



IBP1064_06

MODELO DE MARKOV DE CADEIA ESCONDIDA APLICADO À FILTRAGEM DIGITAL DE PERFIS Gustavo V. L. Moisés¹, João B. R. do Val²

Copyright 2006, Instituto Brasileiro de Petróleo e Gás - IBP

Este Trabalho Técnico foi preparado para apresentação na *Rio Oil & Gas Expo and Conference 2006*, realizada no período de 11 a 14 de setembro de 2006, no Rio de Janeiro. Este Trabalho Técnico foi selecionado para apresentação pelo Comitê Técnico do evento, seguindo as informações contidas na sinopse submetida pelo(s) autor(es). O conteúdo do Trabalho Técnico, como apresentado, não foi revisado pelo IBP. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões do Instituto Brasileiro de Petróleo e Gás, seus Associados e Representantes. É de conhecimento e aprovação do(s) autor(es) que este Trabalho Técnico seja publicado nos Anais da *Rio Oil & Gas Expo and Conference 2006*.

Resumo

Na indústria do petróleo é muito importante conhecer as características litológicas de uma fácies dentro de uma formação geológica. Como os custos associados à testemunhagem e a testes de produção são muito elevados, uma análise preliminar das fácies é feita através de perfis litológicos. Com base nestes perfis, um geólogo consegue identificar com razoável precisão as características litológicas das fácies, suas propriedades físicas, bem como a natureza dos fluidos, por ventura nelas contidos. Dentro de uma mesma fácies, as propriedades das rochas tendem a ser semelhantes e mudanças abruptas nestas propriedades sinalizam mudanças de fácies, o que permite a aplicação do modelo de Markov de cadeia escondida. Assumindo que as mudanças de fácies são modeladas por uma cadeia de Markov, aplicamos algoritmos baseados nos métodos de Monte Carlo para filtrar digitalmente perfis com ruído Gaussiano. Apresentamos uma aplicação com perfil de resistividade nuclear-magnética e consideramos um filtro em batch, Mínimo Erro Quadrático Médio, e o filtro de partículas de Monte Carlo. Para detectar mudanças litológicas durante a perfuração é necessário que os perfis sejam obtidos em tempo real. Com isso, poderemos otimizar os parâmetros de perfuração, modificando a velocidade de rotação da broca, a força aplicada sobre esta, ou até mesmo identificar zonas de alta pressão e temperatura, minimizando custos operacionais.

Abstract

In the petroleum industry it is very important to know the lithological characteristics of a facie in a geological formation. Since the costs associated to drill core and formation tests are very expensive, a preliminary analysis of the facies is made by well log. Based on these logs, a geologist can identify with reasonable precision the lithological characteristics of the facies, its physical properties, and the nature of fluids that can be found inside of them. In a single facie, the properties of rocks seem to be similar and abrupt changes of these properties signalize a facie change which allows the application of hidden Markov models. Assuming that the changes can be modeled by a Markov chain, one applied algorithms based on Monte Carlo methods to digitally filter well logs with Gaussian noise. One presents an application based on a nuclear-magnetic resistivity log and considers a batch filter, Minimum Mean Square Error, and Monte Carlo particle filter. To detect lithological changes during perforation it is necessary to obtain the logs in real time. With this information, the perforation parameters can be optimized through rotation and force changes, minimizing the operational costs. Besides, the identification of high pressure and temperature zones can be considered.

1. Introdução

Na indústria do petróleo é muito importante conhecer as características litológicas de uma fácies dentro de uma formação geológica. Como os custos associados à testemunhagem e a testes de produção são muito elevados, uma análise preliminar destas fácies é feita através de perfis litológicos.

Inúmeras ferramentas de perfilagem que levam em conta a salinidade, a resistividade nuclear-magnética e até a velocidade de propagação do som nas formações (perfis sísmicos) estão disponíveis para análise dos perfis. Com base

¹ Mestre, Engenheiro de Petróleo - PETROBRAS

² PHD, Engenheiro Elétrico – Departamento de Telemática/FEEC/UNICAMP

nesses perfis, um geólogo consegue identificar com razoável precisão as características litológicas das fácies, suas propriedades físicas, bem como a natureza dos fluidos, por ventura nelas contidos.

Uma característica muito importante na geologia é que dentro de uma mesma fácies, as propriedades das rochas são semelhantes. Ao se passar de uma fácies para outra, mudanças abruptas nestas propriedades são detectadas. Como estas mudanças são acentuadas propomos a aplicação do modelo de Markov de cadeia escondida, modelando as mudanças de fácies através de uma cadeia de Markov. Pretendemos a partir dos sinais dos perfis identificar as mudanças de fácies, assim como filtrar digitalmente estes sinais, eliminando parte do ruído aleatório associado, principalmente, a falhas dos sensores das ferramentas de perfilagem.

Na seção 2 apresentaremos o modelo de Markov de cadeia escondida, culminando com a formulação matemática aplicada a modelos lineares com ruído Gaussiano. Descreveremos na seção 3 as técnicas de filtragem em batch e on-line. Na seção 4, apresentaremos uma aplicação real associada ao perfil de resposta nuclear-magnética e na seção 5, as conclusões.

2. Modelo de Cadeia de Markov Escondida

Para que possamos entender o modelo de cadeia de Markov escondida dividiremos esta seção em três subseções: processo discreto de Markov, o modelo de cadeia de Markov escondida e a formulação matemática aplicada a modelos lineares com ruído Gaussiano.

2.1. Processo Discreto de Markov

O estado discreto θ é descrito num dado instante k por um elemento do conjunto $S=\{1,2,\dots,s\}$. Em espaços regulares de tempo, θ passa por uma mudança de acordo com certas probabilidades, assumindo valores dentro deste conjunto. Em geral, a forma probabilística de θ_k é dada pela especificação de todos os estados que o antecederam. No caso especial dos processos discretos de Markov, apenas o estado predecessor é necessário para descrever probabilisticamente o estado atual, isto é, $p(\theta_k | \theta_{k-1}, \theta_{k-2}, \dots) = p(\theta_k | \theta_{k-1})$, em que θ_k representa o estado de Markov no instante k . Considerando apenas os processos em que o lado direito desta equação é independente do tempo, denominados de cadeia de Markov, o conjunto de probabilidades de transição de estados é dado por $p_{ij} = p(\theta_{k+1} = j | \theta_k = i) > 0, \forall i, j \in S = \{1,2,\dots,s\}$ com $\sum_{j=1}^s p_{ii} = 1$ para cada $i \in S$. Isso significa que a probabilidade de ocorrência de um estado θ_k depende apenas do estado anterior θ_{k-1} . Por notação, temos que a matriz de transição de probabilidade é $P = [p_{ij}; i, j \in S]$. Para entender melhor este processo discreto de Markov, considere um simples modelo com três estados de Markov cujo diagrama pode ser visto na Figura 1.

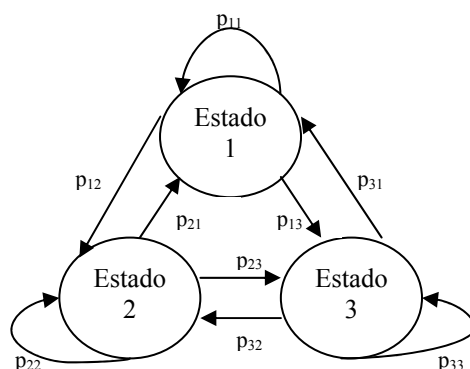


Figura 1. Diagrama de um modelo com três estados de Markov.

Para completar o modelo deste processo, temos que definir o vetor de probabilidade inicial π , em que $\pi_i = p(\theta_1 = i)$ para $i \in S$, de tal forma que $\pi_i > 0$ e $\sum_{i=1}^s \pi_i = 1$. A partir desta seção, teremos θ_k para $\forall k \in \{1,2,3,\dots\}$ como notação para o estado de Markov.

2.2. Modelo de Markov de Cadeia escondida

Nesta seção, iremos desenvolver o modelo de Markov de cadeia escondida em que toda informação sobre os estados de Markov é obtida a partir de uma função de probabilidade associada a eles. Este modelo é imerso em dois processos estocásticos, sendo o primeiro, os estados de Markov, os quais não são observados diretamente, e o segundo, uma sequência de observações obtidas de uma função probabilística do primeiro (processo subjacente). Portanto, os estados de Markov podem ser estimados a partir das observações produzidas pelo segundo processo.

Para compreender melhor o modelo de cadeia escondida vamos utilizar como exemplo, o modelo de bolas e urnas (Rabiner, 1989). Vamos assumir que existem s grandes urnas transparentes em uma sala e que dentro de cada uma delas podemos encontrar um grande número de bolas coloridas. Nós assumimos que existem m cores distintas de bolas e que a quantidade de bolas de cada cor que cada urna contém é conhecida. O processo físico para a obtenção das observações é dado por:

- 1) A primeira urna é sorteada de acordo com um dado processo aleatório;
- 2) Da urna escolhida, uma bola é sorteada aleatoriamente e sua cor é registrada como observação. Somente a cor é anotada, não se conhecendo qual foi a urna escolhida;
- 3) A bola sorteada é devolvida para a mesma urna da qual foi retirada;
- 4) Uma nova urna é selecionada de acordo com o processo de transição aleatório associado à urna atual, sem que se saiba qual foi a urna escolhida.

A partir deste ponto, os procedimentos de 2 até 4 devem ser repetidos até o número de observações desejadas. O processo gera uma sequência finita de cores, que nós podemos modelar como o vetor de observação do modelo de Markov de cadeia escondida. Fica claro no modelo de bolas e urnas que os estados de Markov são definidos pela sequência de urnas e a função probabilística, pela probabilidade de ocorrência das cores que depende da urna (estado de Markov) escolhida. Vale ressaltar que a escolha da urna é feita através da matriz de transição do modelo e que esta escolha depende unicamente da urna escolhida anteriormente.

Para entender como aplicar o modelo de cadeia de Markov escondida para identificar mudanças nas fácies litológicas vamos utilizar o modelo de bolas e urnas como paralelo. Consideremos que as fácies representam estados de Markov e que as probabilidades de mudança de uma fácies para outra são conhecidas e determinadas pela formação geológica em que estas se encontram. Neste caso, as fácies podem ser comparadas as urnas e o vetor de observações, sequência de cores, aos sinais dos perfis.

Se considerarmos a representação por espaço de estados a tempo discreto, temos que o modelo matemático genérico que representa o modelo de fácies é dado por duas equações, a que representa a evolução das propriedades ao longo de uma mesma fácies, Equação 1, e a que representa a forma com que a ferramenta de perfilagem capta estas propriedades através dos sinais, Equação 2, com

$$x_{k+1} = f_{\theta_{k+1}}(x_k, v_{k+1}, u_{k+1}) \quad (1)$$

$$y_k = g_{\theta_k}(x_k, z_k, u_k), \quad (2)$$

em que θ_k representa a fácies em que se encontra a ferramenta, x_k as propriedades das rochas, y_k o vetor de observações ou sinais dos perfis, u_k uma entrada determinística conhecida, v_k e z_k ruídos aleatórios no instante k , que pode representar tempo ou posição no espaço.

2.3. Formulação do Problema

Simplificamos o modelo genérico através da utilização de sistemas lineares com ruído Gaussiano. Desta forma, o sistema linear sujeito a saltos Markovianos, a tempo discreto $\mathcal{G}$ pode ser descrito por

$$x_{k+1} = A_{\theta_{k+1}} x_k + B_{\theta_{k+1}} v_{k+1} + F_{\theta_{k+1}} u_{k+1} \quad (3)$$

$$y_k = C_{\theta_k} x_k + D_{\theta_k} z_k + G_{\theta_k} u_k, \quad (4)$$

sendo que $x_k \in R^{n_x}$ é a parcela contínua do vetor de estados ou estado contínuo do sistema, $y_k \in R^{n_y}$ é o vetor de observação, $u_k \in R^{n_u}$ é uma entrada conhecida determinística, $v_k \in R^{n_v}$ e $z_k \in R^{n_z}$ são sequências de ruído branco gaussiano com média nula cujas variâncias são as identidades I_{n_v} e I_{n_z} .

A, B, C, D, F e G representam coleções de matrizes que são funções do estado da cadeia de Markov θ_k e evoluem de acordo com a realização da cadeia de Markov finita. Se $\theta_k = i$, temos $A_{\theta_k=i} = A_i, B_{\theta_k=i} = B_i, C_{\theta_k=i} = C_i, D_{\theta_k=i} = D_i, F_{\theta_k=i} = F_i$ e $G_{\theta_k=i} = G_i$.

Assumimos que x_0 é uma distribuição gaussiana com média $\bar{x}_0$ e variância P_0 e consideramos que x_0, v_k e z_k são mutuamente independentes para todo k. Por conveniência de notação, temos $y_{1:k} \equiv \{y_1, y_2, \dots, y_k\}$, $x_{0:k} \equiv \{x_0, x_1, \dots, x_k\}$ e $\theta_{1:k} \equiv \{\theta_1, \theta_2, \dots, \theta_k\}$, representando o conjunto de vetores de observação, os estados contínuos e os estados de Markov, respectivamente.

3. Filtragem

As técnicas de filtragem têm como objetivo estimar o vetor de estados de um sistema dinâmico a partir do vetor de observação. Nós adotamos a abordagem via representação no espaço de estados para modelar os sistemas dinâmicos a tempo discreto, devido a sua conveniência e praticidade. Esta abordagem é baseada no vetor de estados do sistema, que contém toda a informação relevante requerida para descrever o sistema sobre investigação. Dividiremos este vetor em duas parcelas, a contínua, representada por x_k e a discreta, por θ_k . Para maiores detalhes sobre algoritmos de filtragem de sistemas lineares baseados em métodos de Monte Carlo para cadeia de Markov vide Moisés (2005).

Nós dividimos as técnicas de filtragem em dois grandes sub-grupos dados por algoritmos em *batch* e por filtros recursivos. O primeiro grupo calcula as estimativas dos estados condicionadas a um conjunto fixo de observações e o segundo, por sua vez, considera os filtros recursivos que recuperam os estados através da sequência de observações obtida até o instante em que se deseja calcular a estimativa. O estudo e análise destes grupos podem ser vistos nas sub-seções 3.1 e 3.2.

3.1. Filtragem em *batch*

A palavra *batch* em inglês significa batelada, conjunto de bens produzidos num dado intervalo de tempo com uma determinada especificação e configuração da máquina ou processo que o produziu. Com base nesta definição temos que filtragem em *batch* representa as técnicas de filtragem que geram estimativas dos estados não observáveis a partir de um conjunto de observações $y_{1:T} \equiv \{y_1, y_2, \dots, y_T\}$. Estes algoritmos tendem a considerar que se temos acesso a duas observações, podemos obter estimativas mais precisas, se compararmos ao caso em que temos acesso apenas a uma observação.

Quando não se tem acesso aos estados, o único dado consistente é a sequência de observações $y_{1:T}$ e as estimativas dos estados de Markov $\theta_{1:T}$ são obtidas, em princípio, a partir da distribuição marginal a *posteriori* $p(\theta_{1:T} | y_{1:T})$. Vamos entender melhor a dificuldade de encontrarmos as estimativas do estado de Markov θ_k tendo acesso apenas às observações imersas em ruído y_k . Inicialmente, temos acesso à função de distribuição de probabilidade do estado contínuo no instante $k=0$, que é representada por uma gaussiana cuja média é $\bar{x}_0$ e a variância P_0 . Como não temos acesso ao estado de Markov no tempo $k=1$, temos s possíveis sistemas lineares, representando todos os estados possíveis da cadeia de Markov. Para sistemas lineares imersos em ruído branco e gaussiano com vetor de estados de entrada também gaussiano, temos que a observação de saída também é representada por uma função de distribuição de probabilidade gaussiana.

Com isso, ao final do instante $k=1$ teremos s gaussianas, representando a observação de saída, uma para cada estado de Markov. Como não temos acesso ao verdadeiro estado de Markov, teremos s^k distribuições gaussianas no instante k . Infelizmente, podemos perceber que o custo computacional do cálculo exato destas estimativas cresce exponencialmente com o número de observações, o que acarreta um custo computacional proibitivo (Doucet et al., 2001). Por este motivo, um algoritmo que aproxima a distribuição marginal a *posteriori* $p(\theta_{1:T} | y_{1:T})$ e que possui custo computacional linearmente crescente com o número de observações será levado em consideração.

Nesta seção, apresentaremos o algoritmo do Mínimo Erro Quadrático Médio que utiliza os métodos de Monte Carlo proposto por Doucet e Andrieu (2001) para amostrar esta distribuição marginal, tendo como ferramentas o filtro de Kalman (Anderson e Moore, 1979) na recuperação do estado contínuo x_k e a dinâmica conhecida da cadeia Markov representada pela probabilidade inicial π e pela matriz de transição P .

O algoritmo para obter as estimativas do Mínimo Erro Quadrático Médio (MMSE) dos estados $\theta_{1:T}$ pode ser descrito por:

- Sorteie aleatoriamente $\theta_{1:T}^{(0)} = \{\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_T^{(0)}\}$;
- Na l -ésima iteração, $l \geq 1$, amostre $\theta_k^{(l)}$ para $k = 1, \dots, T$ segundo $\theta_k^{(l)} \prec p(\theta_k | y_{1:T}, \theta_{-k}^{(l)})$, em que $\theta_{-k}^{(l)} \equiv \{\theta_1^{(l)}, \dots, \theta_{k-1}^{(l)}, \theta_{k+1}^{(l-1)}, \dots, \theta_T^{(l-1)}\}$.

Como para cada iteração temos um conjunto de estados de Markov podemos calcular através do filtro de Kalman as estimativas do conjunto de estado contínuo por iteração $x_{1:T}^{(l)} \equiv \{x_1^{(l)}, \dots, x_T^{(l)}\}$. Após N iterações, a estimativa MMSE dos estados contínuos $x_{1:k}$ pode ser calculada por

$$\bar{x}_{1:T} \equiv \frac{1}{N} \sum_{l=1}^{N-1} x_{1:T}^{(l)}. \quad (5)$$

Uma vantagem deste método é o desenvolvimento da decomposição da distribuição $p(\theta_k | y_{1:T}, \theta_{-k}^{(l)})$, que leva em consideração os estados de Markov $\theta_1^{(l)}, \dots, \theta_{k-1}^{(l)}$ da iteração l e os estados de Markov $\theta_{k+1}^{(l-1)}, \dots, \theta_T^{(l-1)}$ da iteração $l-1$. A seguinte decomposição de $p(y_{1:T} | \theta_{1:T})$ permite o cálculo eficiente da distribuição $p(\theta_k | y_{1:T}, \theta_{-k})$ para $k = 1, \dots, T$

$$p(y_{1:T} | \theta_{1:T}) = p(y_{1:k-1} | \theta_{1:k-1}) p(y_k | y_{1:k-1}, \theta_{1:k}) \int p(y_{k+1:T} | \theta_{k+1:T}, x_k) p(x_k | y_{1:k}, \theta_{1:k}) dx \quad (6)$$

Os dois primeiros termos do lado direito da equação 6 podem ser calculados através da utilização da recursão em avanço baseada no filtro de Kalman (Anderson e Moore, 1979). Já o terceiro termo pode ser calculado, utilizando uma recursão em retrocesso desenvolvida por Doucet e Andrieu (2001).

3.2. Filtro Recursivo

Em muitos problemas, uma estimativa dos estados do sistema é necessária a cada instante de tempo em que uma nova observação é incorporada, sendo conveniente o uso de um filtro recursivo. Nesta abordagem, nem todo o conjunto de informações precisa ser armazenado e, além do mais, não é necessário reprocessar a informação existente quando uma nova observação estiver disponível. No filtro recursivo toda a informação recebida é processada sequencialmente.

O filtro recursivo consiste essencialmente de dois estágios: predição e atualização. O passo de predição utiliza o modelo do sistema para prever a função de distribuição de probabilidade dos estados a frente entre uma medição e outra. Como os estados contínuos estão sujeitos a perturbações desconhecidas, modeladas como ruído aleatório, a predição geralmente translada, alarga ou deforma a verdadeira função de distribuição de probabilidade. A operação de atualização utiliza-se da última medição para modificar e corrigir a função distribuição de probabilidade. Isso pode ser obtido de maneira precisa através do teorema de Bayes, que é o mecanismo de atualização do conhecimento sobre os estados baseado na nova informação adquirida.

Atualmente, um filtro recursivo muito em voga é o filtro de partículas utilizado tanto em abordagens lineares quanto não lineares. O filtro de partículas de Monte Carlo desenvolvido por Doucet et al. (2001) será apresentado para que fique claro o seu funcionamento e simplicidade.

Filtro de partículas de Monte Carlo é um algoritmo recursivo de filtragem baseado em simulações de Monte Carlo para calcular a distribuição condicional estimada dos estados x_k e θ_k dado o histórico de observações $y_{1:k}$. Este filtro de partículas pode ser visto aproximadamente como uma grade aleatória adaptativa, onde as partículas (valores da grade) evoluem aleatoriamente no tempo de acordo com uma regra de simulação. Cada partícula possui um peso associado à verossimilhança que é atualizado de acordo com a regra de Bayes a cada observação.

O filtro de partículas combina com propriedade a amostragem sequencial por importância e os métodos de Monte Carlo para cadeia de Markov (MCMC) associados a um esquema seletivo que replica/descarta partículas a cada instante de tempo. Além de sua simplicidade, o que torna este filtro ainda mais atrativo é que a taxa com que o erro das estimativas converge para zero independente da dimensão do estado de Markov.

Supondo que no instante $k-1$, N_p amostras aleatórias $\theta_{1:k-1}^{(l)}, l=1, \dots, N_p$ distribuídas aproximadamente de acordo com $p(\theta_{1:k-1} | y_{1:k-1})$, o filtro de partículas de Monte Carlo evolui no instante k conforme o procedimento descrito abaixo.

a. Passo da amostragem sequencial por importância

- i. Para $l=1, \dots, N_p$, amostre $\tilde{\theta}_k^{(l)} \sim Y(\theta_k | y_{1:k}, \theta_{1:k-1}^{(l)})$ e faça $\tilde{\theta}_{1:k}^{(l)} \equiv \{\tilde{\theta}_{1:k-1}^{(l)}, \tilde{\theta}_k^{(l)}\}$;
- ii. Para $l=1, \dots, N_p$, calcule os pesos incrementais sobre uma constante de normalização

$$w_k^{(l)} \propto \frac{p(y_k | y_{1:k}, \tilde{\theta}_{1:k}^{(l)}) p(\tilde{\theta}_k^{(l)} | \tilde{\theta}_{k-1}^{(l)})}{Y(\tilde{\theta}_k^{(l)} | y_{1:k}, \tilde{\theta}_{1:k-1}^{(l)})} \quad (7)$$

- iii. Para $l=1, \dots, N_p$, normalize os pesos incrementais

$$\tilde{w}_k^{(l)} = \frac{\tilde{w}_k^{(l)}}{\sum_{j=1}^{N_p} \tilde{w}_k^{(j)}} \quad (8)$$

b. Passo de Seleção:

- i. Obter N_p partículas filhas $(\theta_{1:k}^{(l)}, l=1, \dots, N_p)$ a partir da replicação/descarte das partículas $(\tilde{\theta}_{1:k}^{(l)}, l=1, \dots, N_p)$, considerando os pesos incrementais normalizados $\tilde{w}_k^{(l)}$;

c. Cálculo das estimativas

$$E[\theta_k = i | y_{1:k}] \approx \frac{\sum_{l=1}^{N_p} 1_{\theta_k^{(l)}=i}}{N_p} \quad (9)$$

$$E[x_k | y_{1:k}] \approx \frac{\sum_{l=1}^{N_p} m_{k|k}^{(l)}}{N_p}, \quad (10)$$

sendo N_p e $m_{k|k}^{(l)}$ o número total de partículas e a estimativa do estado contínuo obtido do filtro de Kalman (Anderson e Moore, 1979), no instante k da partícula l , respectivamente.

4. Resistividade Nuclear-Magnética

Um exemplo ilustrativo de aplicação de sistemas lineares sujeitos a saltos Markovianos à identificação de fácies litológicas é a utilização do perfil de resposta nuclear-magnética (Salomão, 2006; Ruanaidh e Fitzgerald, 1996) cujos ruídos podem ser modelados por distribuições Gaussianas. Como temos acesso apenas às observações ruidosas provenientes do perfil, podemos recuperar estimativas dos estados contínuos (resistividade nuclear-magnética), assim como detectar mudanças de fácies.

A resistividade nuclear-magnética de uma formação A , conjunto de fácies litológicas, pode ser representada por uma distribuição normal com $\text{rad}^{(A)}$ e $\text{var}^{(A)}$, representando a média e a variância da resistividade da formação. Consideramos que a resistividade nuclear-magnética dentro de uma mesma fácies é constante e que a resistividade de uma fácies não influencia a das fácies vizinhas, variáveis independentes.

Para completar o modelo vamos considerar ora um ruído branco gaussiano de observação com variância $\text{var}^{(\text{obs})}$ e ora uma observação errônea (*outlier*) dada pelo mau funcionamento da ferramenta de perfilagem. Podemos

caracterizar o *outlier* por uma gaussiana com média e variância dadas por $\text{rad}^{(\text{out})}$ e $\text{var}^{(\text{out})}$, respectivamente. Quatro estados de Markov foram utilizados no modelo:

- Estado 1 - Evolução da resistividade dentro de uma mesma fácies com observação ruidosa;
- Estado 2 - Evolução da resistividade dentro de uma mesma fácies com *outlier*;
- Estado 3 - Mudança de fácies com observação ruidosa;
- Estado 4 - Mudança de fácies com *outlier*.

Embora não haja ação de controle, nós definimos u_k para $\forall k$ dada por $u_k = [\text{rad}^{(A)}; \text{rad}^{(\text{out})}]$ como artifício matemático para inserir as médias não nulas do *outlier* $\text{rad}^{(\text{out})}$ e da resistividade nuclear-magnética da formação $\text{rad}^{(A)}$. As matrizes do sistema G são dadas por

$$\begin{aligned} A_1 = A_2 = 1; A_3 = A_4 = 0; B_1 = B_2 = 0; B_3 = B_4 = \text{var}^{(A)}; C_1 = C_3 = 1; C_2 = C_4 = 0; D_1 = D_3 = \text{var}^{(\text{obs})}; \\ D_2 = D_4 = \text{var}^{(\text{out})}; F_1 = F_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; F_3 = F_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; G_1 = G_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; G_2 = G_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \end{aligned} \quad (11)$$

As médias, variâncias e a matriz de transição do modelo dependem da formação e da ferramenta de perfilagem que será utilizada. Consideramos um perfil real (Ruanaidh e Fitzgerald, 1996) originário da *Schlumberger* que foi coletado durante uma operação de perfuração para óleo *Log-While-Drilling* cujos parâmetros identificados por Salomão (2006) são dados por $\text{rad}^{(A)} = 115.000$, $\text{var}^{(A)} = 10.000$, $\text{var}^{(\text{obs})} = 2.500$, $\text{rad}^{(\text{out})} = 85.000$ e $\text{var}^{(\text{out})} = 12.500$.

Ao invés de uma cadeia de Markov com quatro estados, Salomão (2006) utiliza duas cadeias de Markov com dois estados cada uma. Por isso, tivemos que transformar as duas matrizes de transições em uma única, considerando que a cadeia do modelo do sistema é independente da de medição. A matriz de transição para a cadeia de Markov com quatro estados, considerando as transições identificadas em Salomão (2006) é dada por

$$P = \begin{bmatrix} 0,992813 & 0,003487 & 0,003687 & 0,000013 \\ 0,199260 & 0,797040 & 0,000740 & 0,002960 \\ 0,992813 & 0,003487 & 0,003687 & 0,000013 \\ 0,199260 & 0,797040 & 0,000740 & 0,002960 \end{bmatrix}. \quad (12)$$

O perfil real possui $T=4050$ com T representando o número total de amostras. Para finalizar o modelo, consideramos que o primeiro estado de Markov seja sempre igual ao estado 1, isto é, $\pi = [1 \ 0 \ 0 \ 0]$, $x_0 = \text{rad}^{(A)} = 115.000$ e $P_0 = \text{var}^{(A)} = 10.000$. Lembramos que x_k representa a resistividade nuclear-magnética na amostra k .

Com a velocidade constante e a posição inicial da ferramenta podemos transformar as amostras de tempo em posições no espaço. O gráfico contendo os perfis apresenta as amostras com sinal negativo para representar a passagem da ferramenta de cima para baixo.

Em cada um dos algoritmos estudados alguns parâmetros precisam ser definidos. No Mínimo Erro Quadrático Médio (MMSE), utilizamos $N=200$ e no filtro de partículas (FP) consideramos $N_p=500$ com N_p representando o número total de partículas.

Para verificar o funcionamento dos algoritmos, plotamos sobre o mesmo gráfico a observação y_k e o estado contínuo estimado $\bar{x}_k$ obtido de cada um dos algoritmos. Neste caso, com a observação y_k representada pelos pontos vermelhos e estado contínuo estimado $\bar{x}_k$ pelo traço contínuo em azul, conseguimos através de inspeção visual evidenciar algumas diferenças entre os algoritmos.

Na figura 2, podemos comparar o funcionamento do filtro de partículas (FP) com o algoritmo do mínimo erro quadrático (MMSE). Tanto o FP quanto o MMSE conseguem detectar com maior sensibilidade os saltos, evidenciando as sub-fácies. Neste caso, verificamos que o filtro de partículas tem uma maior dificuldade em relação ao algoritmo em *batch* de eliminar os *outliers*. Isso se deve a característica *on line* do filtro de partículas que leva em consideração apenas as observações $y_{1:k}$ obtidas até um dado instante k , ao contrário dos algoritmos em *batch* que utilizam uma janela de observações $y_{1:T}$.

O MMSE só pode ser aplicado quando temos acesso ao conjunto total de pontos, sendo indicado para a análise geológica. Este algoritmo permite uma identificação mais precisa das mudanças de fácies e aprimora a qualidade dos sinais dos perfis através da filtragem digital, o que elimina parte do ruído aleatório associado, principalmente, a falhas dos sensores das ferramentas de perfilagem.

Para detectar mudanças litológicas durante a perfuração é necessário que os perfis sejam obtidos em tempo real. Com isso, precisamos de um filtro recursivo, indicando o uso do filtro de partículas de Monte Carlo. Poderemos

durante a perfuração detectar mudança de fácies. Quando esta mudança for detectada, os parâmetros de perfuração poderão ser otimizados através da modificação da velocidade de rotação da broca, da força sobre esta e do peso do fluido de perfuração. Isso minimizará os custos operacionais, pois poderemos aumentar o tempo de uso da broca e minimizar o tempo de intervenções com a diminuição do tempo de perfuração e o número de manobras.

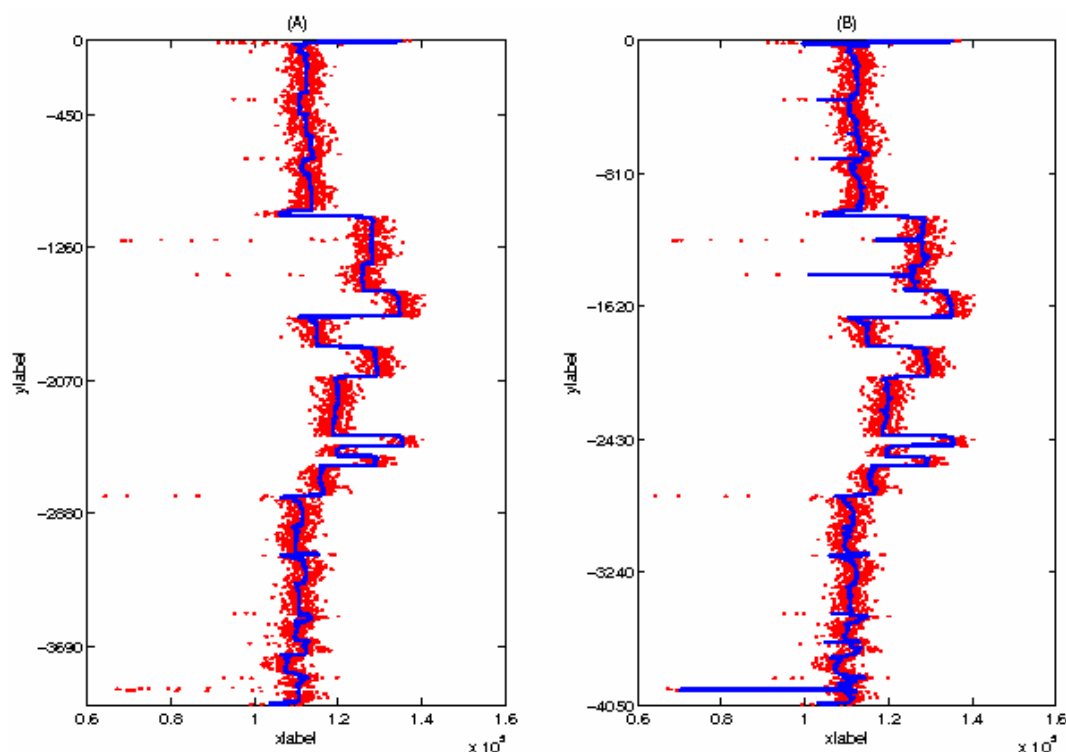


Figura 2. Perfil associado ao (A) Mínimo Erro médio Quadrático; (B) Filtro de partículas de Monte Carlo.

5. Conclusão

O modelo de Markov de cadeia escondida associado aos métodos de Monte Carlo é uma ferramenta eficaz na detecção de mudanças de fácies litológicas dentro de uma formação geológica e deve ser utilizado para otimizar custos operacionais, assim como aprimorar a qualidade das informações adquiridas a partir dos sinais das ferramentas de perfilagem. Além disso, pode auxiliar a detecção de zonas de alta pressão e temperatura a partir dos dados de pressão e temperatura e dos parâmetros de perfuração.

6. Agradecimentos

Esta pesquisa teve apoio da FAPESP, processos 03/03781-1 e 03/06736-7, do CNPq, processo 302345/04-3 e da PETROBRAS S.A..

7. Referências

- ANDERSON, B. D. O., MOORE, J. *Optimal Filtering*, Prentice-Hall, Englewood Cliffs-NJ, 1979.
- DOUCET, A., ANDRIEU, C. Iterative algorithms for state estimation of jump Markov linear systems, *IEEE Transactions on Signal Processing*, v. 49, n. 6, p. 1216–1227, 2001.
- DOUCET, A., GORDON, N. J., KRISHNAMURTHY, V. Particle filters for state estimation of jump Markov linear systems, *IEEE Transactions on Signal Processing*, v. 49, n. 3, p. 613–624, 2001.
- MOISES, G. V. L. *Filtragem via Métodos de Monte Carlo para Processos Lineares com Saltos Markovianos*, Tese de Mestrado, UNICAMP, 2005.

- RABINER, L. R. A tutorial on hidden Markov models and selected application in speech recognition, *Proceedings of the IEEE*, n. 2, p.257–286, 1989.
- RUANAIDH, J. J. K. , FITZGERALD, W. J. *Numerical Bayesian Methods Applied to Signal Processing*, Springer-Verlag, New York-NY, 1996.
- SALOMÃO, R. Filtro de Partículas para Análise de Sinal Geofísico Obtido Durante Perfuração de Poços de Petróleo, Tese de Mestrado, UNICAMP, 2006.