo acima são afins. Os processos de difusão afins são muito comuns em Finanças, pois são “bem equipados” no sentido que possuem a componente de difusão, a componente dos saltos e ainda permitem agregar componentes de volatilidade estocástica ou volatilidades variando no tempo tais com os processos ARCH/GARCH. Em Duffie e Kan (1996) e em Duffie



et alii (2000) há uma resenha sobre os processos de difusão afins⁴. Foi considerado que a distribuição ν é normal⁵.

A seguir, foi derivada a expressão do preço futuro de um contrato que tem maturação em τ . O preço deste contrato no instante t é dado por $F_{\tau,t} = E_t^Q[S_\tau]$. Q representa a Medida Martingal Equivalente sob a qual os processos descritos nas equações em (8) foram escritos. O valor esperado acima pode ser calculado com a transformada proposta em Duffie e Kan (1996), ou ainda, a mesma transformada pode ser encontrada em Duffie et alii (2000), Proposição 1. Os detalhes desta derivação estão em Aiube (2005). O preço futuro, escrito sob a MME, (equação das observações) é dado por:

$$\ln(F_{\tau,t}) = f(\tau) + e^{-k_\chi(\tau-t)}\chi_t + \xi_t + A(\tau-t) + B(\tau-t) \quad (9)$$

onde:

$$A(\tau-t) = (\mu_\xi^*(\tau-t) - \frac{\lambda_\chi}{k_\chi}(1 - e^{-k_\chi(\tau-t)}) + \frac{\sigma_\chi^2}{4k_\chi}(1 - e^{-2k_\chi(\tau-t)}) + \frac{\rho\sigma_\chi\sigma_\xi}{k_\chi}(1 - e^{-k_\chi(\tau-t)}) + \frac{1}{2}\sigma_\xi^2(\tau-t))$$

$$B(\tau-t) = \varpi \int_t^\tau \left[\exp\left(\mu_\nu^* e^{-k_\chi(\tau-z)} + \frac{1}{2}\sigma_\nu^2 e^{-2k_\chi(\tau-z)}\right) - 1 \right] dz$$

$\mu_\xi^* = \mu_\xi - \lambda_\xi$ e ainda μ_ν^* é a média do tamanho dos saltos sob a MME.

A inclusão de saltos no modelo gerou o aparecimento do termo $B(\tau-t)$ que não existia no modelo de Schwartz e Smith (2000). Para obter a equação deste último (equação (4)) basta retirar os saltos, removendo o termo $B(\tau-t)$.

A equação em (9) é escrita na forma da equação de observação como:

$$y_{\tau,t} = f(\tau) + e^{-k_\chi(\tau-t)}\chi_t + \xi_t + A(\tau-t) + B(\tau-t) + u_t$$

onde: $y_{\tau,t} = \ln(F_{\tau,t})$; $u_t \sim N(0,s)$ e s é o desvio padrão do ruído das observações.

A variável de observação é o preço de um contrato futuro com vencimento em τ . São observadas séries históricas de preços futuros para diferentes maturidades. As observações são agregadas em um painel. O procedimento de agregar os dados em painéis é o mesmo dos artigos mencionados na seção 2. A seguir serão apresentados mais detalhes sobre os dados.

O logaritmo do preço futuro é linear em relação aos estados como pode-se observar da equação (9). Os dois processos estocásticos descritos em (8) constituem as equações de transição ou de evolução dos estados. O problema não é Gaussiano porque uma das variáveis de estado (χ_t) segue um processo que possui saltos. A estimação dos estados é feita através do filtro de partículas. Os hiperparâmetros a serem estimados são dados pelo vetor Θ , este vetor será especificado abaixo em cada modelo. O filtro de partículas inicia o processo de filtragem com o vetor de parâmetros Θ_0 . As variáveis de estado e as observações filtradas são estimadas para cada instante de tempo. Através da decomposição do erro de previsão é estimada a verossimilhança do modelo. Ao final da primeira filtragem este valor é avaliado conforme a equação abaixo:

$$\ln L(\Theta) = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |\mathbf{G}_t| - \frac{1}{2} \sum_{t=1}^T \varepsilon_t \mathbf{G}_t^{-1} \varepsilon_t$$

⁴Além dos artigos mencionados, o site <http://www4.ncsu.edu/unity/users/p/pfackler/www/affine/> contém vários outros artigos e links relacionados a utilização de processos afins em Finanças.

⁵ Duffie e Kan (1996) apontam que as distribuições exponencial, binomial e normal são convenientes por levarem a soluções fechadas. Em outras situações deve-se proceder a solução numérica.

onde G_t representa a matriz de covariância dos erros das observações em t ; $\varepsilon_t = y_t - y_{t|t-1}$ é o erro de previsão em t , para $t = 1, 2, \dots, T$; e N o número de contratos no painel. O algoritmo utilizado na maximização da verossimilhança foi o BHHH (Berndt, Hall, Hall e Hausman) e os programas foram implementados em MATLAB.

Com a finalidade de estabelecer comparações, foram analisados dois modelos pela metodologia do filtro de partículas. O primeiro é o modelo Gaussiano de Schwartz e Smith (2000) descrito na seção anterior. O segundo modelo é o proposto acima que inclui saltos e, portanto é não Gaussiano.

Os modelos foram implementados usando dados empíricos dos contratos futuros da *commodity* petróleo WTI negociados NYMEX. Os dados observados dos contratos futuros para diferentes maturidades foram organizados em um painel. Os dados foram coletados de maneira que o painel fosse completo, ou seja, não há dados faltantes⁶. A maturidade dos contratos abrange do primeiro contrato futuro (com vencimento um mês à frente), denominado de F1, até o contrato com vencimento em sete anos, denominado por F84. O painel contém sete contratos futuros. O período de observação foi de 23/04/1997 a 20/12/2004, totalizando 403 observações. A Tabela 3 apresenta algumas propriedades estatísticas destes contratos.

Tabela 3 – Estatísticas do painel de dados utilizado para a estimação

Contrato	Média (US\$/bbl)	Desvio Padrão (US\$/bbl)	Curtose
F1	26,31	8,71	3,27
F5	25,38	7,77	4,08
F10	24,31	6,93	4,71
F17	23,33	6,21	5,41
F27	22,62	5,60	5,95
F48	22,19	5,16	6,20
F84	21,95	4,69	5,93

No modelo sem saltos o vetor Θ é dado por $\Theta = (\mu_\xi^*, k_\chi, \lambda_\chi, \sigma_\chi, \sigma_\xi, \rho, s_1, \dots, s_7)$, onde s_i é desvio padrão do ruído das observações do i -ésimo contrato. No caso do petróleo bruto a componente sazonal não é significativa. Foram realizados testes considerando a sazonalidade representada por uma soma de senos e cossenos. A sazonalidade trimestral mostrou que os coeficientes das funções trigonométricas são estatisticamente insignificantes. A Tabela 4 apresenta os resultados da estimação do modelo de Schwartz e Smith (2000), ou seja, o modelo sem saltos. O resultado mostra que a volatilidade de curto prazo (σ_χ) neste modelo é substancialmente superior que aquela dos preços de equilíbrio (σ_ξ).

No modelo com saltos o vetor Θ é $\Theta = (\mu_\xi^*, k_\chi, \lambda_\chi, \sigma_\chi, \sigma_\xi, \rho, \varpi, \mu_\nu^*, \sigma_\nu, s_1, \dots, s_7)$. A Tabela 5 mostra os resultados obtidos. No modelo sem saltos o choque nos preços estava sendo explicado, principalmente, pela força da reversão à média (parâmetro k_χ) e pela volatilidade de curto prazo (parâmetro σ_χ). A inclusão dos saltos no modelo coloca em ação outro parâmetro capaz de explicar os choques nos preços. Se comparado ao modelo sem saltos, nota-se que a volatilidade σ_χ tem o seu valor reduzido de 36% ao ano para aproximadamente 18% ao ano. De fato este resultado era esperado como pode ser constatado em Das (1998) que incluiu saltos na dinâmica da taxa de juros. Também pode-se notar que na velocidade de reversão k_χ houve redução. A presença dos saltos absorve, em parte, o que estes parâmetros representavam no modelo sem saltos. Isto é, grande parte dos choques existentes na série de preços estava traduzido anteriormente em termos de k_χ e σ_χ . Agora os choques são também explicados pelos saltos.

⁶Tratamos o problema montando painéis completos (sem ausência de informações de preços futuros). Pode-se trabalhar com painéis incompletos, para tal veja Sørensen (2002) e Cortazar e Naranjo (2003).

**Tabela 4** – Estimação dos hiperparâmetros do modelo sem saltos

Parâmetro	Valor	Erro Padrão
μ_{ξ}^*	0,0097	0,0007
k_{χ}	1,4070	0,0155
λ_{χ}	0,0780	0,0783
σ_{χ}	0,3606	0,0124
σ_{ξ}	0,1500	0,0041
ρ	0,2188	0,0446
s_1	0,0092	0,0001
s_2	0,0477	0,0000
s_3	0,0103	0,0000
s_4	0,0092	0,0000
s_5	0,0100	0,0001
s_6	0,0200	0,0006
s_7	0,0156	0,0007
Verossimilhança: $\ln L(\Theta) = 8152$		

Foi realizada a análise da estrutura a termo dos preços para os dois modelos. Isto é, a diferença entre o valor filtrado e aquele observado. A Tabela 6 mostra estes resultados. Cada número representa a média dos erros para os contratos do painel.

A inclusão de saltos no modelo resultou em um menor erro médio para o dados do painel, ou seja, o modelo com saltos descreve mais apropriadamente a evolução dos preços. Em Aiube (2005) pode-se observar que esta mesma melhora nos resultados ocorre quando a distribuição dos saltos é do tipo exponencial. Além disso, nota-se também que a distribuição normal (para ν) é mais apropriada que a exponencial para os preços do petróleo.

5. CONCLUSÕES

Este artigo apresentou inicialmente uma comparação entre os processos de filtragem das metodologias do filtro de partículas e do filtro de Kalman. A comparação foi realizada tanto com dados sintéticos como com dados de mercado e sempre fazendo uso de um modelo Gaussiano. A solução numérica pelo filtro de partículas mostrou resultados próximos daqueles obtidos pela solução exata do filtro de Kalman. Portanto, pode-se concluir que o filtro de partículas apesar de computacionalmente mais intenso é viável em termos de tempo de resposta e apresenta resultados coerentes em termos de filtragem, quando comparado ao filtro de Kalman.

O modelo de Schwartz e Smith (2000) é muito flexível e pode acomodar inúmeras extensões. Além disso, este modelo tem sido usado muito freqüentemente na análise de diferentes *commodities* tais como o gás natural, energia elétrica, produtos agrícolas e petróleo.

O artigo propõe uma extensão do modelo de Schwartz e Smith (2000) incorporando saltos nas variações de curto prazo. Na derivação da expressão dos preços futuros foi feita uma aplicação das transformadas apresentadas por Duffie e Kan (1996) e Duffie et alii (2000) para processos estocásticos afins. A dinâmica proposta com saltos restringe o uso do filtro de Kalman como metodologia de estimação. Aplicou-se então o filtro de partículas e a comparação dos modelos com e sem saltos mostrou que a inclusão dos saltos explica melhor a estrutura a termo dos preços do petróleo.

A maior contribuição do artigo está relacionada à análise de viabilidade do uso do filtro de partículas como metodologia alternativa ao filtro de Kalman clássico. O filtro de partículas é uma metodologia que tem sido usada em vários ramos do conhecimento e certamente é muito promissora em Finanças, pois

Tabela 5 – Estimação dos hiperparâmetros do modelo com saltos

Parâmetro	Valor	Erro Padrão
μ_{ξ}^*	0,0101	0,0000
k_{χ}	1,0155	0,0000
λ_{χ}	0,0100	0,0000
σ_{χ}	0,1775	0,0000
σ_{ξ}	0,1423	0,0000
ρ	0,1800	0,0002
ϖ	0,3222	0,0000
μ_{ν}^*	0,3714	0,0000
σ_{ν}	0,1001	0,0001
s_1	0,0103	0,0000
s_2	0,0123	0,0000
s_3	0,0096	0,0000
s_4	0,0100	0,0000
s_5	0,0477	0,0000
s_6	0,0304	0,0000
s_7	0,0145	0,0000
Verossimilhança: $\ln L(\Theta) = 11859$		

Tabela 6 – Estrutura a termos dos preços

Modelo	RMSE	MAE	MAPE
	US\$/bbl	US\$/bbl	%
Sem saltos	1,0331	0,8454	3,5931
Com Saltos	0,8566	0,6720	2,7936

permite lidar com os modelos cada vez mais complexos que buscam melhor explicar os fenômenos relacionados à dinâmica dos dados financeiros.

Referências Bibliográficas

- Aiube, F. A. L. (2005). *Modelagem dos preços futuros de commodities: abordagem pelo filtro de partículas*. Tese de Doutorado, Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.
- Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.
- Brennan, M. & Schwartz, E. (1985). Evaluating Natural Resource Investments. *Journal of Business*, 58(2):135–157.
- Brennan, M. J. (1991). The Price of Convenience and the Valuation of Commodity. In Lund, D. & Øksendal, B., editors, *Stochastic Models and Option Values*. North-Holland.
- Brock, W., Dechert, W. D., & Scheinkman, J. (1987). A Test for Independence Based on the Correlation Dimension. Technical report, University of Wisconsin, Department of Economics; University of Houston e University of Chicago.



- Cortazar, G. & Naranjo, L. (2003). A multi-factor Stochastic Model and Estimation Procedure for the Valuation and Hedging of Commodity Contingent Claims. Technical report, Pontificia Universidade Catolica de Chile.
- Das, S. R. (1998). Poisson-Gaussian Processes and the Bond Markets. Technical Report 6631, National Bureau of Economic Research. Disponível em <http://www.nber.org/papers/w6631>.
- Deng, S. (1998). Stochastic Models of Energy Commodity Prices and Their Applications: Mean-reversion with Jumps and Spikes. Technical report, University of California at Berkeley.
- Dixit, A. & Pindyck, R. (1994). *Investment under uncertainty*. Princeton University Press Princeton, NJ.
- Doucet, A., De Freitas, N., & Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York.
- Doucet, A., Godsill, S. J., & Andrieu, C. (2000). On Sequential Simulation-Based Methods for Bayesian Filtering. *Statistics and Computing*, 10(3):197–208.
- Duffie, D. & Kan, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance*, 6(4):379–406.
- Duffie, D., Pan, J., & Singleton, K. J. (2000). Transformation Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68(6):1343–1376.
- Durbin, J. & Koopman, S. J. (2002). *Time Series Analysis by State Space Methods*. Oxford University Press.
- Gibson, R. & Schwartz, E. S. (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *The Journal of Finance*, 45(3):959–976.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2):327–343.
- Javaheri, A., Lautier, D., & Galli, A. (2003). Filtering in finance. *Willmott Magazine*, 5:67–83.
- Lucia, J. J. & Schwartz, E. S. (2001). Electricity Prices and Power Derivatives: evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5(1):5–50.
- Manoliu, M. & Tompaidis, S. (2000). Energy Futures Prices: Term Structure Models with Kalman Filter Estimation. Technical report, University of Texas, Austin.
- Merton, R. C. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183.
- Schwartz, E. & Smith, J. E. (2000). Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46(7):893–911. Adendo: Short-term variations and long-term dynamics in commodity prices: Incorporating a stochastic growth rate.
- Schwartz, E. S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52(3):923–973.
- Sørensen, C. (2002). Modeling Seasonality in Agricultural Commodity Futures. *Journal of Futures Markets*, 22(5):393–426.
- Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5(2):177–188.

A. O FILTRO DE PARTÍCULAS

Este Apêndice trata da apresentação do filtro de partículas em seus conceitos básicos. O filtro de partículas é um método numérico de integração. É adequado para lidar com problemas não lineares e não Gaussianos. Desde a década de sessenta, grande atenção tem sido devotada a estes problemas. Entretanto somente com o aumento do poder computacional foi possível tornar o seu uso mais corrente. Os conceitos aqui apresentados estão baseados em Doucet et alii (2001). Veja também Doucet et alii (2000). Foi mantida a mesma notação dos artigos técnicos que tratam dos métodos sequenciais de Monte-Carlo como pode ser observado através das referências aqui listadas.

A.1. Introdução

Em vários ramos do conhecimento é comum o problema relacionado à necessidade de estimar uma variável desconhecida ou não observável a partir de um conjunto de dados. Também é freqüente que estes dados ou observações sejam gerados em tempo real e desejável que a estimação seja feita em tempo real. Na grande parte dos casos é conhecida a distribuição anterior (ou *a priori*) do fenômeno ou da variável não observável. Com o uso do Teorema de Bayes é possível calcular a distribuição posterior (*a posteriori*). Esta distribuição permite a inferência sobre a variável que se deseja estimar. Em resumo, trata-se do mesmo problema que o filtro de Kalman aborda. O filtro de Kalman pode ser visto como um caso particular em que o modelo é linear e as distribuições das variáveis de estado e observação são Gaussianas. O filtro de Kalman é a solução analítica exata para problemas Gaussianos. Portanto, no algoritmo do filtro de Kalman apenas os dois primeiros momentos da distribuição das variáveis de estado são relevantes. A aplicabilidade do filtro de partículas está relacionada aos mesmos problemas de controle em sistemas dinâmicos que trata o filtro de Kalman. Entretanto, aborda uma classe mais ampla de problemas, pois não tem as mesmas restrições do filtro de Kalman. Por outro lado, possui a desvantagem de ser computacionalmente muito mais intenso. Exemplos de aplicações do filtro de partículas são diversos no ramo da engenharia. Esta pesquisa propõe o seu uso para estimar variáveis de estado não observáveis em mercados financeiros. Um simples exemplo é o caso da volatilidade de ativos. Esta variável não é diretamente observada no mercado. É extraída das séries financeiras sendo dispensável ressaltar sua importância para os mercados. A cada nova informação do preço de um ativo financeiro que normalmente, vem corrompida por um ruído, este parâmetro pode ser filtrado e estimado em tempo real. Outro exemplo está relacionado aos mercados futuros de *commodities*. A cada nova informação de preço de negociação no mercado futuro para uma *commodity*, o preço à vista que, via de regra, não é observável, pode ser estimado. Portanto, em caso de Gaussianidade, como dito acima, existem soluções fechadas para a distribuição posterior e este conjunto de equações recursivas é o filtro de Kalman.

Entretanto, nem sempre os fenômenos podem ser tratados por modelos lineares. Uma solução sempre à vista é linearizar tais modelos para que possa recair sob condições mais fáceis para se trabalhar. É o caso do filtro de Kalman estendido. Ainda mais, nem sempre é possível trabalhar com modelos Gaussianos. Ou seja, os modelos Gaussianos restringem o conjunto de fenômenos que podem ser analisados. A não linearidade e a não Gaussianidade são propriedades comumente encontradas. Nestes casos a determinação da distribuição posterior depende de intensos cálculos numéricos para ser determinada, uma vez que não há solução analítica fechada. Esquemas de aproximações numéricas têm sido pesquisados nos últimos trinta anos com esta finalidade. Os métodos sequenciais de Monte-Carlo (MC) destacam-se com vantagens sobre os demais no que se refere ao cálculo da distribuição posterior. Além disso, o aparecimento de computadores potentes possibilitou o rápido crescimento desta disciplina. Os algoritmos sequenciais de MC apareceram na literatura sob diversas denominações: filtros *bootstrap*, condensação, filtro de partículas, filtro de Monte-Carlo. Aqui todos serão tratados por filtro de partículas.



A.2. Definições básicas

Seja $\{x_t, t \in \mathbb{N}\}$ a variável de estado, não observável, que segue um processo de Markov. Seja $\{y_t, t \in \mathbb{N}\}$ a variável observável. Ainda sejam $p(x_0)$ a distribuição inicial da variável de estado e $p(x_t|x_{t-1})$ a equação de transição que descreve a evolução da variável de estado. A distribuição marginal da variável de observação é dada por $p(y_t|x_t)$. Ou seja, as observações $\{y_t, t \in \mathbb{N}\}$ são condicionalmente independentes dado $\{x_t, t \in \mathbb{N}\}$. Em resumo:

$p(x_0)$ distribuição inicial da variável de estado,

$p(x_t|x_{t-1})$ equação de transição da variável de estado,

$p(y_t|x_t)$ distribuição marginal da variável de observação.

Dentro do ambiente Bayesiano toda informação relevante sobre $\{x_0, x_1, \dots, x_t\}$ dadas as observações até t , podem ser obtidas a partir da distribuição posterior $p(x_0, x_1, \dots, x_t|y_1, y_2, \dots, y_t)$. Em várias aplicações há o interesse em determinar recursivamente, em tempo real, esta distribuição, $p(x_{0:t}|y_{1:t})$. Além disso, deseja-se saber a distribuição marginal $p(x_t|y_{1:t})$, conhecida como distribuição filtrada. Ainda mais, deseja-se saber valores esperados tais como:

$$I(f_t) = \int f_t(x_{0:t})p(x_{0:t}|y_{1:t})dx_{0:t} \quad (10)$$

onde $f_t: \mathbb{R}^{t+1} \rightarrow \mathbb{R}$ é integrável com respeito a $p(x_{0:t}|y_{1:t})$. Assim estabelecido, o problema acima é conhecido como problema de filtragem Bayesiano ou problema ótimo de filtragem.

Em um instante de tempo t qualquer, o Teorema de Bayes fornece a distribuição posterior:

$$p(x_t|y_{1:t}) = \frac{p(y_{1:t}|x_{0:t})p(x_t)}{\int p(y_{1:t}|x_{0:t})p(x_{0:t})dx_{0:t}} \quad (11)$$

Ainda pode-se obter uma fórmula recursiva para avaliar a distribuição posterior:

$$p(x_{0:t+1}|y_{1:t+1}) = p(x_{0:t}|y_{1:t}) \frac{p(y_{t+1}|x_{t+1})p(x_{t+1}|x_t)}{p(y_{t+1}|y_{1:t})} \quad (12)$$

A distribuição marginal $p(x_t|y_{1:t})$ também pode ser obtida de forma recursiva a partir das seguintes equações:

$$p(x_t|y_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1} \quad (13)$$

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t} \quad (14)$$

A equação (13) é a equação de previsão e a equação (14) é a equação de atualização. Estas duas equações constituem a base dos algoritmos sequenciais de Monte-Carlo.

Quando as equações de transição e observação são lineares e as distribuições são Gaussianas as integrais em (13) e (14) possuem soluções analíticas fechadas. A solução ótima obtida é exatamente o filtro de Kalman.

A.3. Amostragem por importância

O cálculo da integral em (10) é função da densidade posterior. Da mesma forma, toda a descrição da variável não observável $\{x_{0:t}, t \in \mathbb{N}\}$ é obtida da densidade posterior. Dentro do ambiente Bayesiano a distribuição posterior desempenha papel fundamental.

À exceção de casos lineares e Gaussianos, o cálculo da distribuição posterior e dos estimadores Bayesianos são proibitivamente complexos. Para transpor esta dificuldade o filtro de partículas adota uma abordagem baseada em simulação cuja técnica básica é denominada amostragem por importância. O objetivo é estimar a densidade de probabilidade posterior e a ideia central do filtro de partículas é representar tais densidades por conjunto de partículas.

Adota-se uma distribuição $\pi(x_{0:t}|y_{1:t})$ denominada distribuição por importância. A amostragem será feita a partir desta distribuição e serão geradas amostras iid. Assim, seja a integral em (10):

$$I(f_t) = \int f_t(x_{0:t})p(x_{0:t}|y_{1:t})dx_{0:t}$$

É válida a seguinte identidade:

$$I(f_t) = \frac{\int f_t(x_{0:t})p(x_{0:t}|y_{1:t})dx_{0:t}}{\int p(x_{0:t}|y_{1:t})dx_{0:t}}$$

Multiplicando e dividindo por $\pi(x_{0:t}|y_{1:t})$ numerador e denominador:

$$I(f_t) = \frac{\int f_t(x_{0:t}) \frac{p(x_{0:t}|y_{1:t})}{\pi(x_{0:t}|y_{1:t})} \pi(x_{0:t}|y_{1:t}) dx_{0:t}}{\int p(x_{0:t}|y_{1:t}) \frac{\pi(x_{0:t}|y_{1:t})}{\pi(x_{0:t}|y_{1:t})} dx_{0:t}}$$

Definindo $\omega(x_{0:t}) = \frac{p(x_{0:t}|y_{1:t})}{\pi(x_{0:t}|y_{1:t})}$ resulta em:

$$I(f_t) = \frac{\int f_t(x_{0:t})\omega(x_{0:t})\pi(x_{0:t}|y_{1:t})dx_{0:t}}{\int \omega(x_{0:t})\pi(x_{0:t}|y_{1:t})dx_{0:t}} \quad (15)$$

onde $\omega(x_{0:t})$ é denominado peso de importância. Um estimador da integral em (15) é dado por:

$$\hat{I}(f_t) = \frac{\frac{1}{N} \sum_{i=1}^N f_t(x_{0:t}^{(i)})\omega(x_{0:t}^{(i)})}{\frac{1}{N} \sum_{j=1}^N \omega(x_{0:t}^{(j)})} = \sum_{i=1}^N f_t(x_{0:t}^{(i)})\omega_t^{*(i)}$$

onde $\omega_t^{*(i)} = \frac{\omega(x_{0:t}^{(i)})}{\sum_{j=1}^N \omega(x_{0:t}^{(j)})}$ são os pesos de importância normalizados. Sob condições fracas, a lei forte

dos grandes números é aplicável, isto é, $\hat{I}_N(f_t) \xrightarrow{qc} I(f_t)$. O método acima apesar de ser um método genérico de integração de Monte-Carlo (MC), não possui uma forma recursiva.

A.4. Amostragem por importância sequencial

Seja o caso em que a distribuição por importância em t é obtida da distribuição por importância em $t-1$:

$$\pi(x_{0:t}|y_{1:t}) = \pi(x_{0:t-1}|y_{1:t-1})\pi(x_t|x_{0:t-1},y_{1:t}) \quad (16)$$

Rekursivamente pode-se obter:

$$\pi(x_{0:t}|y_{1:t}) = \pi(x_0) \prod_{k=1}^t \pi(x_k|x_{0:k-1},y_{1:k}) \quad (17)$$

Agora os pesos de importância são obtidos por:

$$\omega_t^{*(i)} = \omega_{t-1}^{*(i)} \frac{p(y_t|x_t^{(i)})p(x_t^{(i)}|x_{t-1}^{(i)})}{\pi(x_t^{(i)}|x_{0:t-1},y_{1:t})} \quad (18)$$



A escolha da distribuição por importância é essencial, pois ela determinará a eficiência e a complexidade do filtro de partículas.

Se for adotada como distribuição por importância a distribuição posterior, então a distribuição por importância será dada por $\pi(x_{0:t}|y_{1:t}) = p(x_t|x_{0:t}, y_{1:t})$. Seria a condição ideal no sentido de que a variância dos pesos por importância é minimizada e em consequência melhores estimativas podem ser alcançadas. Entretanto, o cálculo dos pesos é muito complexo pois envolve a avaliação de $p(y_t|x_{t-1}, y_{1:t-1})$ através de integrações onerosas. Este fato exclui a distribuição posterior como uma boa candidata para ser usada como distribuição por importância.

Se for adotada como distribuição por importância a distribuição anterior, a equação (17) transforma-se em:

$$\pi(x_{0:t}|y_{1:t}) = p(x_0) \prod_{k=1}^t p(x_k|x_{k-1}) \quad (19)$$

Trata-se de uma situação especial, pois é de fácil implementação. Quando a distribuição por importância é a distribuição anterior, a amostragem por importância recebe a denominação de amostragem por importância sequencial. Neste caso os pesos por importância são dados por:

$$\omega_t^{*(i)} = \omega_{t-1}^{(i)} p(y_t|x_t^{(i)}) \quad (20)$$

Agora a função $p(y_t|x_t^{(i)})$ é a função de verossimilhança em t , que é relativamente fácil de calcular. A desvantagem do seu uso é que ela não carrega nenhuma informação das observações e assim as partículas geradas podem ser oriundas das caudas da distribuição posterior. Desta forma, com o transcorrer do tempo o sistema degenera-se rapidamente. Em outras palavras, a variância incondicional dos pesos de importância aumenta com o tempo.

A.5. Estratégia de reamostragem e o filtro *bootstrap*

A amostragem sequencial por importância acarreta a degeneração do procedimento. A solução adotada para evitar este problema foi criar uma etapa adicional. Trata-se da etapa de seleção ou de reamostragem. Uma vez calculados os pesos por importância, antes de fazer a evolução para o instante seguinte, é realizada a seleção. As partículas de maior peso (maior importância), são selecionadas. De acordo com o seu peso é realizada uma nova amostragem (da distribuição anterior). Assim, as partículas de maior importância dão origem a um maior número de partículas. As partículas de menor importância desaparecem e não originam “descendentes”. Este procedimento é denominado filtro de partículas *bootstrap* (FPB). Ao amostrar proporcionalmente à verossimilhança, o filtro de partículas focaliza nas regiões da distribuição onde a verossimilhança é maior, ou seja, onde as boas aproximações significam possuir maior importância. As principais vantagens do FPB são: (i) facilidade de implementação; (ii) modularidade (basta atualizar a densidade por importância e o peso por importância); (iii) é paralelizável; (iv) a etapa de reamostragem independe da complexidade do modelo e não requer nenhuma alteração; (v) a sua aplicabilidade prescinde das condições de linearidade e Gaussianidade do modelo que está sendo analisado. O quadro abaixo resume as principais etapas do algoritmo do filtro de partículas padrão.

1. Inicialização

Tome um conjunto de partículas da distribuição anterior $p(x_0)$ e obtenha

$$\{(x_0^{*(i)}, \omega_0^{(i)}), \quad i = 1, \dots, N\}$$

Faça $t = 1$

2. Etapa de Avaliação

a) Calcule os novos pesos

$$\omega_{t-1}^{(i)} \propto \omega_{t-2}^{(i)} p(y_{t-1} | x_{t-1}^{*(i)}) \quad i = 1, \dots, N$$

b) Normalize os pesos

$$\tilde{\omega}_{t-1}^{(i)} = \frac{\omega_{t-1}^{(i)}}{\sum_{j=1}^N \omega_{t-1}^{(j)}} \quad i = 1, \dots, N$$

3. Etapa de reamostragem ou seleção

Reamostre $x_t^{(i)}$ com probabilidade $\tilde{\omega}_{t-1}^{(i)}$, obtendo N partículas aleatórias iid $x_t^{(j)}$, aproximadamente distribuídas conforme $p(x_t | y_{1:t})$.

Faça $\tilde{\omega}_t^{(i)} = \frac{1}{N} \quad i = 1, \dots, N$

4. Etapa de Evolução

Avance os estados no tempo de t-1 para t usando a equação de evolução dos estados $x_t^{*(i)} \sim \pi(\cdot) = p(x_t | x_{t-1}^{(i)})$ para $i = 1, \dots, N$

5. Retorne a etapa 2.

