Journal of Neuroscience Methods 257 (2016) 147–158

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Fast estimation of diffusion tensors under Rician noise by the EM algorithm

Jia Liu[a,*], Dario Gasbarra[b], Juha Railavo[c]

[a] Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box (MaD), FI40014, Finland
[b] Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FI00014, Finland
[c] Helsinki University Hospital, Finland

HIGHLIGHTS

- We originally implement the EM algorithm under data augmentation version in DTI.
- We propose a fast computational scheme for diffusion tensor estimation under the Rician noise model.
- The proposed EM approach is superior in terms of computational burden and estimating accuracy.
- Performance is shown by both mathematical interpretation and numerical comparison.

ARTICLE INFO

Article history:
Received 7 January 2015
Received in revised form
26 September 2015
Accepted 27 September 2015
Available online 9 October 2015

Keywords:
Data augmentation
Fisher scoring
Maximum likelihood estimator
Maximum a posteriori estimator
Rician Likelihood
Reduced computation

ABSTRACT

Diffusion tensor imaging (DTI) is widely used to characterize, in vivo, the white matter of the central nerve system (CNS). This biological tissue contains much anatomic, structural and orientational information of fibers in human brain. Spectral data from the displacement distribution of water molecules located in the brain tissue are collected by a magnetic resonance scanner and acquired in the Fourier domain. After the Fourier inversion, the noise distribution is Gaussian in both real and imaginary parts and, as a consequence, the recorded magnitude data are corrupted by Rician noise.

Statistical estimation of diffusion leads a non-linear regression problem. In this paper, we present a fast computational method for maximum likelihood estimation (MLE) of diffusivities under the Rician noise model based on the expectation maximization (EM) algorithm. By using data augmentation, we are able to transform a non-linear regression problem into the generalized linear modeling framework, reducing dramatically the computational cost. The Fisher-scoring method is used for achieving fast convergence of the tensor parameter. The new method is implemented and applied using both synthetic and real data in a wide range of b-amplitudes up to 14,000 s/mm^2. Higher accuracy and precision of the Rician estimates are achieved compared with other log-normal based methods. In addition, we extend the maximum likelihood (ML) framework to the maximum a posteriori (MAP) estimation in DTI under the aforementioned scheme by specifying the priors. We will describe how close numerically are the estimators of model parameters obtained through MLE and MAP estimation.

1. Introduction

Diffusion tensor imaging (DTI) is a powerful tool to detect, in vivo, the white matter anatomy and structures of the brain. The raw MR-data are collected by a magnetic resonance scanner and consist of spectral measurement from the displacement distribution of water molecules constrained into cellular structures. Diffusion anisotropy characterizes the nervous fibers.

After the Fourier inversion, the MR-signals are corrupted by a complex Gaussian noise, and consequently, the recorded measurement magnitudes, referred as diffusion weighted magnetic resonance imaging (DW-MRI) data, will follow the Rician distribution. The complex noise is composed of two components, where the real and imaginary parts are still independently Gaussian (Henkelman, 1985; Koay et al., 2009; Zhu et al., 2007). The simplest method for diffusion tensor estimation (DTE) is based on the linearized log-normal regression model, where the residual variance is assumed

* Corresponding author. Tel.: +358 294151407.
E-mail address: jia.liu@jyu.fi (J. Liu).

http://dx.doi.org/10.1016/j.jneumeth.2015.09.029
0165-0270/

to be either constant (the least squares) or depending on the signal amplitude (the weighted least squares). These Gaussian noise models fail to fit the high frequency data, which carry information about the higher order diffusion characteristics. In the existing literature (Rajan et al., 2011; Veraart et al., 2011; Andersson, 2008) on the ML-estimation of diffusion tensors under the Rician noise, the maximization algorithm involves repeated computation of modified Bessel functions. By using data augmentation we are able to replace the Rician likelihood by a Poisson likelihood which is standard in the generalized linear modeling (GLM) framework.

Such simplification reduces dramatically the computational burden of the Fisher-scoring maximization algorithm. This applies also at high b-amplitudes, where in the low signal regime measurements below a threshold are customarily coded as zeros. In the standard LS or WLS approaches, zero-measurements are problematic since they cannot be fitted by a log-normal distribution, and simply discarding them induces selection bias. The appropriately modeled noise level provides capability of data correction in further insights, e.g. removing artefacts from the raw data.

This paper is structured as follows. Section 2 describes the noise in MRI and data augmentation, specifying the statistical model for DTE. In Section 3 we discuss the implementation of the EM and the Fisher-scoring algorithms in the DTI context. In addition, we also specify priors for the parameters and discuss the computation of the maximum a posteriori estimator (MAPE) under the same scheme. Section 4 illustrates the results from both synthetic and real data. Section 5 details the method comparisons. In Section 6 we conclude with an overview of the methods and the undergoing developments. Theoretical details are left for the appendices.

2. GLM for MRI observations

2.1. Rician noise in MRI

In magnetic resonance imaging (MRI), we usually need to take the noise in the raw MR-acquisitions into account. The complex valued noise ε is composed of two *i.i.d.* Gaussian random variables with zero mean and variance σ^2, one for the real and the other one for the imaginary component. After the Fourier inversion, the signal intensity $S \geq 0$ is corrupted by a complex Gaussian noise, and $Y = |S + \varepsilon|$ will be observed. Consequently, the observed MR-signal magnitudes follow a Rician distribution resulting in the likelihood function

$$p_{S,\sigma^2}(y) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2+S^2}{2\sigma^2}\right) I_0\left(\frac{yS}{\sigma^2}\right), \tag{1}$$

where I_α is the α-order modified Bessel function of first kind. For $\alpha = 0$ it has also the following representation in terms of Gaussian hypergeometric series (Jeffrey and Zwillinger, 2007):

$$I_0(2\tau) = {}_0F_1(1, \tau^2) = \sum_{n=0}^{\infty} \frac{\tau^{2n}}{(n!)^2}. \tag{2}$$

Let $t = S^2/(2\sigma^2)$, then Eq. (1) gives

$$P_{t,\sigma^2}(Y \in dy) = \frac{y}{\sigma^2} \exp\left(-t - \frac{y^2}{2\sigma^2}\right) I_0\left(\frac{y}{\sigma}\sqrt{2t}\right) dy \tag{3}$$

with $\tau = yS/(2\sigma^2) = \sqrt{2t}y/(2\sigma)$.

2.2. Data augmentation

We follow the strategy presented in Gasbarra and Liu (2014) implementing augmented data N from a Poisson distribution with mean $t > 0$. The likelihood for the observed data can be transformed from the Rician likelihood equation (3) to a joint augmented density

$$\begin{aligned} P_{t,\sigma^2}(N = n, Y^2 \in dy^2) &= P_{t,\sigma^2}(N = n, X \in dx) \\ &= P_t(N = n) P_{\sigma^2}(X \in dx | N = n) \\ &= \frac{(tx)^n}{(n!)^2 (2\sigma^2)^{n+1}} \exp\left(-t - \frac{x}{2\sigma^2}\right) dx, \end{aligned} \tag{4}$$

where X is from the conditional distribution Gamma$(N+1, 1/(2\sigma^2))$ given N. Eq. (4) provides a transformation from a non-linear regression problem to the GLM framework

$$f_{\xi,\phi}(z) = c(z,\phi) \exp\left(\frac{z\xi - a(\xi)}{\phi}\right) \tag{5}$$

with z corresponding to the response in general, see McCullagh and Nelder (1989) for more details.

3. Method

3.1. DW-MRI and parametrization

In DW-MRI, the signal is modeled as the first equality

$$S(\mathbf{q}) = S_0 \exp(-bd(\mathbf{g})) = S_0 \exp\left(Z\theta\right), \tag{6}$$

where the control vector $\mathbf{q} \in \mathbb{R}^3$ is determined by the sequence of gradient pulses, $b = |\mathbf{q}|^2$, and $\mathbf{g} = \mathbf{q}/|\mathbf{q}| \in \mathcal{S}^2$ is a vector of unit length. The MR-signal decays exponentially with respect to the b-amplitude. Depending on the gradient direction $\mathbf{g}$ the decay is modeled by the reflection symmetric diffusivity function $d : \mathcal{S}^2 \to \mathbb{R}^+$.

Great efforts have been devoted to modeling the diffusivity, and in general we can have parametrization as the second equality in Eq. (6). In the simplest model the diffusivity is expressed by a symmetric and positive definite rank-2 tensor $D \in \mathbb{R}^{3\times 3}$, giving

$$\log S(\mathbf{q}) = \log S_0 - b\mathbf{g}^\top D\mathbf{g} = \log S_0 + Z\theta,$$

where in the left hand side the diffusion tensor is parametrized as

$$\theta = (\theta_1, \ldots, \theta_6)^\top := \left(D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}\right)^\top$$

with a design matrix

$$Z = Z(\mathbf{q}) = -b(\mathbf{g}_x^2, \mathbf{g}_y^2, \mathbf{g}_z^2, 2\mathbf{g}_x\mathbf{g}_y, 2\mathbf{g}_x\mathbf{g}_z, 2\mathbf{g}_y\mathbf{g}_z).$$

In high angular resolution models (HARDI) (see, e.g. Barmpoutis et al., 2009), the diffusivity is modeled with a totally symmetric Cartesian tensor D of order $n \in \mathbb{N}$, as

$$d(\mathbf{g}) := \sum_{\ell_1=1}^{3}\sum_{\ell_2=1}^{3}\cdots\sum_{\ell_{2n}=1}^{3} D_{\ell_1,\ell_2,\ldots,\ell_{2n}} g_{\ell_1} g_{\ell_2} \cdots g_{\ell_{2n}}.$$

3.2. EM in MLE

In the optimization of the likelihood, we employ the EM (expectation-maximization) algorithm, which is one among the iterative methods in the MLE or in the maximum a posteriori estimation (MAPE). The EM algorithm proceeds in two steps and shortens the computational complexity by using augmented data. In terms of our case, in the E-step we calculate the expectation of the log-likelihood w.r.t. the conditional distribution of N given by the observations and other parameters with fixed values. In the M-step, we find the ML parameter of S_0^2 and σ^2 by maximizing the augmented log-likelihood quantities. The computational details are listed in Appendix A.