Nonparametric Wavelet Regression with Correlated Errors

Porto, Rogério
Bank of Brazil
SBS Qd. 1 Bl. C Lote 32 Ed. Sede III 10 andar
70073-901 Brasília, DF, Brazil
E-mail: rogerio@bb.com.br

Morettin, Pedro
University of São Paulo, Institute of Mathematics and Statistics
Rua do Matão 1010
05508-090 São Paulo, SP, Brazil
E-mail: pam@ime.usp.br

Percival, Donald
University of Washington, Applied Physics Laboratory and Department of Statistics
1013 NE 40th Street
98105-6698 Seattle, WA, USAC
E-mail: dbp@apl.washington.edu

Aubin, Elisete
University of São Paulo, Institute of Mathematics and Statistics
Rua do Matão 1010
05508-090 São Paulo, SP, Brazil
E-mail: aubin@ime.usp.br

Introduction

A mathematical problem of considerable interest is to approximate a continuous function $f(t)$, $t \in [0,1]$, based upon samples $f(t_i)$, $i = 1, \ldots, n$. We do not observe $f(t_i)$ directly, but only in the presence of correlated zero mean noise $\{\epsilon(t_1), \ldots, \epsilon(t_n)\}$, which we assume throughout to obey a multivariate Gaussian distribution. The data are $\{(t_1, y(t_1)), \ldots, (t_n, y(t_n))\}$, where $y(t_i) = f(t_i) + \epsilon(t_i)$, for $i = 1, \ldots, n$, and our objective is to extract the signal f from the data using an estimator $\hat{f}$ with low integrated mean squared error (IMSE), defined as

$$R(\hat{f}, f) = E||\hat{f} - f||_2^2 = \int_0^1 E(\hat{f}(x) - f(x))^2 dx.$$

Wavelet shrinkage methods have been very successful in signal extraction and nonparametric regression, but most methods are focused on equispaced samples (i.e., over a regular grid $t_i = i/n$) with independent and identically distributed (IID) errors. The equispaced assumption has been relaxed to handle unequally spaced samples with a fixed design (Cai and Brown, 1998), a uniformly distributed design (Cai and Brown, 1999) and a general random design (Sardy et al., 1999; Kerkyacharian and Picard, 2004), but these extensions are restricted to IID errors. Wavelet shrinkage methods have also been adapted to handle correlated errors, in the context of equispaced samples (Johnstone and Silverman, 1997) and of unequally spaced samples with a fixed design (Porto et al., 2008).

In this paper, we investigate wavelet shrinkage for certain unequally sampled designs in the presence of correlated errors. The sampling schemes that we consider are stochastic, where either the sample points t_i are uniformly distributed in $[0,1]$ or they come from a jittering; i.e., $t_i = (2i-1)/(2n) + j_i$, where j_i are IID uniform $[-1/(2n), 1/(2n)]$ random variables. Stochastic sampling techniques are of interest because they can overcome certain aliasing problems associated with sampling on a regular grid (Dippé and Wold, 1985). We show that under our assumptions the samples can be treated as if they were equispaced with correlated noise (Johnstone and Silverman, 1997), and hence we can apply the VisuShrink procedure (Donoho and Johnstone, 1994) with level-dependent thresholds.

Wavelets and wavelet shrinkage

An orthonormal wavelet basis is generated from dilation and translation of a "father" wavelet ϕ (or scaling function) and a "mother" wavelet ψ. Let

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k) \text{ and } \psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k).$$

Denote the inner product by $\langle \cdot, \cdot \rangle$. For a given square-integrable function f on $[0, 1]$, let

$$c_{j,k} = \langle f, \phi_{j,k} \rangle \text{ and } d_{j,k} = \langle f, \psi_{j,k} \rangle .$$

The function f can be expanded into a wavelet series as

$$f(x) = \sum_{k=0}^{2^{j_0}-1} c_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(x).$$

An orthonormal wavelet basis has an associated exact orthogonal discrete wavelet transform W that transforms sampled data into discrete wavelet coefficients. Let

$$\begin{aligned} \tilde{\theta} = Wy \quad &= \Big(\tilde{c}_{j_0,0}, \ldots, \tilde{c}_{j_0,2^{j_0}-1}, \tilde{d}_{j_0,0}, \ldots, \tilde{d}_{j_0,2^{j_0}-1}, \ldots, \\ &\quad \tilde{d}_{J-1,0}, \ldots, \tilde{d}_{J-1,2^{J-1}-1}\Big)^T \end{aligned}$$

be the coefficients of the discrete wavelet transform. Define the soft threshold function by

$$\eta_S(d, \lambda) = \text{sgn}(d)(|d| - \lambda)_+,$$

for some threshold λ.

Now suppose that the error vector $e = (e_1, \ldots, e_n)^T$ have a multivariate Gaussian distribution with mean 0 and covariance matrix Γ. Also, assume that the errors are stationary so that Γ has entries $\gamma_{|r-s|}$. Let $z = We$ be the wavelet transform of the error vector. Neglecting boundary effects, within each level $z_{j,k}$ will be a portion of a stationary process with level-dependent variance $\sigma_j^2 = \text{Var}(z_{j,k})$ (Johnstone and Silverman, 1997).

Wavelet shrinkage for random design with correlated errors

Consider a sample $(t_1, y(t_1)), (t_2, y(t_2)), \ldots, (t_n, y(t_n))$ from some stochastic sampling scheme with respective order statistics $0 \le t_{(1)} < t_{(2)} < \cdots < t_{(n)} \le 1$ that satisfy

(1) $$\text{Var}\left(t_{(i)}\right) \le \frac{1}{n} \text{ and } \left| E\left(t_{(i)}\right) - \frac{i}{n} \right| \le \frac{1}{\sqrt{n}}$$

for $i = 1, \ldots, n$. Given the data, assume the model

(2) $$y_i = f(t_{(i)}) + e_i,$$

where $y_i \equiv y(t_{(i)})$ and the errors $e_i = e(t_{(i)})$ are such that

(3) $$\text{Cov}\left(e(t_{(i)}), e(t_{(j)})\right) \le \gamma(|i - j|)$$

and

(4) $$\lim_{n\to\infty} \sum_{u=-(n-1)}^{n-1} |\gamma(u)| < \infty.$$

Let $\hat{f}(t)$ be the estimator of $f(t)$ for all $t \in [0,1]$, where

$$(5) \qquad \hat{f}(t) = \sum_{k=0}^{2^{j_0}-1} \hat{c}_{j_0,k}\phi_{j_0,k}(t) + \sum_{j=j_0}^{J'-1}\sum_{k=0}^{2^j-1} \hat{d}_{j,k}\psi_{j,k}(t);$$

$\hat{d}_{j,k}$ is given by

$$(6) \qquad \hat{d}_{j,k} = \eta_S(\tilde{d}_{j,k}, \lambda_j),$$

where $\lambda_j = \sigma_j\sqrt{2\log n}$; and J' is the largest integer such that $2^{J'} \leq K\sqrt{n/\log n}$ for some chosen constant $K > 0$. The following theorem states our main result.

Theorem 1 *Suppose that model (2) is valid, the conditions of (1) are met and $e_i = e(t_{(i)})$ are stationary Gaussian noise with zero mean satisfying the conditions of (3) and (4). Suppose also that the mother wavelet ψ has r vanishing moments and is compactly supported. Then the estimator $\hat{f}$ given by (5) achieves within a logarithmic factor almost the optimal convergence rate over the range of Hölder classes $\Lambda^\alpha(M)$ with $\alpha \in (0,r]$ in the sense that*

$$\sup_{f\in\Lambda^\alpha(M)} E||\hat{f} - f||_2^2 \leq C\left(\frac{\log n}{n}\right)^{\alpha/(1+\alpha)}$$

and

$$\sup_{f\in\Lambda^\alpha(M)} \frac{1}{n}\sum_{k=1}^{n} E||\widehat{f(t_k)} - f(t_k)||_2^2 \leq C\left(\frac{\log n}{n}\right)^{\alpha/(1+\alpha)},$$

for all $M \in (0,\infty)$.

The conditions of (3) and (4) occur in diverse applications, and specific cases of interest where the conditions of (1) also occur are given by two propositions in our paper.

Simulations

Our simulation results show that the IMSE on random designs is bigger than those on equispaced design in all the cases. The IMSE for jittering fall between those for uniform and equispaced in almost all the cases. However, the jittered sampling yields almost the same results as the equispaced design so that the efect of small timing errors is small, mainly for bigger sample sizes. Visually, the reconstruction with uniform design is a little more wrinkled than the equispaced and jittered designs. The jittering is visually almost indistinguishable from the equispaced design.

Application

As an example of the application of our methodology, let us consider the problem of estimating the light curve for the variable star RU Andromeda using data obtained from the American Association of Variable Star Observers (AAVSO) International Database at
`www.aavso.org`. For our example, we focused on the 256 successive observations recorded from Julian Day 2,440,043 to 2,441,592 (July 5, 1968 to October 1, 1972).

Figure 1(c) shows the estimated light curve using threshold (6). Note that this light curve differs from the one in Figure 1(a) mainly in the first half of the series, evidently due to the autocorrelated errors. Figure 1(d) shows the sample autocorrelation sequence for the residuals from the fitted curve. The fact that this sequence damps down rapidly is an indication that assuming the conditions of (3) and (4) is reasonable here.

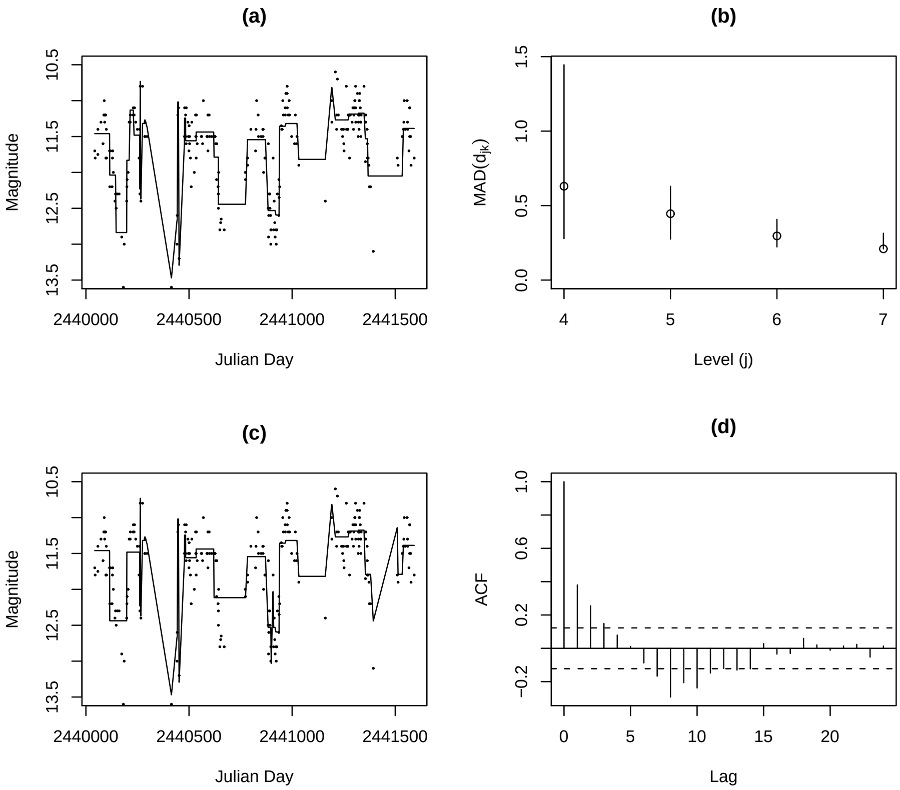

Figure 1: (a) Data points and estimated light curve through VisuShrink. (b) Mean absolute deviation (MAD) from zero of the wavelet coefficients at each resolution level j. Level $j = 7$ is the finest. Endpoints of the error bars are the .025 and .975 quantiles of MAD obtained from 500 samples (with replacement) of the wavelet coefficients at each resolution level j. (c) Data points and estimated light curve considering correlated errors. (d) Residuals sample autocorrelation function and 95% confidence interval.

REFERENCES

T. T. Cai and L. Brown. Wavelet shrinkage for nonequispaced samples. Ann. Statist., 26(5):1783–1799, 1998.

T. T. Cai and L. Brown. Wavelet estimation for samples with random uniform design. Statistics and Probability Letters, 42:313–321, 1999.

M. A. Z. Dippé and E. H. Wold. Antialiasing through stochastic sampling. ACM SIGGRAPH Computer Graphics, 19(3):69–78, 1985.

D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. Biometrika, 81:425–455, 1994.

I. M. Johnstone and B. W. Silverman. Wavelet threshold estimators for data with correlated noise. J. R. Stat. Soc. Ser. B Stat. Methodol., 59(2): 319–351, 1997.

G. Kerkyacharian and D. Picard. Regression in random design and warped wavelets. Bernoulli, 10(6):1053–1105, 2004.

R. F. Porto, P. A. Morettin, and E. C. Q. Aubin. Wavelet regression with correlated errors on a piecewise Hölder class. Statist. Probab. Lett., 78: 2739–2743, 2008.

S. Sardy, D. B. Percival, A. G. Bruce, H. Gao, and W. Stuetzle. Wavelet shrinkage for unequally spaced data. Stat. Comput., 9:65–75, 1999.