Expectation–maximization algorithm with a nonlinear Kalman smoother for MEG/EEG connectivity estimation

Narayan Puthanmadam Subramaniyam[1], Filip Tronarp[2], Simo Särkkä[2] and Lauri Parkkonen[1]

[1] Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland
[2] Department of Electrical Engineering and Automation, Aalto University, Espoo, Finland

Abstract— **Current techniques to determine functional or effective connectivity from magnetoencephalography (MEG) and electroencephalography (EEG) signals typically involve two sequential steps: 1) estimation of the source current distribution from the sensor data, for example, by minimum-norm estimation or beamforming, and 2) fitting a multivariate autoregressive (MVAR) model to estimate the AR coefficients, which reflect the interaction between the sources. Here, we introduce a combination of the expectation–maximization (EM) algorithm and a nonlinear Kalman smoother to perform joint estimation of both source and connectivity (linear and nonlinear) parameters from MEG/EEG signals. Based on simulations, we show that the proposed approach estimates both the source signals and AR coefficients in linear models significantly better than the traditional two-step approach when the signal-to-noise ratio (SNR) is low ($\leq$1) and gives comparable results at higher SNRs ($>$1). Additionally, we show that nonlinear interaction parameters can be reliably estimated from MEG/EEG signals at low SNRs using the EM algorithm with sigma-point Kalman smoother.**

Keywords— **MEG, EEG, functional connectivity, nonlinear Kalman smoother, expectation–maximization**

I. Introduction

The transfer and integration of information between anatomically segregated brain areas can be analyzed in terms of functional or effective connectivity, which quantifies the directed, linear or nonlinear, interactions between several brain regions [1, 2]. Magnetoencephalography (MEG) and electroencephalography (EEG) can capture signals originating from such interaction at high temporal resolution. Functional connectivity can be estimated from MEG/EEG data using model-driven techniques such as dynamic causal modeling (DCM) [3] and data-driven methods such as Granger causality that use multivariate autoregressive (MVAR) models [4].

When data-driven methods are used to estimate connectivity [5, 6], a conventional approach involves first solving the MEG/EEG inverse problem to estimate cortical signals at a set of regions of interest (ROIs). Then, an MVAR-model is fit (using, e.g., least-squares method or Yule–Walker (YW) algorithm). Finally, using the MVAR coefficients, connectivity measures like partial directed coherence (PDC) or directed transfer function (DTF) are estimated in the frequency domain to find significant interactions [7]. The limitation of such a sequential approach is that it completely ignores the information provided by the connectivity structure for the identification of source parameters. Furthermore, methods like PDC are not robust to nonlinear interactions [8]. Recent attempts have been made at joint estimation of source and connectivity parameters. Cheung and colleagues used a state-space formulation to estimate MVAR parameters using an expectation–maximization (EM) approach; however, they only considered linear interactions [9]. More recently, Fukushima and colleagues proposed estimation of high-dimensional MVAR models using functional magnetic resonance imaging (fMRI) priors on the spatial patterns of source activity and a sparse prior on MVAR coefficients, again considering only linear interactions [10].

In this work we propose an algorithm to jointly estimate the source parameters along with both linear and nonlinear interaction parameters. To this end, we use an EM algorithm together with a Kalman smoother (KS) [11] to estimate parameters of linear interactions. For nonlinear interactions, we use a sigma-point KS with EM for approximate parameter estimation [12]. The EM algorithm gives a maximum-likelihood (ML) estimate or its approximation (when the model is nonlinear) using an iterative approach [11–13]. We use a state-space model (SSM) that includes a dynamic model describing the interaction between the source signals and a measurement model describing how source signals are mapped to MEG/EEG sensors. The SSM includes both observational and process noise, which are assumed to be additive and zero-mean Gaussian. We compare our approach to the traditional sequential estimation of connectivity parameters using a minimum-norm estimate (MNE) [14] for source models and the YW algorithm for the interaction parameters (i.e., the AR coefficients). For all tests, we used simulated MEG recordings generated with a realistically-shaped three-compartment boundary element method (BEM) head model.

© Springer Nature Singapore Pte Ltd. 2018
H. Eskola et al. (eds.), *EMBEC & NBC 2017*,
IFMBE Proceedings 65,
DOI: 10.1007/978-981-10-5122-7_191

II. METHODS

A. Bayesian state-space model

Consider a source-space partitioned into n cortical regions. Let the output of each cortical region at time $t \in \mathbb{Z}$ be denoted as $x_{j,t}$, where $j \in \mathbb{N}^+$, and represent the source amplitude of the representative dipole of j'th region at time t. The temporal evolution of such an n-dimensional cortical state vector $\mathbf{x}_t = [x_{1,t}, x_{2,t}, \cdots, x_{n,t}]^T$ and interaction between them is modeled using a p'th-order, multivariate autoregressive (MVAR) model

$$\mathbf{x}_t = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{x}_{t-k} + \mathbf{q}_t \tag{1}$$

where matrix $\mathbf{A}_k$ contains the AR coefficients (interaction parameters) at time $t-k$ and is given as

$$\mathbf{A}_k = \begin{pmatrix} a_{11,k} & a_{12,k} & \cdots & a_{1n,k} \\ a_{21,k} & a_{22,k} & \cdots & a_{2n,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1,k} & a_{n2,k} & \cdots & a_{nn,k} \end{pmatrix} \tag{2}$$

and the process noise is given as $\mathbf{q}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. The corresponding noisy MEG observations can be described by the measurement model

$$\mathbf{y}_t = \mathbf{L}\mathbf{x}_t + \mathbf{r}_t, \quad \mathbf{r}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{3}$$

where $\mathbf{L}$ is the $m \times n$ lead-field matrix that maps the n-dimensional cortical state vector to the m-dimensional MEG measurement vector $\mathbf{y}_t = [y_{1,t}, y_{2,t}, \cdots, y_{m,t}]^{\mathsf{T}}$.

B. Expectation–Maximization algorithm

Given the measurement sequence $\mathbf{Y}_{1:T} = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_t\}$, and estimates for the parameter $\Theta^i = \{\mathbf{A}^i, \mathbf{Q}^i, \mathbf{R}^i\}$ from the previous iteration i, the expectation step involves the computation of $\mathcal{Q}$,

$$\mathcal{Q} \equiv E_{X|Y}[\log p(\mathbf{Y}_{1:T}, \mathbf{X}_{1:T} | \Theta)] \tag{4}$$

and the maximization-step involves the differentiation of $\mathcal{Q}$ with respect to model parameter Θ. The update for $\Theta = \{\mathbf{A}, \mathbf{Q}, \mathbf{R}\}$, in case of linear dynamical system (see Equation 1), is given by

$$\begin{aligned} \mathbf{A}^{i+1} &= \Big(\sum_{t=2}^{T} \mathbf{V}_{t,t-1|T}\Big)\Big(\sum_{t=2}^{T} \mathbf{V}_{t-1|T}\Big)^{-1} \\ \mathbf{Q}^{i+1} &= \frac{1}{T-1}\Big(\sum_{t=2}^{T} \mathbf{V}_{t|T} - \mathbf{A}^{i+1}\sum_{t=2}^{T} \mathbf{V}_{t-1|T}\Big) \\ \mathbf{R}^{i+1} &= \frac{1}{T}\sum_{t=1}^{T}\Big[(\mathbf{y}_t - \mathbf{L}\hat{\mathbf{x}}_{t|T})(\mathbf{y}_t - \mathbf{L}\hat{\mathbf{x}}_{t|T})^{\mathsf{T}} + \mathbf{L}\mathbf{P}_{t|T}\mathbf{L}^{\mathsf{T}}\Big] \end{aligned} \tag{5}$$

where,

$$\begin{aligned} \mathbf{V}_{t|T} &\equiv E(\mathbf{x}_t \mathbf{x}_t^{\mathsf{T}} | \mathbf{Y}_{1:T}) = \mathbf{P}_{t|T} + \hat{\mathbf{x}}_{t|T}\hat{\mathbf{x}}_{t|T}^{\mathsf{T}} \\ \mathbf{V}_{t,t-1|T} &\equiv E(\mathbf{x}_t \mathbf{x}_{t-1}^{\mathsf{T}} | \mathbf{Y}_{1:T}) = \mathbf{P}_{t,t-1|T} + \hat{\mathbf{x}}_{t|T}\hat{\mathbf{x}}_{t-1|T}^{\mathsf{T}} \end{aligned} \tag{6}$$

and the expectation-step employs a linear Kalman filter and smoother to compute $\hat{\mathbf{x}}_{t|T}$ and $\mathbf{P}_{t|T}$ [11]. Now suppose the system is modeled by non-linear dynamics according to

$$\mathbf{x}_t = \mathbf{A}\mathbf{f}(\mathbf{x}_{t-1}) + \mathbf{q}_t. \tag{7}$$

The smoothing density may then be approximated by sigma-point smoothers [11], which enables an approximate implementation of the EM iteration [12]. Since the measurement equation is unchanged, its EM update is the same as in the linear case. The update for $\mathbf{Q}$ and $\mathbf{A}$ are, however, altered according to

$$\mathbf{A}^{i+1} = \left(\sum_{t=2}^{T} E_{X|Y}[\mathbf{x}_t \mathbf{f}^{\mathsf{T}}(\mathbf{x}_{t-1})]\right) \times \left(\sum_{t=2}^{T} E_{X|Y}[\mathbf{f}(\mathbf{x}_{t-1})\mathbf{f}^{\mathsf{T}}(\mathbf{x}_{t-1})]\right)^{-1} \tag{8}$$

$$\mathbf{Q}^{i+1} = \frac{1}{T-1}\sum_{t=2}^{T} E_{X|Y}[(\mathbf{x}_t - A\mathbf{f}(\mathbf{x}_{t-1}))(\mathbf{x}_t - A\mathbf{f}(\mathbf{x}_{t-1}))^{\mathsf{T}}] \tag{9}$$

where the expectations are computed using the smoothing solution of $(\mathbf{A}^i, \mathbf{Q}^i, \mathbf{R}^i)$; For an efficient sigma-point implementation, see Ref. [12].

C. Simulation setup

We used a three-compartment realistically-shaped BEM head model (the sample data provided by the MNE software [15]) that included inner skull, outer skull and skin with conductivities 0.3 S/m, 0.06 S/m and 0.3 S/m, respectively. The source space was created using octahedron subdivision and fixed source orientation, resulting in approximately 8000 source dipoles distributed evenly across the cortical mantle. The active dipoles were placed in three ROIs: 1) Broadmann area 1 (BA1; part of primary somatosensory cortex, left hemisphere), 2) Broadmann areas 41 and 42 (BA41 and BA42; auditory cortex, left hemisphere) and 3) Broadmann area 17 (BA17; visual area V1, right hemisphere); see Figure 1. Source activity was simulated using a neural mass model [16], and the connectivity between the dipoles was simulated using MVAR models involving both linear and nonlinear interactions [8].