NeuroImage 175 (2018) 272–285

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Bayesian uncertainty quantification in linear models for diffusion MRI

Jens Sjölund [a,b,c,*], Anders Eklund [b,c,d], Evren Özarslan [b,c], Magnus Herberthson [e], Maria Bånkestad [f], Hans Knutsson [b,c]

[a] Elekta Instrument AB, Kungstensgatan 18, Box 7593, SE-103 93, Stockholm, Sweden
[b] Department of Biomedical Engineering, Linköping University, Linköping, Sweden
[c] Center for Medical Image Science and Visualization (CMIV), Linköping University, Sweden
[d] Department of Computer and Information Science, Linköping University, Linköping, Sweden
[e] Department of Mathematics, Linköping University, Linköping, Sweden
[f] RISE SICS, Isafjordsgatan 22, Box 1263, SE-164 29, Kista, Sweden

ARTICLE INFO

Keywords:
Diffusion MRI
Uncertainty quantification
Signal estimation

ABSTRACT

Diffusion MRI (dMRI) is a valuable tool in the assessment of tissue microstructure. By fitting a model to the dMRI signal it is possible to derive various quantitative features. Several of the most popular dMRI signal models are expansions in an appropriately chosen basis, where the coefficients are determined using some variation of least-squares. However, such approaches lack any notion of uncertainty, which could be valuable in e.g. group analyses. In this work, we use a probabilistic interpretation of linear least-squares methods to recast popular dMRI models as Bayesian ones. This makes it possible to quantify the uncertainty of any derived quantity. In particular, for quantities that are affine functions of the coefficients, the posterior distribution can be expressed in closed-form. We simulated measurements from single- and double-tensor models where the correct values of several quantities are known, to validate that the theoretically derived quantiles agree with those observed empirically. We included results from residual bootstrap for comparison and found good agreement. The validation employed several different models: Diffusion Tensor Imaging (DTI), Mean Apparent Propagator MRI (MAP-MRI) and Constrained Spherical Deconvolution (CSD). We also used in vivo data to visualize maps of quantitative features and corresponding uncertainties, and to show how our approach can be used in a group analysis to downweight subjects with high uncertainty. In summary, we convert successful linear models for dMRI signal estimation to probabilistic models, capable of accurate uncertainty quantification.

Introduction

Diffusion magnetic resonance imaging (dMRI) permits the noninvasive assessment of tissue microstructure. By fitting a model of the dMRI signal in each voxel it is possible to derive various quantitative features. Examples include voxel-based scalar indices such as fractional anisotropy (FA) or return to origin probability (RTOP), as well as the pairwise probability of a virtual fiber being traced between two points. Using such measures it is possible to perform statistical group analyses. For the reliability of such tests it is essential to quantify the uncertainty of the relevant measures. This fact is well-established in the field of functional magnetic resonance imaging (fMRI) — where it is common to, for example, downweight subjects with a high variance (Chen et al., 2012; Woolrich et al., 2004) — but not so for dMRI; the most popular approach for doing FA group comparisons (Smith et al., 2006) ignores uncertainty in FA.

A large body of research has been devoted to formulating new diffusion models that address the inability of diffusion tensor imaging (DTI) (Basser et al., 1994) to resolve crossing and kissing fibers. A common trait among several of the most widely used methods is that they expand the signal in an appropriately chosen functional basis. Often, the coefficients of the expansion are then determined by some variation of linear least-squares (see Table 1 for a non-exhaustive list). This is precisely the type of models we are concerned with in this paper — linear (in the coefficients) models fitted with least-squares. Incidentally, some non-parametric models (Andersson and Sotiropoulos, 2015; Sjölund et al., 2017) also belong to this class, although we will not elaborate on the connection here. Our key observation is that linear models fitted with least-squares are amenable to a probabilistic reinterpretation. We will show that it follows almost immediately that, under the same assumptions as in the fitting, it is possible to determine the full posterior

* Corresponding author. Elekta Instrument AB, Kungstensgatan 18, Box 7593, SE-103 93, Stockholm, Sweden.
E-mail address: jens.sjolund@elekta.com (J. Sjölund).

https://doi.org/10.1016/j.neuroimage.2018.03.059
Received 16 November 2017; Received in revised form 16 February 2018; Accepted 25 March 2018
Available online 29 March 2018

Table 1
An assortment of linear models used in dMRI. We have used the notation $\mathbf{q} = q\,\hat{\mathbf{q}}$ where applicable. For clarity, some details have been omitted.

Method	y	$\phi(\mathbf{x})$	$\mathbf{c}$	W	Λ
Diffusion Tensor Imaging (DTI) (weighted least-squares) (Basser et al., 1994; Salvador et al., 2005)	$\log S(\mathbf{q})$	$1,\ q_i q_j$	$\log S_0,\ D_{ij}$	$\mathrm{diag}(S^2)$	0
Diffusion Kurtosis Imaging (DKI) (weighted least-squares) (Veraart et al., 2013)	$\log S(\mathbf{q})$	$1,\ q_i q_j,\ q_i q_j q_k q_l$	$\log S_0,\ D_{ij},\ K_{ijkl}$	$\mathrm{diag}(S^2)$	0
Q-space Trajectory Imaging (QTI) (Westin et al., 2016)	$\log S(B)$	$1,\ B,\ B^{\otimes 2}$	$\log S_0,\ \langle D\rangle,\ \mathbb{C}$	$\mathrm{diag}(S^2)$	0
Constrained Spherical Deconvolution (CSD) (Tournier et al., 2007)	$S(\hat{\mathbf{q}})$	$\chi(\hat{\mathbf{q}})Y_l(\hat{\mathbf{q}})$	c_l	I	L
Q-Ball Imaging (QBI) (Descoteaux et al., 2007; Tuch, 2004)	$S(\hat{\mathbf{q}})$	$Y_l(\hat{\mathbf{q}})$	c_l	I	$\int_{S^2} \lVert\Delta_b S\rVert^2 d\Omega$
MAP-MRI with Laplacian regularization (MAPL) (Fick et al., 2016; Özarslan et al., 2013)	$S(\mathbf{q})$	$\Phi_n(u,q)$	c_n	I	$\int_{\mathbb{R}^3} \lVert\Delta S\rVert^2 d\mathbf{q}$
Spherical Polar Fourier (SPF) (Assemlal et al., 2009)	$S(\mathbf{q})$	$R_k(q)Y_l(\hat{\mathbf{q}})$	c_{kl}	I	$\Lambda_R + \Lambda_Y$

y: response variable.
$\mathbf{x}$: input variables.
$\phi(\mathbf{x})$: basis function.
$\mathbf{c}$: coefficients.
W: inverse noise correlation matrix.
Λ: regularization matrix.
S: signal.
S_0: non-diffusion weighted signal.
χ: single fiber response function.
Y_l: real spherical harmonics, cf. (Descoteaux et al., 2007).
L: matrix determined iteratively using a sparsifying heuristic.
Δ: Laplace operator.
Δ_b: Laplace-Beltrami operator.
Φ_n: Hermite functions scaled by a factor u.
R_k: Gaussian Laguerre polynomials.
Λ_R: diagonal matrix penalizing higher radial orders.
Λ_Y: diagonal matrix penalizing higher angular orders.

distribution of the coefficients — not just a point estimate.

This is, of course, not the first time someone has taken a probabilistic view on signal estimation in dMRI. It is well-known that the noisy signal in MRI follows a Rician, or more generally a non-central Chi, distribution (Gudbjartsson and Patz, 1995), which on the other hand is approximately Gaussian when the signal-to-noise ratio is, at least, moderately high ($\gtrsim$3). In DTI (Basser et al., 1994) and diffusion kurtosis imaging (DKI) (Jensen et al., 2005), it is common to fit a linear model to the logarithm of the signal. The resulting log-Rician distribution is again approximately normal for moderate signal-to-noise ratios, but with signal-dependent (heteroscedastic) noise (Salvador et al., 2005). In such cases, weighted least-squares has been shown to work well (Rawlings et al., 2001; Veraart et al., 2013). Error propagation in DTI fitted with nonlinear least-squares has also been investigated (Koay et al., 2007).

Bayesian methods (Behrens et al., 2003; Gelman et al., 2013; Gu et al., 2017; Wegmann et al., 2017) are distinctly different from methods using least-squares to find point estimates. By assuming parametric probability distributions for the likelihood and for every parameter in the model, Bayesian methods make it possible to derive the probability distribution of any quantity of interest, at least in principle. Actually evaluating such distributions, however, typically relies extensively on sampling methods such as Markov Chain Monte Carlo.

Bootstrapping is a frequentistic alternative to Bayesian models. The general idea is to approximate the underlying probability distribution with an empirical one. Samples are drawn with replacement and for each draw the parameter of interest is calculated. By repeating this procedure, it is possible to approximate the sampling distribution of the relevant parameter. Naturally, the quality of the approximation degrades as the number of empirical samples decreases. Residual bootstrap (Chung et al., 2006) and wild bootstrap (Whitcher et al., 2008) are two forms of model-based bootstrapping that have been applied to dMRI. In residual bootstrap, the normalized residuals (after fitting a model) are resampled, and the fitting procedure is repeated for this new draw, after which the parameter value of interest is recorded. All residuals are assumed to have identical distributions and resampling is done freely among them. In wild bootstrap, on the other hand, modified residuals are randomly added or subtracted to the fitted point where they originated from, without being distributed to other design points. Model-based bootstrapping is only reliable insofar as the model can adequately describe the measured diffusion signals (Yuan et al., 2008). Basically all of the methods described above have been applied to tractography (Behrens et al., 2007; Berman et al., 2008; Haroon et al., 2009; Jeurissen et al., 2011; Jones, 2003, 2008).

In this work, instead of starting out by assuming more or less contrived prior distributions for the coefficients and the likelihood, we look at methods that are tried and tested and see what the corresponding priors are. Surprisingly, it results in a simple closed form expression for the posterior distribution of coefficients, and by extension also of all parameters linear in the coefficients. In other cases there is no need to repeat the whole fitting procedure, as in bootstrapping methods. Since the posterior is available in closed form it is very efficient to sample from it directly.

Theory

To recapitulate, we focus on linear models fitted with linear least-squares. This might sound restrictive at first but — as can be seen in Table 1 — it encompasses several of the most widely used models in