



Métodos Baseados em Núcleos e Máquinas de Suporte Vetorial em Aplicações de Geofísica de Petróleo

Geovan Tavares, Hélio Lopes, Alex Laier, Renner Castro - PUC-Rio e Rogério Santos, Amin Murad - Petrobras

Copyright 2007, SBGf - Sociedade Brasileira de Geofísica

This paper was prepared for presentation at the 10th International Congress of The Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 19-22 November 2007.

Contents of this paper were reviewed by the Technical Committee of the 10th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction, or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

Kernel based methods and support vector machines are recognized today as important methods in areas such as artificial intelligence and bioinformatics, among others. Increasingly in the last years applications to geophysics is becoming widespread, emphasis made to the oil industry. In this paper we propose a new method to distinguish between reservoir and non reservoir characteristics in a seismic sample. Volumetric multiattribute seismic and well data are used in the training process.

Introdução

Sísmica de petróleo tem tido uma dos mais importantes desenvolvimentos tecnológicos na indústria mundial, principalmente aquele relacionado à análise de multiatributos desejavelmente correlacionados com multipropriedades petrofísicas observadas na escala de perfis de poços através de perfis elétricos. Tal correlação esbarra em limitações naturais oriundas das diferenças de escala de observação de leitura de amostras diretas (poços) e remotas (sísmica). Atributos sísmicos têm se mostrado de grande valor na estimativa de propriedades geológicas e petrofísicas de reservatórios de hidrocarbonetos, acompanhando, e por vezes liderando, o crescente desenvolvimento de algoritmos matemáticos com o objetivo de se derivar correlações que sugiram, com redução de incertezas, elos entre aquilo que se observa com a propagação da onda sísmica (10m a 50m) com a informação de perfis elétricos (15 a 20cm). Necessidades econômicas de diversas áreas em reservatórios de hidrocarbonetos apontam para uma demanda de crescente precisão na correlação entre tais diferentes medidas, com seu uso na obtenção de melhores estimativas em poços de porosidades, saturações e espessuras, entre outras, que possam ser extrapoladas pelo dado sísmico. Métodos atuais dessa correlação poço-sísmica possuem por vezes algumas fragilidades de predição, pois nem sempre se consegue derivar um único atributo sísmico que seja, direta e unicamente, correlacionável à uma determinada propriedade petrofísica. Adicionalmente, quando se consegue essa boa aproximação, nem sempre ela é linear, impondo equacionamentos matemáticos por vezes bastante complexos para uma determinada regra de correlação. Dessa forma, torna-se valiosa a existência de uma ferramenta estatística supervisionada que use o máximo número de atributos sísmicos correlacionáveis a diferentes propriedades petrofísicas de poços, em estimativas simultâneas, globais e decisivas, envolvendo

as duas escalas de observação. Para tal decisão com multiatributos 3D, usa-se, invariavelmente, uma grande quantidade de informações e, então, um passo natural é aplicação de inteligência artificial, através de métodos matemáticos de aprendizagem, ou treinamento. Métodos baseados em Núcleos (Kernel Based Methods) e as Máquinas de Suporte Vetorial (Support Vector Machines, ou simplesmente SVM's) fazem parte de mais moderna classe de algoritmos de aprendizagem estatística, introduzida por Vapkin (1988). Tais métodos vêm sendo utilizados com sucesso em diversas aplicações em geofísica, tanto em classificação binária quanto em classificação multi-classes, Kuzman (2003), Li (2005), Li et alli (2005), Zhao et alli (2005). A grande vantagem dos métodos baseados em núcleo e em máquinas de suporte vetorial é que eles são métodos não-lineares de classificação que utilizam algoritmos computacionais eficientes de otimização não-linear.

Muitos autores têm se dedicado a métodos com redes neurais nas mais diversas áreas de aplicação (geofísica, biologia, economia, finanças, reconhecimento de padrões em imagens, etc.). Tanto redes neurais quanto as SVM têm o mesmo objetivo que é o de achar relações entre os dados que minimizam o erro da classificação. Redes neurais tentam minimizar o risco empírico, isto é, reduzir a probabilidade de erro na classificação dentro dos dados de treinamento. Por sua vez, as SVMs foram criadas para minimizar o erro estrutural, de tal forma que a probabilidade de erro de classificação dos dados, que não são de treinamento, seja minimizada, com isso, elas aproximam melhor a função de classificação ideal.

Enquanto em redes neurais, o treinamento é feito de uma forma iterativa, onde cada passo tenta obter continuamente melhores resultados no ajuste da função de classificação, na maioria dos casos, torna-se difícil determinar quando finalizar o processo iterativo, ou especificar quando se chega a uma boa aproximação. As SVMs tratam os dados global e simultaneamente, proporcionando uma solução ótima quando o treinamento finaliza. Ainda comparativamente, nos treinamentos de redes neurais há uma auto-modificação, alterando os pesos nas arestas conectadas. Durante esse processo de atualização, a solução ótima pode não chegar ao ponto ótimo global, pois o método pode parar num ponto crítico local. O principal motivo é que os dados de treinamento são alimentados na rede um após o outro e um subconjunto dos dados, podendo gerar uma grande influência nos pesos das arestas, desprezando a contribuição dos demais subconjuntos. Por outro lado, as máquinas de suporte vetorial possuem, necessariamente, um ótimo global para um processo de classificação. O problema de treinamento se reduz, então, a um problema de programação quadrática, qual seja minimizar um funcional quadrático sujeito a restrições lineares com desigualdade e igualdades.

Quando se deseja correlacionar multiatributos sísmicos volumétricos com propriedades petrofísicas oriundas de perfis de poços, deve-se buscar uma classificação sísmica associada, por exemplo a rochas não reservatório, reservatórios de areia fina, de areia media, de areia grossa, carbonatos, etc.. O método de aprendizagem estatística supervisionada aqui apresentado tem obtido elevados índices de acerto. Para descrever, ilustrativamente, a idéia básica desses métodos, vamos considerar um problema de classificação binária, discriminando-se dois grupos das litologias citadas: reservatórios e não-reservatórios.

Método

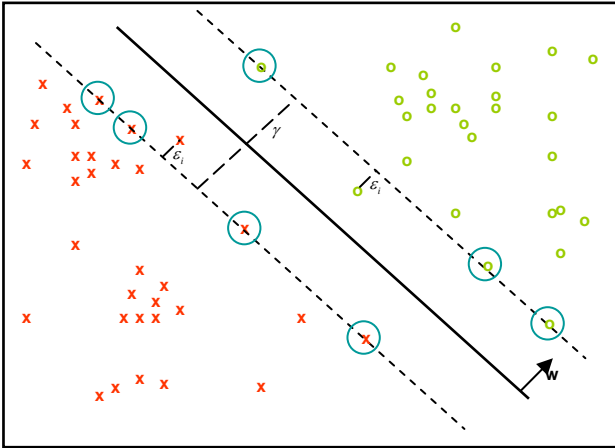


Figura 1. Dados e parâmetros a serem otimizados para construção de um hiperplano separador em dimensão alta.

O que procuramos, indicado pela figura 1, é maximizar a largura da faixa γ na direção w , centrado em b , e com penalidade $C \cdot \varepsilon_i$ para cada dado x_i dentro da faixa. Os vetores dentro dos círculos são os vetores de suporte.

O problema de otimização pode ser descrito como:

$$\text{MAX}_{w, \gamma} \quad \gamma - C \sum_{i=1}^l \varepsilon_i$$

$$\text{Sujeito à } y_i \langle w, \Phi(x_i) \rangle \geq \gamma - \varepsilon_i, \quad \varepsilon_i \geq 0 \quad \text{e} \quad \|w\|^2 = 1$$

onde $\Phi(x_i)$ é escolhido para mergulhar os dados $\{x_i\}$ em dimensão alta com o objetivo de separar os dados por hiperplanos determinados por vetores de suporte w .

Este problema se resolve no dual, isto quer dizer que é função somente dos vetores de suporte, com a seguinte formulação:

$$\text{MAX}_{\alpha} \quad W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$\text{Sujeito à } \sum_{i=1}^l \alpha_i y_i = 0 \quad \text{e} \quad 0 \leq \alpha_i \leq C \quad i=1, \dots, l$$

Tem-se:

$$\alpha_i \left[y_i \langle w^*, \Phi(x_i) \rangle - (\gamma - \varepsilon_i) \right] = 0 \quad \rightarrow \quad \alpha_i \neq 0 \quad \forall i \in SV$$

Esse método de classificação é usado para construir um hiperplano separador, em um espaço vetorial de dimensão infinita, que é chamado de espaço característico F , sendo um espaço de Hilbert (ou um espaço vetorial completo, separável munido de um produto interno). Assim a função que classificaria um novo ponto x , como pertencendo a classe $+1$ ou a -1 , teria a seguinte fórmula:

$$f(x) = \text{sgn}(\langle w, \Phi(x) \rangle + b),$$

onde $\langle \cdot, \cdot \rangle$ denota o produto interno, w é um vetor em F , b é uma constante real e $\varphi : X \rightarrow F$ é uma transformação não-linear que mapeia os exemplos x_i para o espaço característico F , ou seja leva $x \in \mathbb{R}^n$ em $\varphi(x) \in F$.

Podemos substituir o vetor w por $\sum_{i=1..L} \alpha_i \Phi(x_i)$, obtendo a seguinte expressão para a função classificadora:

$$f(x) = \text{sgn}(\langle \sum_{i=1..L} \alpha_i \Phi(x_i), \Phi(x) \rangle + b) =$$

$$\text{sgn}(\langle \sum_{i=1..L} \alpha_i \langle \Phi(x_i), \Phi(x) \rangle + b)$$

Nesse ponto, torna-se conveniente definir o que seja um núcleo.

Um núcleo $K: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função simétrica semi-definida positiva que pode ser escrita na forma:

$$K(x, z) = \langle \Phi(x), \Phi(z) \rangle$$

Existem várias outras funções núcleo que são muito importantes em geofísica, tais como o núcleo definido a partir de análise de componentes principais, ou aqueles núcleos construídos através das transformações de Fourier ou Wavelets. Além disso, pode-se compor e operar núcleos de modo que eles continuem sendo núcleos.

Assim, a função de classificação f , definida anteriormente, pode ser re-escrita através do uso de uma função núcleo:

$$f(x) = \text{sgn}(\langle \sum_{i=1..L} \alpha_i K(x_i, x) \rangle + b)$$

Dessa forma, o objetivo fundamental dos métodos baseado em núcleo é o de determinar qual seria o valor do vetor $\alpha = (\alpha_1, \dots, \alpha_L)$ e da variável b que melhor separe esses pontos para uma determinada escolha da função núcleo k . As máquinas de suporte vetorial são algoritmos que permitem a escolha dessas variáveis através de um processo de programação quadrática.

Considerando um conjunto de treinamento $\{(x_i, y_i), i = 1, \dots, L\}$ cujo i -ésimo exemplo $x_i \in X \subset \mathbb{R}^n$ possui o rótulo $y_i \in \{+1, -1\}$. No caso descrito, o vetor x_i significa um intervalo de uma, ou várias, propriedades do poço. A notação $y_i = +1$ representa uma rocha que, nesse intervalo, tem característica de reservatório e $y_i = -1$ seria uma rocha não-reservatório.

Na prática, é conveniente se oferecer boas opções para construir a função de núcleo. Para se atingir esse objetivo, uma boa interface gráfica para o usuário é fundamental para a definição da função núcleo, com a possibilidade de uso da transformação de Fourier, de

uma transformação Wavelets, ou de uma transformação S. Através dessas transformações pode-se facilmente, se realizar uma mudança de escala, tão necessária quando se deseja correlacionar informações de poços e sísmica. Adicionalmente, pode-se separar os dados através de uma análise de frequências do sinal de uma maneira diferente daquela proposta Zhao et alli (2005). Adequada também é a combinação das transformadas citadas como exemplo, com análise de componentes principais (PCA), melhorando-se ainda mais a precisão da função classificadora.

Existem diversos algoritmos para desenvolver um classificador multi-classes usando SVM. Eles montam, basicamente, uma árvore binária de decisão, utilizando como critério os classificadores binários citados anteriormente. Tais classificadores podem seguir basicamente duas estratégias: uma classe contra todas as demais, e uma classe contra cada uma separadamente. Com esses classificadores construídos, a árvore binária de decisão pode então ser montada e o classificador pode responder a qual classe um determinado dado de entrada se enquadraria melhor.

Arquitetura do sistema

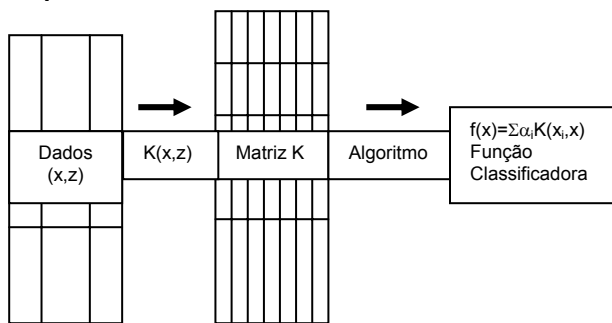


Figura 2. Máquina de Suporte Vetorial

Os métodos baseados em núcleos são bastante modulares, pois possuem duas partes que são independentes: a primeira parte consiste num conjunto de algoritmos que permite a construção das funções núcleo. A segunda parte se refere aos algoritmos de classificação que, através de um certo processo de otimização, define o vetor $\alpha = (\alpha_1, \dots, \alpha_L)$ e a variável b que melhor separam os pontos de treinamento.

Resultados

Construiu-se um modelo geológico com duas faces (reservatório e não reservatório) utilizando o método aqui apresentado. Nesse modelo o acerto foi de 99% dos poços. Para fazer o aprendizado, utilizamos os poços em profundidade e atribuímos um rótulo y_i para cada amostra sísmica i de um cubo de atributos interceptado por cada poço. Nesse caso, conseguiu-se rotular 550 amostras. Para cada amostra rotulada, tomou-se uma janela centrada em 9 amostras do traço sísmico. Cada amostra rotulada por i passa então a possuir 21 atributos, a saber: sua coordenada geográfica (x, y, z) ; a amplitude de cada elemento numa janela de 9 amostras do traço, nela centrado, e a parte real da transformada de Fourier dessa

janela de 9 amostras de amplitudes. Aplicou-se, nesse modelo, um núcleo gaussiano e um discriminante linear no espaço característico como o algoritmo classificador.

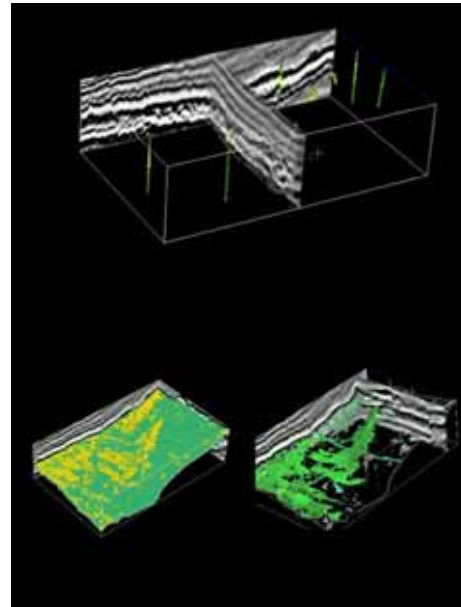


Figura 3. Sessão sísmica, classificação e reconstrução 3D.

Conclusão

Nenhum algoritmo de classificação é adequado para todo e qualquer dado. Aqui é proposta uma implementação operacional de um classificador estatístico de dados sísmicos que proporciona excelentes resultados. A combinação de análise de agrupamento (clustering analysis) associada ao espaço de Fourier de sinais, usando tanto métodos baseados em núcleos quanto máquinas de suporte vetorial, representa uma poderosa e inédita ferramenta de construção de um consistente elo entre a classificação de facies sedimentares ou petrofísicas e a sua relação com a sísmica.

As principais vantagens dos métodos baseados em núcleo são:

1. A necessidade de poucos dados para treinar funcionalmente o modelo.
2. A definição de núcleos pode ser visualizada e testada de uma forma bastante interativa.
3. A saída do classificador são volumes ou superfícies e assim a continuidade das fácies podem ser naturalmente alcançadas.

Agradecimentos

Agradecemos à Petrobras pelo contínuo suporte ao Projeto Sismolitofacies, através de convênio tecnológico com a PUC-Rio, dentro do qual são desenvolvidas as teorias aqui ilustradas.

Referências

Kuzman, H., (2003) A support vector machine for AVO interpretation, 2003 SEG Abstracts, AVO 1., 181-184.

Li, J., and J. Castagna (2004), Support Vector Machine (SVM) pattern recognition to AVO classification, *Geophys. Res. Lett.*, 31, L02609, GL018299.

Li, J. (2005), Multiattributes pattern recognition for reservoir prediction, *Geophys. Innovation: 2005 CSEG National Convention*, 205-208.

Shawe-Taylor J., N. Christianini, (2004), *Kernel Based Methods for Pattern Analysis*, Cambridge, 2004.

Vapnik, V., (1998) *Statistics learning theory*, John Wiley & Sons Inc.

Zhao, B, Zhou H., (2005) Nonlinear Classification of AVO Attributes Using SVM, *Nonlinear Geophysics*.