Statistics > Methodology

A Bayesian Model of NMR Spectra for the Deconvolution and Quantification of Metabolites in Complex Biological Mixtures

William Astle, Maria De Iorio, Sylvia Richardson, David Stephens, Timothy Ebbels

(Submitted on 11 May 2011 (v1), last revised 15 May 2012 (this version, v3))

Nuclear Magnetic Resonance (NMR) spectra are widely used in metabolomics to obtain profiles of metabolites dissolved in biofluids such as cell supernatants. Methods for estimating metabolite concentrations from these spectra are presently confined to manual peak fitting and to binning procedures for integrating resonance peaks. Extensive information on the patterns of spectral resonance generated by human metabolites is now available in online databases. By incorporating this information into a Bayesian model we can deconvolve resonance peaks from a spectrum and obtain explicit concentration estimates for the corresponding metabolites. Spectral resonances that cannot be deconvolved in this way may also be of scientific interest so we model them jointly using wavelets.

We describe a Markov chain Monte Carlo algorithm which allows us to sample from the joint posterior distribution of the model parameters, using specifically designed block updates to improve mixing. The strong prior on resonance patterns allows the algorithm to identify peaks corresponding to particular metabolites automatically, eliminating the need for manual peak assignment. We assess our method for peak alignment and concentration estimation. Except in cases when the target resonance signal is very weak, alignment is unbiased and precise. We compare the Bayesian concentration estimates to those obtained from a conventional numerical integration method and find that our point estimates have sixfold lower mean squared error.

Finally, we apply our method to a spectral dataset taken from an investigation of the metabolic response of yeast to recombinant protein expression. We estimate the concentrations of 26 metabolites and compare to manual quantification by five expert spectroscopists. We discuss the reason for discrepancies and the robustness of our methods concentration estimates.

Subjects: **Methodology (stat.ME)**; Quantitative Methods (q-bio.QM); Applications (stat.AP)

Cite as: **arXiv:1105.2204 [stat.ME]**
(or **arXiv:1105.2204v3 [stat.ME]** for this version)

Submission history

From: William Astle [view email]
[v1] Wed, 11 May 2011 14:24:04 GMT (175kb)
[v2] Tue, 28 Feb 2012 23:48:33 GMT (840kb)
[v3] Tue, 15 May 2012 18:36:56 GMT (832kb)

Which authors of this paper are endorsers?