Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 62 (2018) 683–691

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Modified kernel regression method for the denoising of X-ray pulsar profiles

Jianing Song [a], Jiawei Qu [b], Guodong Xu [a,*]

[a] Research Centre of Satellite Technology, Harbin Institute of Technology, Harbin 150001, China
[b] New H3C Group, Hangzhou 310052, China

Received 2 October 2017; received in revised form 16 April 2018; accepted 18 April 2018
Available online 3 May 2018

Abstract

In this paper, a modified kernel regression algorithm is proposed to reduce the noise of pulsar profiles autonomously. Taking advantage of the classical autonomous kernel regression the presented algorithm based on the second-order derivative compensation is developed to improve the performance of the Nadaraya-Watson kernel estimator. The periodic extension technique is introduced to eliminate the boundary issue inherent in kernel regression means. Four indexes are utilized to explore the performance of the proposed method via both emulated and real data. Additionally, other widely accepted denoising methods based on wavelet transformation and empirical model decomposition are simulated to make a comparison. The experimental results have shown that the proposed algorithm achieves a higher quality profile than the compared methods, which will help to discover the emission mechanism of pulsars. According to our experiments, we would like to point out that a smooth profile cannot guarantee an accurate measurement (time of arrival, TOA), which indicates that it is not necessary to denoise the epoch folding profile for estimating the TOA information in X-ray pulsar-based navigation systems.

Keywords: Signal denoising; X-ray pulsar photon; Kernel regression; Profile recovery

1. Introduction

Pulsar is a type of rapidly-rotating neutron star that emits radiation at varying intensities. The radio emission is continuous through its magnetic axis, which is in a different direction to its rotation axis. Thus, an observer sees a pulse of radiation each time the beam sweeps across his line-of-sight (Lorimer and Kramer, 2012). Since the pulse period equals the rotation period of the spinning neutron star, the period of pulsar signals is quite stable. With the benefit of these unique characteristics, X-ray pulsars are valuable in both scientific research and engineering applications. Scientists believe that X-ray pulsars provide most effective tests of general relativity (Antoniadis et al., 2013) and can be used to probe relativistic effects (Lyne et al., 2004; Taylor et al., 1979). Additionally, X-ray pulsars provide new possible navigation and timing schemes for spacecraft autonomous navigation (Sheikh et al., 2006; Emadzadeh and Speyer, 2011b).

The fundamental observation of X-ray pulsars in spacecraft is X-ray photons. Afterwards, an empirical profile of a pulsar is recovered from a series of photon time of arrivals (TOAs) for future analysis. These pulsar profiles have long been recognized as an important clue to lead a better understanding of the pulsar phenomenon in astronomy (Gil et al., 1993; Hakobyan et al., 2017). In addition, the measurement of time of arrival (TOA) can be obtained by comparing the empirical profile with a standard one, which helps to determine the position of spacecraft

* Corresponding author.
E-mail addresses: hitsjn@hit.edu.cn (J. Song), xgdong_61@163.com (G. Xu).

https://doi.org/10.1016/j.asr.2018.04.027
0273-1177/

autonomously in X-ray pulsar-based navigation system. Nevertheless, the signal-to-noise ratio (SNR) of the observed profile is very low due to the faint source flux density. A high SNR profile is important to study the emission mechanisms of pulsars (Gotthelf et al., 1999; Wang et al., 2017a,b) and to estimate an accurate TOA by intuition. Therefore, it is necessary to develop noise reduction methods of pulsar profiles.

Previous literature has demonstrated pulsar profile denoising algorithms based on wavelet transformation and empirical model decomposition. These methods can be seen as the parametric regression method (Wang et al., 2017b) which the performance heavily relies on the accuracy of the assumed model and sometimes needs human visual inspection. Thus, Wang et al. (2017b) first introduced the kernel regression method for addressing the denoising issue of pulsar profiles autonomously. Instead of involving a specific model, kernel regression is a non-parametric method based on the data itself (Takeda et al., 2007) and kernel functions are used as weights to estimate a given point from all measurements.

The concept of kernel methods appeared during the mid-twenty century (Nadaraya, 1964; Watson, 1964) and then became well-known with the emergence of machine learning. Recently, two-dimensional kernel regression has become popular in image processing (Takeda et al., 2007; Zhang et al., 2013). While there are two major challenges using kernel regression technique to signal denoising problems: (1) the bandwidth selection of kernel function (2) low performance near the boundary (Wand and Jones, 1994). The bandwidth is a positive number in kernel functions and the choose of it is a famous trade-off between the bias and variance of the estimator (Köhler et al., 2014; Wand and Jones, 1994). A kernel regression method performs badly or even fails if the bandwidth is selected improperly. The boundary is the variable which is near the end of the support. In the boundary region, the estimated variance and bias become worse because of lacking observations (Baszczyńska, 2016). In order to improve the performance near the boundary, local polynomial kernel estimators (LPKEs) are proposed, which a particular point is estimated by "locally" fitting a pth degree polynomial (Wand and Jones, 1994; Cleveland, 1979). Comparing with the Nadaraya-Watson kernel regression (NWKR) estimator (Wand and Jones, 1994), the best fitting parameters of LPKEs are first given by a weighted least square problem in each estimated point. At the cost of considerable computation, LPKEs have attractive results and boundary properties. To the authors' knowledge, though previous work has studied pulsar profile denoising based on kernel regression methods, neither of above challenges has been fully presented.

This paper addresses the two challenges in pulsar profile denoising issues. A modified kernel regression algorithm using the second-derivative compensation is studied to reduce the bias of regression functions. The boundary effect is eliminated by periodic extension with small computation. Four indexes, including SNR, Pearson correction coefficient (PCC), mean and standard deviation of TOA, are utilized to evaluate the performance of the proposed method and to study whether the accuracy of TOA and the SNR of empirical profiles has a positive correlation. Through simulation experiments, the proposed method outperforms current pulsar profile denoising methods such as wavelet-based methods and empirical model decomposition methods, as well as has better robustness on bandwidth selection than the NWKR method. Moreover, our experimental results suggest that a clear positive relationship between the accuracy of TOA and the SNR of profiles does not exist.

The remainder of the paper is organized as follows. Section 2 introduces the classical kernel regression technique in one-dimension. Section 3 derives a modified kernel regression method based on the properties of pulsar profiles. The performance of the proposed method is analyzed via computer simulations in Section 4 and concluding remarks are discussed in Section 5.

2. Related works

2.1. Classical kernel regression and its properties

Considering a nonlinear model is

$$y_i = m(x_i) + \varepsilon(x_i), \qquad i = 1, 2, \ldots, n \tag{1}$$

where $m(\cdot)$ is called the regression function, $\varepsilon(x_i)$ is the noise that obeys Gaussian distribution, and x_i is the measurement epoch corresponding to measurement y_i.

For LPKEs, the regression function $\hat{m}(x; p, h)$ at the point x is obtained by fitting a pth order polynomial. The coefficient of the polynomial $\{\beta_n\}$ can be solved by the following optimization problem (Wand and Jones, 1994)

$$\min_{\beta_n} \sum_{i=1}^{n} \{y_i - \beta_0 - \cdots - \beta_p (x_i - x)^p\}^2 K_h(x_i - x) \tag{2}$$

where $K_h(\cdot)$ is the kernel function and symmetric, h is the bandwidth of $K_h(\cdot)$. Since the shape of the kernel function has a small effect on the estimating result (Wand and Jones, 1994). The widely used Gaussian kernel is selected in this paper, giving by (Wand and Jones, 1994)

$$K_h(t) = \frac{1}{\sqrt{2\pi}h} \exp\left(-\frac{t^2}{2h^2}\right) \tag{3}$$

By integration operating, the Gaussian kernel $K_h(\cdot)$ satisfies

$$\int K_h(t)dt = 1, \quad \int tK_h(t)dt = 0, \quad \int t^2 K_h(t)dt = h^2 \tag{4}$$

After obtaining $\{\beta_n\}$, the regression function $\hat{m}(x; p, h)$ based on LPKEs can be expressed as (Wand and Jones, 1994)

$$\hat{m}(x; p, h) = \mathbf{e}_1 (\mathbf{X}_x^T \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^T \mathbf{W}_x \mathbf{Y} \tag{5}$$