Journal of Neuroscience Methods 240 (2015) 170–178

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Autoregressive model in the Lp norm space for EEG analysis

Peiyang Li[a,1], Xurui Wang[a,1], Fali Li[a], Rui Zhang[a], Teng Ma[a], Yueheng Peng[b], Xu Lei[c], Yin Tian[d], Daqing Guo[a], Tiejun Liu[a], Dezhong Yao[a], Peng Xu[a,*]

[a] Key Laboratory for NeuroInformation of the Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China
[b] School of Microelectronics and Solid-State Electronics, University of Electronic Science and Technology of China, Chengdu, China
[c] School of Physiology, Southwest University, Chongqing, China
[d] School of Bioinformatics, Chongqing university of Posts and Telecommunications, Chongqing, China

HIGHLIGHTS

- Designed an Lp ($p \leq 1$) norm-based residual model to estimate autoregressive (AR) parameters.
- The Lp ($p \leq 1$) norm AR model estimates parameters more robustly than an L2 norm-based AR mode for time series with outliers.
- The Lp ($p \leq 1$) norm AR holds a lower relative error of AR parameters and higher prediction accuracy than L2 norm-based methods.
- A resting EEG power spectrum estimated by the Lp ($p \leq 1$) norm AR model is less influenced by ocular artifacts compared with L2 norm-based AR.

ARTICLE INFO

Article history:
Received 6 September 2014
Received in revised form 8 November 2014
Accepted 10 November 2014
Available online 18 November 2014

Keywords:
Autoregressive model
Lp norm
EEG
Power spectrum

ABSTRACT

The autoregressive (AR) model is widely used in electroencephalogram (EEG) analyses such as waveform fitting, spectrum estimation, and system identification. In real applications, EEGs are inevitably contaminated with unexpected outlier artifacts, and this must be overcome. However, most of the current AR models are based on the L2 norm structure, which exaggerates the outlier effect due to the square property of the L2 norm. In this paper, a novel AR object function is constructed in the Lp ($p \leq 1$) norm space with the aim to compress the outlier effects on EEG analysis, and a fast iteration procedure is developed to solve this new AR model. The quantitative evaluation using simulated EEGs with outliers proves that the proposed Lp ($p \leq 1$) AR can estimate the AR parameters more robustly than the Yule–Walker, Burg and LS methods, under various simulated outlier conditions. The actual application to the resting EEG recording with ocular artifacts also demonstrates that Lp ($p \leq 1$) AR can effectively address the outliers and recover a resting EEG power spectrum that is more consistent with its physiological basis.

1. Introduction

Power spectral density (PSD) and variable states are the two important measures for characterizing the physiological information underlying EEGs. There are two main approaches to estimating these measurements, nonparametric and parametric methods. Nonparametric methods (Antoniou, 2006), such as Fourier transforms and periodograms, use the observed data to directly perform the estimation. However, this type of approach is usually problematic because of leakage and frequency resolution in PSD estimates, which require a large number of samples. In the parametric approach, a random signal is characterized by the parameters estimated from the finite record of the data, so there is no need to make assumption about how the data were generated. Currently, three types of parametric models, the autoregressive (AR) model, the moving average (MA) model and the autoregressive moving average (ARMA) model, are used for related estimations (Antoniou, 2006), among which the AR model is by far the most widely used due to the following merits. First, with a suitable order, it can approximate any stationary random process. Second, the AR model is suitable for representing spectra with narrow peaks. Third, the AR model has a set of very simple linear equations for parameter estimation so that many efficient algorithms are available (Jain and Dandapat, 2005). Compared with AR, both the MA and ARMA models, as a general rule, require more coefficients to represent the

* Corresponding author. Tel.: +86 028 83206978.
E-mail address: xupeng@uestc.edu.cn (P. Xu).
[1] Both the authors equally contribute to this work.

http://dx.doi.org/10.1016/j.jneumeth.2014.11.007
0165-0270/

signal spectrum information. In various studies, AR has been proven to be able to build the more meaningful spectrum information of EEG than the FFT-based analysis (Güler et al., 2001).

In recent years, there have been a large number of applications using AR-related models in various EEG-related studies such as EEG system identification, EEG power spectrum estimation, EEG network analysis, and brain computer interfaces (Chen et al., 2013; Qiu, 2011; Wang et al., 2014b). Wang et al. (2014b) proposed a corrected approach to remove the ocular artifact using AR-based system identification, resulting in substantial performance improvement. In Chen et al. (2013), AR is used to extract the phase and frequency information from intracranial EEGs. Dahal et al. (2014) used a time varying AR model to extract features to delineate the attention and distraction in the motor driving experiment. Wang et al. (2010) developed an improved feature extraction method based on the multivariate adaptive autoregressive (MVAAR) model for the classification of motor imagery.

In real-world applications, EEGs will usually be contaminated with outliers due to eye blinks or head movement, which will greatly influence the AR estimation. Most of the previous work in this area paid less attention to the outlier effect, even though the adopted schemes such as moving averaging and sparse constraint may actually compress the outlier effect to some degree (Songsiri, 2013). Theoretically, the previously used AR model and its variants are based on the L2 norm structure, and the L2 norm will exaggerate the outlier effect due to the square property of the L2 norm (Blankertz et al., 2007; Songsiri, 2013; Xu et al., 2010). Compared with the L2 norm, the Lp ($p \leq 1$) norm has been proven to be robust to outliers, and it has been widely used in a diversity of signal processing applications such as denoising, EEG inverse problem, MRI/CT reconstruction, feature extraction (Chartrand, 2009; Li et al., 2013; Lustig et al., 2007; Xu et al., 2007). In consideration of the merits of the Lp ($p \leq 1$) norm to compress the outlier effect, we will restructure the residual equation for AR parameter estimation in the Lp ($p \leq 1$) norm space and establish a fast iteration procedure to solve this new AR model.

2. Materials and methods

2.1. Autoregressive model

The AR model is usually expressed as:

$$x(n) = -\sum_{k=1}^{q} w_k x(n-k) + u(n) \tag{1}$$

where $u(n)$ is the input sequence to the system and is usually considered to be zero-mean white Gaussian noise with a variance σ_w^2. $x(n)$ is the observed data, representing the output sequence. $\{w_k, 1 \leq k \leq q\}$ is the corresponding AR parameters with q being the order of the AR model. The system transfer function is given by

$$H(z) = \frac{B(z)}{C(z)} = \frac{1}{1 + \sum_{k=1}^{q} w_k z^{-k}} \tag{2}$$

where the $C(z)$ and $B(z)$ represent the poles and zeros of the system response, respectively. Based on the transfer function, the AR based spectrum estimation at frequency f has the form

$$P_{xx}(f) = \sigma_w^2 |H(f)|^2 = \frac{\sigma_w^2}{\left|1 + \sum_{k=1}^{q} w_k \, e^{-2\pi k f}\right|^2} \tag{3}$$

In essence, the AR model minimizes the residual errors for all the observed samples:

$$E[u_n^2] = \sum_{i=1}^{N} u_n^2 = \sum_{i=1}^{N} \left| x(i) - \sum_{k=1}^{q} w_k x(i-k) \right|^2 \tag{4}$$

Let $W = [w_1, w_2, \ldots, w_q]^T$; $Y = [x(q+1), x(q+2), \ldots, x(N)]^T$, with N being the length of signal; $||\bullet||_2$ denotes the L2 norm of a matrix or a vector; and $A \in R^{(N-q)\times q}$ be the delay array:

$$A = \begin{bmatrix} x(q) & x(q-1) & \cdots & x(1) \\ x(q+1) & x(q) & \cdots & x(2) \\ \vdots & \vdots & & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-q) \end{bmatrix} \tag{5}$$

Eq. (4) can be formatted as

$$\arg\min_{W} f(W) = ||Y - AW||_2^2 \tag{6}$$

By taking the derivative of (6) with respect to W under the condition $df/dw = 0$, we can obtain the follow formulation:

$$2A^T AW - 2A^T Y = 0 \tag{7}$$

and the objective parameters W can be estimated as

$$W = (A^T A)^{-1} A^T Y \tag{8}$$

In addition to the least square algorithm, other approaches such as the Yule–Walker (Y-W) equations and the Burg method (Antoniou, 2006) are also used to estimate the AR parameters that can minimize the sum of the residual errors in (4). No matter what scheme is used to estimate the AR parameters, the inherent L2 norm structure in (4) and (6) indicates that the influence of outliers will be exaggerated due to the square property of the L2 norm, resulting in the biased AR parameters.

2.2. Lp ($p \leq 1$) norm based autoregressive model

In real-world applications, outliers will create an unexpected effect on related analyses such as spectrum estimation, signal prediction. To improve the robustness of the AR parameters estimation, some schemes such as sparse constraint with Lp ($p \leq 1$) norm terms are proposed in various AR variant versions to alleviate the noise effect (Ping-bo and Zhi-ming, 2006). However, most of them mainly focus on the imposing restrictions on the parameters, leaving the main structure of the objective function in the L2 norm space. Unfortunately, the L2 norm object function will inevitably exaggerate the outlier effect no matter how the AR parameters are emphasized. We will define the AR object function in the Lp ($p \leq 1$) norm space, aiming to improve the AR robustness to the outlier effect.

The AR object function is defined in Lp ($p \leq 1$) norm space as

$$\begin{aligned} W^* &= \arg\min_{W} f^*(W) = \arg\min_{W} ||Y - AW||_p^p \\ &= \arg\min_{W} \sum_{i=1}^{N-q} |x_{q+i} - A(i,:)W|^p \end{aligned} \tag{9}$$

where $||\bullet||_p$ denotes the Lp ($p \leq 1$) norm of a vector; we refer to this model as the Lp-AR estimation. The gradient for this function is

$$g = p \sum_{i=1}^{n-q} |x_{q+i} - A(i,:)W|^{p-1} sgn(i)(-A^T(i,:)) \tag{10}$$