

LLC-KBDM: uma abordagem de aprendizado de máquina no ajuste de dados de espectroscopia por RM

Danilo Mendes Dias Delfino da Silva e Fernando Fernandes Paiva

Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brasil.

Introdução: A análise espectral de sinais de Ressonância Magnética (RM) pode ser utilizada para calcular a concentração de compostos visíveis a esta técnica. Em aplicações *in vivo*, é possível determinar a concentração de um determinado metabólito. Para tal, é necessário obter a área sob a curva dos respectivos picos no domínio da frequência. Desta forma, quando há superposição, não é possível calcular diretamente esta área. Dois grandes grupos de soluções para este problema existem na literatura: os métodos baseados em ajuste de uma base adquirida experimentalmente ou simulada; e os métodos baseados em ajuste do sinal em um modelo baseado em funções matemáticas. Este trabalho faz uso do *Krylov Basis Diagonalization Method* (KBDM), que realiza o ajuste do sinal, do ponto de vista espectral, em funções lorentzianas [1].

Métodos: O KBDM é um algoritmo que tem como produto final M tuplas contendo os parâmetros de cada pico, onde cada tupla é formada pelas amplitudes, taxas de relaxação, frequências e fases. Se o sinal possui K picos lorentzianos, os $M - K$ picos restantes serão nulos. De fato, isto ocorre apenas se o sinal se encaixa perfeitamente ao modelo ou não possui ruído. Definitivamente, não é o que ocorre em aplicações de espectroscopia *in vivo* por RM. Desta forma, picos denominados espúrios são criados para compor a contribuição do ruído e as demais imperfeições do modelo.

A nossa abordagem, denominada LLC-KBDM (*Line List Clustering-KBDM*), se baseia na utilização de múltiplos resultados do cálculo do KBDM para um mesmo sinal, porém, variando-se o valor de M em um determinado intervalo. Estes resultados são ligeiramente diferentes e quando representados em um conjunto de variáveis adequadas, é possível perceber a formação de agrupamentos relacionados a cada pico. A contribuição do ruído tende a se espalhar, não formando nenhum ou quase nenhum agrupamento. Estes agrupamentos são obtidos pelo algoritmo DBSCAN [2], que realiza automaticamente a filtragem dos pontos esparsos relacionados ao ruído. O centróide de cada agrupamento é utilizado para se obter a estimativa final dos parâmetros de cada linha.

Resultados e Discussões: A figura 1 mostra um espectro simulado contendo ruído típico em aplicações *in vivo* e o respectivo ajuste utilizando o LLC-KBDM. A figura 2 mostra a comparação dos resíduos do LLC-KBDM com o obtido através dos métodos HLSVD (Hankel-Lanczos Singular Value Decomposition) e HLSVD-PRO, uma versão mais robusta do anterior, para diferentes níveis de ruído. O resíduo é calculado a partir da soma das diferenças absolutas entre cada ponto do espectro simulado sem ruído e o espectro estimado. Foram utilizadas as implementações do software jMRUI para os algoritmos HLSVD e HLSVD-PRO. O LLC-KBDM foi computado a partir de implementação própria desenvolvida em Python.

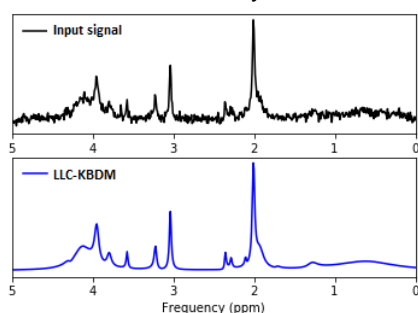


Figura 1: LLC-KBDM aplicado em sinal simulando aquisição em cérebro humano em equipamento de 3T.

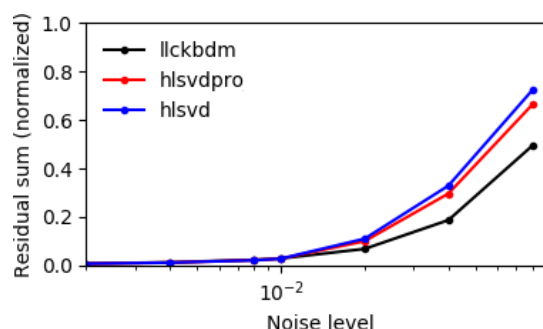


Figura 2: resíduo normalizado para diferentes níveis de ruído.

Conclusões: O LLC-KBDM mostrou-se superior aos métodos tradicionais HLSVD e HLSVD-PRO para o sinal simulado apresentado para diferentes níveis de ruído, conforme mostrado na Figura 2. Em função da utilização de médias através do cálculo dos centroides dos agrupamentos, o LLC-KBDM tende a ser um algoritmo mais estável em relação aos demais algoritmos não supervisionados que utilizam modelos de funções matemáticas no ajuste do sinal.

Referências:

- [1] MANDELSHTAM, V. A. FDM: The filter diagonalization method for data processing in NMR experiments. *Progress in Nuclear Magnetic Resonance Spectroscopy*, v. 38, n. 2, p. 159–196, mar. 2001.
- [2] ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. AAAI Press, 1996.