Journal of Applied Geophysics 112 (2015) 236–241

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Adaptive steering kernel regression for prestack seismogram denoising

Yanan Tian [a,b], Yue Li [b,*], Xuechun Xu [a], Ning Wu [b]

[a] College of Earth Sciences, Jilin University, Changchun, China
[b] Department of Communication and Engineering, Jilin University, Changchun, China

ARTICLE INFO

Article history:
Received 17 June 2014
Accepted 24 November 2014
Available online 2 December 2014

Keywords:
Kernel regression (KR)
Seismic denoising
Nonlinear filtering
Adaptive kernel

ABSTRACT

Noise attenuation is a necessary and persistent problem during prestack seismic data processing. In this paper, we discuss a developed nonlinear method called adaptive steering kernel regression (ASKR) and apply it to seismic random noise attenuation. In classical kernel regression (KR), spatial distance is considered as the only determinant of regression function weights. A kernel with fixed shape is used for all samples, which results in severe distortion of edge information. In the discussed ASKR, the weights are estimated based upon another important determinant — gray distance. The shape of kernel varies with the characteristics of data samples in different regions. Thus, the edge information, which corresponds to reflection events in seismic records, is preserved more effectively and completely. Results on both synthetic records and real seismic data show its feasibility and effectiveness. Moreover, we verify its better performance than classical KR in the aspect of amplitude preservation and noise attenuation.

1. Introduction

Random noise in seismic exploration not only corrupts the effective signals, but also degrades the quality of seismic records, so that it is difficult to obtain desirable signals from collected seismic data. Unlike coherent noise, random noise is incoherent in time, and most of the time unpredictable. One of the main tasks in seismic denoising is to eliminate random noise and useless components to enhance signal-to-noise ratio (SNR) while preserving or recovering important seismic such as signal amplitudes and discontinuous contours or edges (Yilmaz, 2001).

So far, many kinds of denoising methods have been proposed for seismic random noise attenuation, such as median filtering (Liu and Liu, 2009), curvelet transform (Awasthi and Singh, 2012; Kumar and Oueity, 2011 and Shan and Fu, 2009), Wiener filtering (Chen and Benesty, 2006; Levent and Arslan, 2006 and Sekko and Boukrouche, 1999), independent component analysis (ICA) (Liu and Liu, 2003; Lu, 2006 and Thomas and Deville, 2006). Although these methods have received great popularity, some shortcomings still exist in some cases. For example, Wiener filtering is an effective method to estimate desired signals from observations. But its performance fails when the statistical properties of desired signal or random noise are unknown. ICA is another method for signal extraction. However, if signal and noise do not satisfy the model or assumptions, the filtering result becomes unsatisfactory resulting in severe signal distortion and detailed feature loss.

We note that “denoise” is a special case of regression problem, where samples at all desired sample locations are given, but corrupted. Kernel regression is a well established technique, which is widely used in image processing and reconstruction. It is one of non-parametric methods, which generally relies on the observed data and existing information to dictate the parameters of desirable signal model.

In classical kernel regression (KR), the target value of a test input is computed as a weighted average of function values observed at the inputs. The weights of each input are computed using a kernel function and determined only by spatial distance (Wand and Jones, 1995). In contrast, the discussed adaptive steering kernel regression (ASKR) takes gray distance, i.e. radiometric distance between two pixels, into account and considers it as another determinant of the weights. Thus, the size and shape of regression kernel are adaptive according to different sample features. In this way, the edge information, which corresponds to reflection events in seismic records, can be preserved more completely.

The main structure of this paper is managed as follows. First, the basic theories of classical KR and some corresponding concepts are described briefly. Then, we illustrate the ASKR method with corresponding formula derivations and explanations. Finally, both synthetic records and real seismic data are used to verify its feasibility and better effectiveness.

2. Classical kernel regression (KR)

The model of a 2D classical kernel regression (KR) can be given by

$$y_i = z(\mathbf{x}_i) + \varepsilon_i,\ i = 1, 2, \cdots, P, \tag{1}$$

where $z(\cdot)$ is the regression function, y_i is the observed signal, and $\mathbf{x}_i = [x_{1i}, x_{2i}]^T$ is the coordinate of y_i. ε_is are independent and identically

* Corresponding author at: Nanhu Road, Changchun 130012, China.
E-mail address: liyue@jlu.edu.cn (Y. Li).

http://dx.doi.org/10.1016/j.jappgeo.2014.11.020
0926-9851/

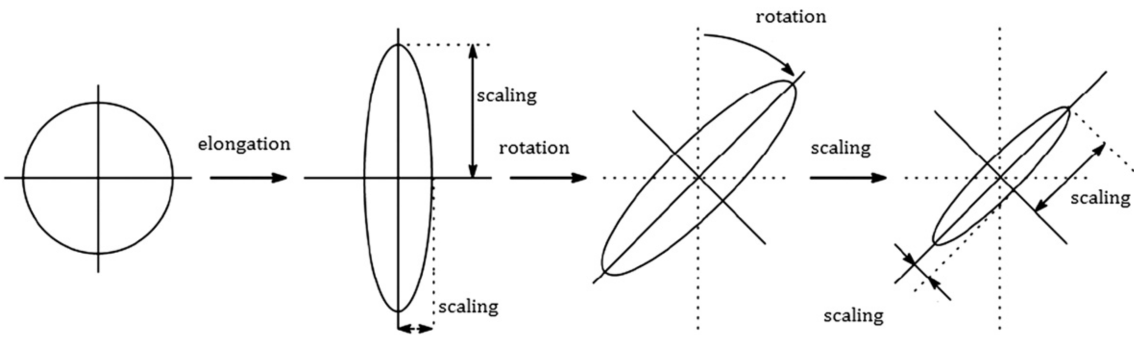

Fig. 1. Schematic illustration of the effects of steering matrix.

distributed estimation errors. As the specific form of $z(\cdot)$ is unspecified, we take the N-term Taylor expansion as its local expansion at point $\mathbf{x}_i$. The order N = 1, 2, 3 correspond to constant, linear and quadratic, respectively (Gijbels and Lambert, 2006; Ruppert and Wand, 1994 and Takeda and Farsiu, 2007). The 2-order Taylor expansion series can be expressed as

$$z(\mathbf{x}_i) = \beta_0 + \boldsymbol{\beta}_1^T(\mathbf{x}_i - \mathbf{x}) + \boldsymbol{\beta}_2^T \mathrm{vech}\left\{(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T\right\}, \tag{2}$$

where vech$(\cdot)$ denotes half-vector operator, which orders a matrix into a vector. And $\boldsymbol{\beta}_n(n = 0, 1, 2)$ is the regression function weights, which can be expressed as

$$\beta_0 = z(\mathbf{x}), \tag{3}$$

$$\boldsymbol{\beta}_1 = \nabla z(\mathbf{x}) = \left[\frac{\partial z(\mathbf{x})}{\partial x_1}, \frac{\partial z(\mathbf{x})}{\partial x_2}\right]^T, \tag{4}$$

$$\boldsymbol{\beta}_2 = \frac{1}{2}\left[\frac{\partial^2 z(\mathbf{x})}{\partial x_1^2}, 2\frac{\partial^2 z(\mathbf{x})}{\partial x_1 \partial x_2}, \frac{\partial^2 z(\mathbf{x})}{\partial x_2^2}\right]^T \tag{5}$$

where $\mathbf{x} = [x_1, x_2]^T$ and ∇ denotes the gradient operator. $\boldsymbol{\beta}_n(n = 0, 1, 2)$ can be computed from the following optimization problem

$$\min_{\{\beta_n\}} \sum_{i=1}^{P} [y_i - z(\mathbf{x}_i)]^2 K_\mathbf{H}(\mathbf{x}_i - \mathbf{x}), \tag{6}$$

where

$$K_\mathbf{H}(\mathbf{x}_i - \mathbf{x}) = \frac{1}{\det(\mathbf{H})} K\left(\mathbf{H}^{-1}(\mathbf{x}_i - \mathbf{x})\right), \tag{7}$$

where K is a 2D kernel function. $\mathbf{H}$ is a 2×2 smoothing matrix, which can be expressed as a simple and efficient model as

$$\mathbf{H}_i = h\mu_i \mathbf{I}, \tag{8}$$

where μ_i is the local density parameter, which is usually set to $\mu_i = 1$. h is the global smoothing parameter. A large h has a strong denoising effect (Chu and Marron, 1991 and Hardle and Vieu, 1992). $\mathbf{I}$ is an identity matrix. Eq. (6) is a weighted least-square optimization problem. So the weights solution can be solved in matrix form as

$$\hat{\beta}_0 = \mathbf{e}_1^T\left(\mathbf{X}_\mathbf{x}^T \mathbf{W}_\mathbf{x} \mathbf{X}_\mathbf{x}\right)^{-1} \mathbf{X}_\mathbf{x}^T \mathbf{X}_\mathbf{x} \mathbf{y}, \tag{9}$$

where $\mathbf{y}$ is a column vector composed of y_is and

$$\mathbf{W}_\mathbf{x} = \mathrm{diag}[K_\mathbf{H}(\mathbf{x}_1 - \mathbf{x}), K_\mathbf{H}(\mathbf{x}_2 - \mathbf{x}), \cdots, K_\mathbf{H}(\mathbf{x}_P - \mathbf{x})],$$
$$\mathbf{X}_\mathbf{x} = \begin{bmatrix} 1 & (\mathbf{x}_1 - \mathbf{x})^T & vech^T\left\{(\mathbf{x}_1 - \mathbf{x})(\mathbf{x}_1 - \mathbf{x})^T\right\} \\ 1 & (\mathbf{x}_2 - \mathbf{x})^T & vech^T\left\{(\mathbf{x}_2 - \mathbf{x})(\mathbf{x}_2 - \mathbf{x})^T\right\} \\ \vdots & \vdots & \vdots \\ 1 & (\mathbf{x}_P - \mathbf{x})^T & vech^T\left\{(\mathbf{x}_P - \mathbf{x})(\mathbf{x}_P - \mathbf{x})^T\right\} \end{bmatrix}.$$

Then $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ can be obtained by substituting Eq. (9) into Eqs. (4) and (5).

3. Adaptive steering kernel regression (ASKR)

From Eq. (6) we know that in classical KR, the weights of regression function only depend on the spatial distance between $\mathbf{x}_i$ and $\mathbf{x}$. In this section, let us discuss a nonlinear ASKR method that enables its structure to have nonlinear effect on data. In the discussed ASKR, the weights are estimated based on spatial distance and radiometric distance between two pixels, i.e. gray distance (Cho and Koschan, 2013; Krishnan

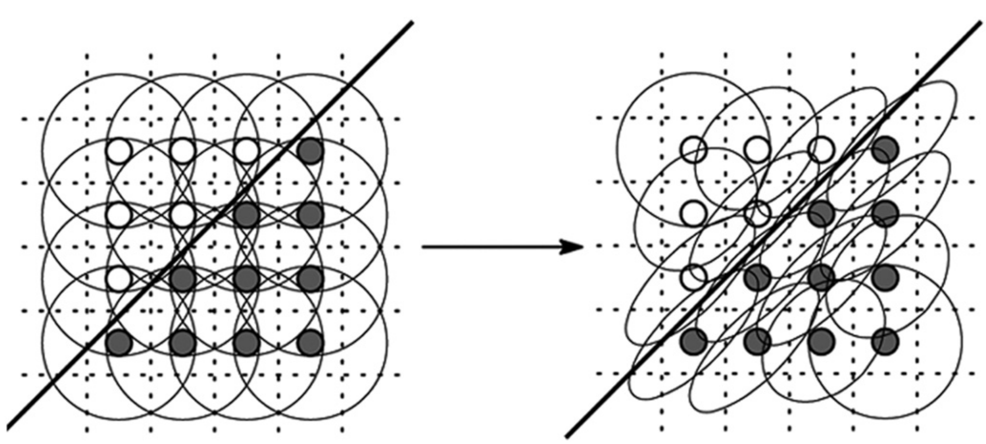

Fig. 2. Shape of kernels. a) classical kernel and b) adaptive kernel.