Journal of Statistical Planning and Inference 164 (2015) 39–53

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

On nonparametric feature filters in electromagnetic imaging

Jian Zhang

School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7NF, UK

ARTICLE INFO

Article history:
Received 21 September 2014
Received in revised form 9 March 2015
Accepted 10 March 2015
Available online 25 March 2015

Keywords:
MEG neuroimaging
Beamforming
Nonparametric feature selection
Source screening and reconstruction

ABSTRACT

Estimation of sparse time-varying coefficients on the basis of time-dependent observations is one of the most challenging problems in statistics. Our study was mainly motivated from magnetoencephalographic neuroimaging, where we want to identify neural activities using the magnetoencephalographic sensor measurements outside the brain. The problem is ill-posed since the observed magnetic field could result from an infinite number of possible neuronal sources. The so-called minimum-variance beamformer is one of data-adaptive nonparametric feature filters to address the above problem in the literature. In this paper, we propose a method of sure feature filtering for a high-dimensional time-varying coefficient model. The new method assumes that the correlation structure of the sensor measurements can be well represented by a set of non-orthogonal variance–covariance components. We develop a theory on the sure screening property of the proposed filters and on when the beamformer-based location estimators are consistent or inconsistent with the true ones. We also derive the lower and upper bounds for the mean filtering errors of the proposed method. The new theory is further supported by simulations and a real data analysis.

1. Introduction

Suppose that we observe an n-dimensional time-course $\mathbf{Y}(t)$ from the model

$$\mathbf{Y}(t) = \int_{\Xi} \mathbf{x}(r, \eta(r))\beta(r, t)dr + \varepsilon(t),$$

where Ξ is a bounded subset of R^3, $\beta(r, t)$ is a latent univariate time source of interest at location r, $\mathbf{x}(r, \eta(r))$ is a design vector with nuisance parameter $\eta(r)$, and $\varepsilon(t)$ is a noise. Assume that $\boldsymbol{\beta}(r, t)$ is sparse, i.e., the temporal variability (called power or the marginal variance) $\mathrm{var}(\beta(r, .)) = 0$ for all $r \in B$ except a few locations (i.e., non-null sources). We want to localize these non-null sources among an infinite number of candidates. Given the limited number of time-courses we observed, the problem is ill-posed and high-dimensional. To simplify it, we discretize the integration, obtaining

$$\mathbf{Y}(t) = \sum_{k=1}^{p} \mathbf{x}(r_k, \eta(r_k))\beta(r_k, t) + \varepsilon(t), \tag{1}$$

where $\Omega = \{r_1, \ldots, r_p\}$ is a sieve (or grid) approximation to the source space. The problem becomes a large-p-small-n problem. Several new methodologies have been developed for addressing large-p-small-n problems in regression settings, including least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996), smoothly clipped absolute deviation

E-mail address: jz79@kent.ac.uk.

http://dx.doi.org/10.1016/j.jspi.2015.03.004
0378-3758/

(SCAD) (Fan and Li, 2001), and correlation screening (SIS) (Fan and Lv, 2008). Many important theoretical results have also been established recently in selection consistency (e.g., Meinshausen and Bühlmann, 2006; Zhao and Yu, 2006; Zhang, 2010). However, all these works focused on finite dimensional features and are therefore not directly applicable to neuroimaging studies, where features are time series.

In this paper, we propose a nonparametric feature filtering procedure for identifying the sparse coefficients. The proposed procedure is general but was initially designed for magnetoencephalography (MEG) neuroimaging. MEG is a technique for mapping brain activity by measuring magnetic fields produced by electrical currents occurring in the brain, using arrays of superconducting quantum interference sensors (Hamalainen et al., 2010). The MEG neuroimaging can be employed to study perceptual and cognitive brain processes, to localize regions affected by pathology, and to determine the function of various parts of the brain. While MEG offers a direct measurement of neural activity with very high temporal resolution, its spatial resolution is relatively low. Concerns over its spatial resolution have raised fundamental issues of methodology and theory. In fact, improving its resolution by virtue of source reconstruction lies at the heart of the entire MEG-based brain mapping enterprise.

In the MEG neuroimaging, $Y_i(t_j)$ is the measurement recorded by the MEG sensor i at time t_j for $1 \leq i \leq n$, $1 \leq j \leq J$, where the time points $t_j = j/\Delta$, the number of the time instants $J = b\Delta$ is determined by the time window b and the sampling rate Δ per second, and the number of the sensors n is of order hundreds. Let $\mathbf{Y}(t_j)$ denote the measurements from all the sensors at time t_j, which are assumed to be induced by potential sources at locations r_k, $1 \leq k \leq p$ along time-invariant orientations $\eta_k \equiv \eta(r_k, t)$, $1 \leq k \leq p$ in the brain respectively. Let $\boldsymbol{\beta}(t_j) = (\beta(r_1, t_j), \ldots, \beta(r_p, t_j))^T$ be the source magnitude vector of these sources at time t_j and $\{\beta(r_k, t_j) : 1 \leq j \leq J\}$ the source time-course at r_k, where the superscript T indicates the matrix transpose. Let $\mathbf{x}_k$ be the output vector of these sensors that would be induced by a unit-magnitude source located at r_k along orientation η_k and $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p)$. As a special case of model (1), the sensor measurements $\mathbf{Y}(t_j)$ may be modeled as

$$\mathbf{Y}(t_j) = \sum_{k=1}^{p} \mathbf{x}_k \beta(r_k, t_j) + \varepsilon(t_j), \tag{2}$$

where $\mathbf{x}_k = l(r_k)\eta_k$, $l(r_k) \in R^{n\times 3}$ (called lead field matrix at location r_k) is derived from a forward physical model of the brain, and $\varepsilon(t_j)$ is the noise vector of the n sensors at time t_j (Sarvas, 1987). To search for unknown sources, a neural activity index for each grid point in the sieve is calculated, creating a source map of brain activity. Important sources can be then identified by filtering the source map. The accuracy of the filtering depends on the sieve size p, and the spatial and temporal dimensions of the MEG measurements (i.e., the number of sensors and the number of time instants). In practice, the sieve size p is often set to a value much larger than n.

The minimum-variance beamforming, a data-adaptive filtering approach to the above source localization problem has been widely used in neuroimaging. In the approach, one scans the source space through a feature space with a series of filters; each is tailored to a particular area of feature (called pass-band) and resistant to confounding effects originating from other areas (called stop-band) (van Veen et al., 1997; Robinson and Vrba, 1998). The scalar minimum variance beamforming aims to estimate the source power at the location r_k by minimizing the sample variance of the projected data $w^T\mathbf{Y}(t_j)$, $1 \leq j \leq J$ with respect to the weighting vector w, subject to the constraint $w^T\mathbf{x}_k = 1$. In the scalar minimum-variance beamformer, the pass-band is defined by linearly weighting sensor arrays with the constraint $w^T\mathbf{x}_k = 1$, while the stop-band is realized via minimizing the variance of the projected data. The estimated power can be normalized to produce a power map over a given temporal window while the beamformer projected data can provide time course information at each source. We rank these candidate sources by their powers and filter out noisy ones by thresholding.

In recent years, a number of simulation studies and theoretical studies have been conducted to evaluate the performance of a beamformer (e.g., Brookes et al., 2008; Sekihara and Nagarajan, 2010). Despite of this, several issues remain to be addressed. Firstly, there is no rigorous statistical theory available to allow one to examine when the estimated source time-courses are consistent with the true ones. In particular, when there are multiple sources, the accuracy is compromised by confounding effects of multiple sources. It is natural to ask when a beamformer will breakdown in presence of multiple sources and how this effect is determined by the spatial and temporal dimensions of a beamformer. Secondly, it is largely unknown in the literature when the beamformer-based filtering procedure can recover the true sources with an overwhelming probability, although a sure filtering procedure for ordinary linear regression models has been developed by Fan and Lv (2008).

To address these issues, we propose a beamformer filtering procedure which is based on the thresholded sensor covariance estimator. The objective of the procedure is to identify a set of sources from sparse source model (2), which have nonzero powers or power increases relative to a reference. We develop a sure filtering theory for the proposed procedure under certain conditions. We show that if the true sources are not too close to each other and if n and J are large enough, then these sources can be recovered in a probability tending to one. Furthermore, we provide mean error bounds for source localization, power estimation and time-course estimation in the procedure. We conduct simulation studies and a real data analysis to assess the performance of the proposed procedure.

The paper is organized as follows. The details of the new beamforming methodology are provided in Sections 2 and 4. The asymptotic properties of the proposed procedure are investigated in Section 3. The simulation results and a real MEG data analysis are presented in Section 5. The conclusions are made in Section 6. The Proof of Theorem 3.1 is deferred to