Statistics and Probability Letters 129 (2017) 120–130

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Robust parameter estimation for stationary processes by an exotic disparity from prediction problem

Yan Liu

Department of Applied Mathematics, Waseda University, Tokyo 169-8555, Japan

ARTICLE INFO

Article history:
Received 7 June 2016
Received in revised form 2 May 2017
Accepted 11 May 2017
Available online 25 May 2017

MSC 2010:
62M10
62F12
62F35

Keywords:
Stationary process
Spectral density
Minimum contrast estimation
Prediction problem
Asymptotic efficiency

ABSTRACT

A new class of disparities from the point of view of prediction problem is proposed for minimum contrast estimation of spectral densities of stationary processes. We investigate asymptotic properties of the minimum contrast estimators based on the new disparities for stationary processes with both finite and infinite variance innovations. The relative efficiency and the robustness against randomly missing observations are shown in our numerical simulations.

1. Introduction

Methods of fitting parametric models to linear time series have been investigated for a long time. One method for parameter estimation is to minimize a certain disparity measure $D(f_\theta, \hat{g}_n)$ between the spectral density f_θ and some nonparametric estimator $\hat{g}_n$ of f_θ. Two disparity measures, the location disparity and the scale disparity, have been mainly considered so far. The *location* disparity $D(f_\theta, \hat{g}_n) = \int_{-\pi}^{\pi} [\Phi(f_\theta(\omega))^2 - 2\Phi(f_\theta(\omega))\Phi(\hat{g}_n(\omega))]d\omega$ with a bijective function $\Phi(\cdot)$ was proposed in Taniguchi (1981). The *scale* disparity $D(f_\theta, \hat{g}_n) = \int_{-\pi}^{\pi} K(f_\theta(\omega)/\hat{g}_n(\omega))d\omega$ with a sufficiently smooth contrast function $K(\cdot)$ was proposed in Taniguchi (1987). Both methods yield consistent parameter estimation.

In this paper, we propose a new consistent disparity as follows.

$$D(f_\theta, \hat{g}_n) = \int_{-\pi}^{\pi} a(\theta) f_\theta^{\alpha}(\omega) \hat{g}_n(\omega) d\omega, \tag{1}$$

where $\alpha \in \mathbb{R} \setminus \{0\}$ and $a(\theta)$ is

$$a(\theta) = \begin{cases} C \cdot \left(\int_{-\pi}^{\pi} f_\theta^{\alpha+1}(\omega) d\omega \right)^{-\frac{\alpha}{\alpha+1}}, & \text{if } \alpha \neq -1, \\ C, & \text{if } \alpha = -1. \end{cases}$$

Here, C is a nonzero generic constant. The disparity (1) originates from minimizing prediction error in L^p of prediction problem of X_0 based on $\{X_t;\ t \in \mathbb{Z} \setminus \{0\}\}$ for stationary processes. Our new disparity is neither included in the class of location

E-mail address: yan.liu@aoni.waseda.jp.

http://dx.doi.org/10.1016/j.spl.2017.05.005

disparities nor scale disparities. The form of the disparity is similar to the power divergence proposed in Renyi (1961), Csiszár (1975) and Fujisawa and Eguchi (2008) for density estimation of independent and identically distributed random variables. As pointed out in Fujisawa and Eguchi (2008), the disparity has the robustness to outliers and contamination under the heavy contaminated models. Under regularity conditions for parameter estimation of stationary processes, we show that the estimator based on the disparity (1) has properties of consistency and asymptotic normality. The finite sample properties of the estimator are investigated in our numerical simulations. To investigate the robustness under contaminated models in simulations, the observation is supposed to be randomly missing from a stationary time series. It is shown that the mean squared error for $\alpha < -1$ is smaller than that for $\alpha = -1$ in both cases when i.i.d. innovations are Cauchy and Gaussian distributed although the case $\alpha = -1$ is theoretically the most efficient if we have full observations.

The paper is organized as follows. In Section 2, we derive the disparity (1) from the prediction problem in L^p. Asymptotic properties of the proposed estimator for stationary processes with both finite and infinite variance innovations are investigated in Section 3. In Section 4, we investigate our new disparity (1) in numerical simulations. Especially, we apply our new disparity (1) to the irregularly observed stationary processes as an application for robust parameter estimation.

2. A new class of disparities from prediction problem

In this section, we derive our disparity (1) from the prediction problem of stationary processes in L^p. Let $\mathbb{Z}$ denote the set of all integers. Suppose $\{X(t), t \in \mathbb{Z}\}$ is a real-valued stationary process with spectral density $g(\omega)$. Let $\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$ and $\mathcal{M}$ denote the closed linear manifold generated by $\{e^{ij\omega}, j \in \mathbb{Z}_0\}$. The linear prediction of X_0 based on $\{X_t, t \in \mathbb{Z}_0\}$ in L^p is the problem of minimizing the prediction error in L^p on $\mathcal{M}$ for $1 < p < \infty$, that is,

$$\inf_{\phi \in \mathcal{M}} \int_{-\pi}^{\pi} |1 - \phi(\omega)|^p g(\omega) d\omega. \tag{2}$$

For example, interpolation problem for time point 0 is formulated by (2) with $p = 2$ (cf. Grenander and Rosenblatt, 1957; Rosenblatt, 2000). As shown in Miamee and Pourahmadi (1988), the best predictor $\phi(\omega)$ is given by

$$\phi(\omega) = 1 - \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} g^{\alpha+1}(\omega) d\omega\right)^{-1} g^{\alpha+1}(\omega),$$

where $\alpha = -p/(p-1)$. In practice, it is usually difficult to know the true density $g(\omega)$ a priori. Suppose we fit a parametrized density $f_\theta(\omega)$ with d-dimensional parameter $\theta \in \Theta \subset \mathbb{R}^d$ to the true spectral density for the prediction problem. Then, noting that $p = \alpha/(\alpha + 1)$, the error in L^p of the prediction by the best predictor $\phi(\omega)$ for $f_\theta(\omega)$ is

$$\begin{aligned} &\int_{-\pi}^{\pi} \left| \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f_\theta^{\alpha+1}(\omega) d\omega\right)^{-1} f_\theta^{\alpha+1}(\omega) \right|^p g(\omega) d\omega \\ &= \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f_\theta^{\alpha+1}(\omega) d\omega\right)^{-\frac{\alpha}{\alpha+1}} \left(\int_{-\pi}^{\pi} f_\theta^{\alpha}(\omega) g(\omega) d\omega\right). \end{aligned} \tag{3}$$

Motivated by (3), we consider an estimation procedure to estimate the parameter θ by the following disparity ($\alpha \in \mathbb{R} \setminus \{0\}$),

$$D(f_\theta, g) = \int_{-\pi}^{\pi} a(\theta) f_\theta^{\alpha}(\omega) g(\omega) d\omega, \tag{4}$$

where $a(\theta)$ is

$$a(\theta) = \begin{cases} C \cdot \left(\int_{-\pi}^{\pi} f_\theta^{\alpha+1}(\omega) d\omega\right)^{-\frac{\alpha}{\alpha+1}}, & \text{if } \alpha \neq -1, \\ C, & \text{if } \alpha = -1, \end{cases}$$

where C is a nonzero generic constant. This notation is used hereafter. We call (4) the *exotic* disparity. Note that when $\alpha = 0$, (4) do not give a disparity between f_θ and g, and further, (2) cannot be regarded as the prediction problem since $p = 0$.

Next, let us investigate the basic property of the exotic disparity. This disparity is not included in the class of either location disparities or scale disparities since the powers of parametrized spectral density and true density are not same in the exotic disparity (4). However, the definition of the disparity can be motivated by the following two examples: (i) the Whittle disparity when $\alpha = -1$; (ii) the estimation procedure minimizing the interpolation error when $\alpha = -2$. The comparison of the efficiency between these two methods are considered in Suto et al. (2016). In addition to the efficiency, the robustness against randomly missing observations by the exotic disparity (4) is also considered in Section 4 in this paper.

In the following, we impose Assumption 1.

Assumption 1.

(i) The parameter space Θ is a compact subset of $\mathbb{R}^d$.
(ii) If $\theta_1 \neq \theta_2$, then $f_{\theta_1} \neq f_{\theta_2}$ on a set of positive Lebesgue measure.
(iii) The parametric spectral density $f_\theta(\omega)$ is three times continuously differentiable with respect to θ and the second derivative $\frac{\partial^2}{\partial\theta\partial\theta^T} f_\theta(\omega)$ is continuous in ω.