2000 Proceedings of the Autumn Conference, Korean Statistical Society

Polynomial Boundary Treatment for Wavelet Regression

Hee-Seok Oh*, **Philppe Naveau*** and **GeungHee Lee****

Abstract

To overcome boundary problems with wavelet regression, we propose a simple method that reduces bias at the boundaries. It is based on a combination of wavelet functions and low-order polynomials. The utility of the method is illustrated with simulation studies and a real example. Asymptotic results show that the estimators are competitive with other nonparametric procedures.

Key Words : Boundary problem, Polynomial regression, Wavelet regression, Polynomial trigonometric regression.

1. Introduction

Consider a classical nonparametric regression problem involving data $y_i = f(i/n) + \varepsilon_i$ with $i = 1, \ldots, n$, where errors are independent and identically distributed normal variables with mean zero and variance σ^2. The true function f is assumed to be square integrable on the interval [0,1].

Most nonparametric smoothing approaches such as kernel smoothers, trigonometric regression and wavelet regression suffer from boundary or edge effects. To overcome boundary problems when dealing with smooth functions, Eubank and Speckman (1990, 1991a) suggested a simple method called polynomial-trigonometric regression, in which they write the estimator of f as a sum of trigonometric function and a low-order polynomial. The latter is expected to account for the boundary problem. They illustrated their approach with examples and showed that the rates of convergence for their estimators over a particular smoothness class of functions are optimal, whether or not the regression function is periodic. However, as noted by these authors, sharps peaks in the regression function can cause a ringing phenomenon and other wiggles, and this limits the usefulness of polynomial-trigonometric regression. To extend this procedure to functions with singularities, a possible solution is to replace the trigonometric part of the estimator by an estimator based on wavelets while keeping the polynomial part to model the boundaries. The choice of wavelets as a replacement of the trigonometric part is a natural one since wavelets are well known to be much more efficient than trigonometric functions when fitting curves with sharp features.

* Geophysical Statistics Projects, National Center for Atmospheric Research, Boulder, Colorado 80307, U.S.A. ** The Bank of Korea. This work was supported in part by grants from the U.K. Engeneering and Physical Council and the U.S. National Science Foundation.

In Section 2, a detailed description of the polynomial-wavelet regression method is presented. The asymptotic properties of the method are discussed in Section 3. In Section 4, the polynomial wavelet regression is tested on a pair of standard case-study functions and a real data example is investigated.

2. Polynomial Wavelet Regression

Let $\{\psi_{j,k}, j\in Z, k\in Z\}$ be an orthonormal wavelet basis for $L^2(R)$ (Daubechies, 1992). Any square integrable function can be then represented by the following expansion:

$$f(x) = \sum_{k=-\infty}^{\infty} c_{0,k}\phi_k(x) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d_{j,k}\psi_{j,k}(x) \tag{1}$$

where $\phi_k(x) = 2^{1/2}\phi(2x-k)$ and $\psi_{j,k}(x) = 2^{j/2}\psi(2^jx-k)$. Here the scaling and detail coefficients are respectively equal to $c_{0,k} = \int_{-\infty}^{\infty} f(x)\phi_k(x)dx$ and $d_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx$. From equation (1), we can express the following classical nonlinear wavelet regression estimator $\widehat{f_W} = \sum_{k=1}^{2^J-1}\widehat{c_{0,k}}\phi_k(x) + \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1}\widehat{d^S_{j,k}}\psi_{j,k}(x)$, where $\widehat{d^S_{j,k}} = sgn(\widehat{d_{j,k}})\max(0, |\widehat{d_{j,k}}| - \lambda)$ denotes the soft-thresholded wavelet coefficients.

Donoho and Johnstone (1994, 1995) showed that $\widehat{f_W}$ with a properly chosen threshold rule, has various important optimality properties. The choice of the shrinkage rule is therefore crucial in wavelet regression. Several approaches to thresholding have been studied. For example, Donoho and Johnstone (1994, 1995) proposed *VisuShrink* and *SureShrink* as minimax approaches, Nason (1996) considered a cross validation method, and Abramovich, Sapatinas and Silverman (1998) looked at a Bayesian thresholding rule, BayesThresh. Such procedures will be used and compared in our simulation study.

To reduce the boundary effects present in classical wavelet regression, we propose the estimator,

$$\widehat{f_{PW}}(x) = \sum_{n=1}^{d}\widehat{\alpha_n}x^n + \sum_{k=1}^{2^J-1}\widehat{c_{0,k}}\phi_k(x) + \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1}\widehat{d^S_{j,k}}\psi_{j,k}(x) \tag{2}$$

where d is a positive integer. This new estimator takes the form, $\widehat{f_{PW}}(x) = \widehat{f_P}(x) + \widehat{f_W}(x)$, where $\widehat{f_P}(x)$ is a polynomial estimator of degree d and $\widehat{f_W}$ is defined by (1). The motivation behind (2) is based on an argument similar to the one given by Eubank and Speckman (1990, 1991a). Suppose that the regression function f satisfies some periodic boundary conditions, for example $f^{(l)}(0) = f^{(l)}(1)$, for $l=0,..,m$. Then the capacity of the wavelet estimator $\widehat{f_W}$ to handle the boundary behaviour of f efficiently will improve as m increases. In real cases, such boundary conditions are rarely satisfied and the estimator $\widehat{f_W}$ is far from optimal at the boundaries; See Fig. 1. Suppose now that $P(x)$ denotes a polynomial of order m such that $P^{(l)}(0) - P^{(l)}(1) = f^{(l)}(0) - f^{(l)}(1)$, and rewrite the true function as $f(x) = g(x) + P(x)$. The difference is that the new function $g(x)$ satisfies $g^{(l)}(0) = g^{(l)}(1)$, and so the periodic properties of the function $g(x)$ are better suited to wavelet regression. Hence, decomposing the signal into a polynomial and wavelet term should improve the fit at the boundaries.

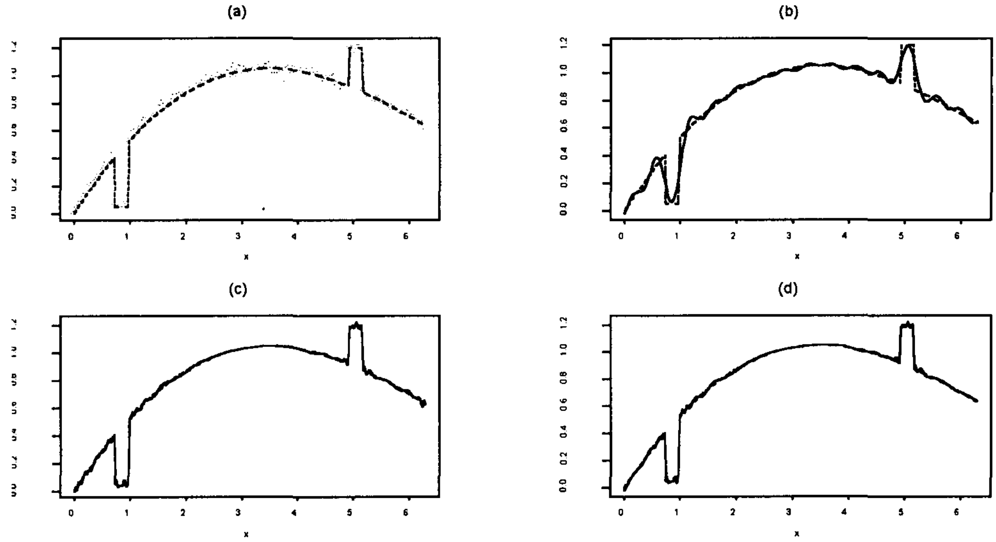

Figure 1 (a) True function and simulated data; (b) The fit by polynomial-trigonometric regression; (c) The fit by wavelet regression; (d) The fit by polynomial-wavelet regression. In all cases, the dotted line denotes the true function.

In order to maintain orthogonality between the set of polynomial basis $x, \ldots, x^d$ and the wavelet basis, the equations $\int x^n\psi(x)dx = \int x^n\phi(x)dx = 0$ have to be satisfied for $n=1,\ldots,d$. Wavelets with such properties were constructed by Daubechies (1992) and named coiflets. Hence, use of a coiflet with at least $d+1$ vanishing moments in (2) implies that the polynomial regression term is orthogonal to the wavelet regression term. A consequence is that the d has to be smaller than the number of vanishing moments. In our application, we will use a coiflet with 5 vanishing moments and $d \le 3$. Eubank and Speckman (1990) fixed $d=2$ for their applications. To estimate the parameters in (2), we first regress the observations y_i on the set $x, \ldots, x^d$ for fixed d and then apply wavelet regression to the residuals of the polynomial regression.

3. Asymptotic Properties

Donoho and & Johnstone (1994) studied the risk of the nonlinear estimator $\widehat{f}_W$ defined by (1), where risk is defined to be $R(\widehat{f}_W, f) = n^{-1}E\| \widehat{f}_W - f \|^2$, i.e. the mean squared error. They derived an upper bound, $R(\widehat{f}_W, f) \le (2\log n + 1)\{ R(S, f) + \sigma^2/n\}$, where $R(S, f)$ is the ideal risk obtained from the thresholding procedure. This means that it is possible to come within a $2\log n$ factor of the performance of ideal wavelet adaptation.

Theorem 1. *Suppose that there exists some positive integer m such that the derivatives* $f^{(l)}(0)$ *and* $f^{(l)}(1)$ *are well defined for l=0,...,m. Then the bias of the estimator* $\widehat{f_{PW}}$ *is reduced,* $|E(\widehat{f_{PW}})-f|^2 \le |E(\widehat{f_W})-f|^2$, *its risk satisfies* $R(\widehat{f_{PW}},f) \le R(\widehat{f_W},f)+\frac{m\sigma^2}{n}$, *and thus the ideal wavelet adaptation inequality becomes* $R(\widehat{f_{PW}},f) \le (2\log n+1)\left\{ R(T,f)+\frac{\sigma^2}{n} \right\}+\frac{m\sigma^2}{n}$.

This theorem shows that the cost associated with a reduction of the bias at the edges is asymptotically small, $m\sigma^2/n$, and the estimator $\widehat{f_{PW}}$ maintains the excellent performance obtained by classical wavelet shrinkage.

4. Simulations and an Application

To assess the numerical performance of (2), we select the Blocks function from the standard test functions of Donoho and Johnstone (1994) as the periodic case, and a function from Fan and Gijbels (1995) as the non-periodic case; See Fig. 2. The quality of each estimator, corresponding to different ways of treating the boundaries as symmetric, periodic or polynomial, was measured by computing the average and the standard error of the mean squared error $R(\hat{f},f)$ over 1000 simulations. Each simulation contains a sample of 256 observations contaminated by Gaussian white noise. As expected, Table 1 indicates that polynomial-wavelet regression provides the best result for the function that departs the most from symmetry and periodicity, i.e. a function from Fan & Gijbels (1995), and performs as well as classical regression for the periodic case, i.e. the Blocks function.

Note that the boundary problem appears locally, near the edges, and so the mean squared error, which is a global measure of quality of fit, does not necessarily describe the gain obtained by polynomial wavelet regression. A natural estimator of local discrepancy is

$$R_\tau(\hat{f},f)=\frac{1}{2\tau}\sum_{i\in N(\tau)}E\{\hat{f}(x_i)-f(x_i)\}^2 \quad \text{for } \tau=1,\ldots,[n/2],\ x_i=i/n,$$

where $N(\tau)=\{1,\ldots,\tau,n-\tau+1,\ldots,n\}$. Figure 3 shows that polynomial wavelet regression applied to the function from Fan and Gijbels (1995) performs much better near boundaries, i.e. with τ small, than does classical wavelet regression, and it clearly indicates that the polynomial term in (2) removes the artificial wiggles observed with classical wavelet regression.

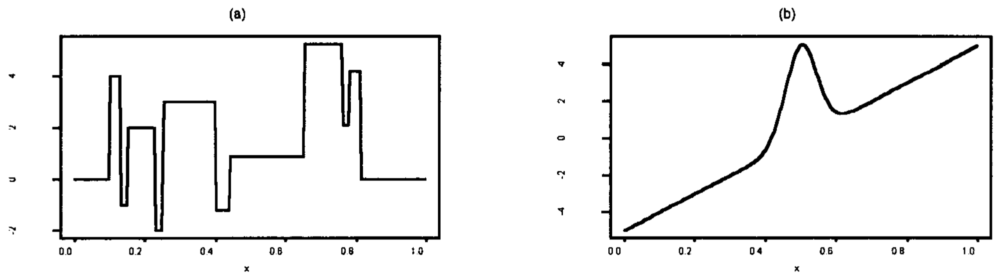

Figure 2 (a) Blocks function and (b) the function from Fan and Gijbels (1995)

Table 1 Estimated average and standard deviation, given in parentheses, of the risk using the SureShrink (Sure), cross validation (CV) and BayesThresh (Bayes) thresholding rules over 1000 replications for different values of noise standard deviation σ^2

Function	Threshold	σ^2	PWR	WR(P)	WR(S)
Blocks	Sure	0.3	0.0946 (0.0178)	0.0958 (0.0183)	0.0971 (0.0181)
		0.5	0.2125 (0.0446)	0.2147 (0.0548)	0.2173 (0.0556)
		0.7	0.8227 (0.0905)	0.8339 (0.0977)	0.8407 (0.0953)
	CV	0.3	0.0755 (0.0102)	0.0752 (0.0100)	0.0769 (0.0101)
		0.5	0.1634 (0.0202)	0.1633 (0.0197)	0.1682 (0.0201)
		0.7	0.2853 (0.0371)	0.2855 (0.0371)	0.2959 (0.0379)
	Bayes	0.3	0.0886 (0.0161)	0.0896 (0.0160)	0.0917 (0.0158)
		0.5	0.2682 (0.0560)	0.2702 (0.0569)	0.2736 (0.0538)
		0.7	0.4645 (0.0722)	0.4716 (0.0736)	0.4722 (0.0698)
F & G	Sure	0.3	0.0182 (0.0043)	0.0360 (0.0065)	0.0222 (0.0047)
		0.5	0.0623 (0.0132)	0.0854 (0.0153)	0.0722 (0.0142)
		0.7	0.0982 (0.0224)	0.1461 (0.0323)	0.1163 (0.0238)
	CV	0.3	0.0188 (0.0043)	0.0384 (0.0057)	0.0331 (0.0065)
		0.5	0.0434 (0.0104)	0.0918 (0.0146)	0.0774 (0.0157)
		0.7	0.0740 (0.0199)	0.1609 (0.0290)	0.1362 (0.0298)
	Bayes	0.3	0.0150 (0.0038)	0.0317 (0.0062)	0.0198 (0.0043)
		0.5	0.0322 (0.0090)	0.0696 (0.0127)	0.0435 (0.0098)
		0.7	0.0516 (0.0141)	0.1144 (0.0277)	0.0702 (0.0158)

F & G: the function from Fan and Gijbels (1995) in Fig. 2(b)
PWR: polynomial-wavelet regression
WR(P): wavelet regression with periodic boundary treatment rule
WR(S): wavelet regression with symmetric boundary treatment rule.

We applied the polynomial-wavelet regression to the voltage drop data discussed by Eubank and Speckman (1990). This dataset of 32 observations represents the voltage drop in the battery of a guided missile motor during flight. Figure 4 shows that polynomial wavelet regression provides a clear improvement near both boundaries.

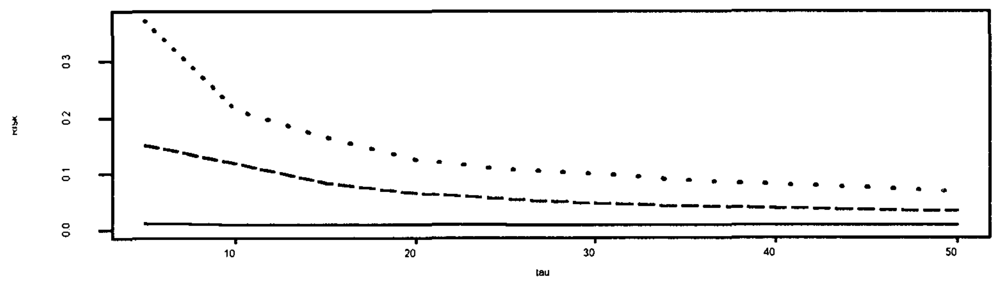

Figure 3 Risk $R_\tau(\hat{f}, f)$ versus τ for estimates of the function from Fan and Gijbels (1995) based on BayesThresh and noise level σ^2=0.5. The solid line is for polynomial-wavelet regression, the dotted line is for wavelet regression with periodic boundary correction, and the dashed line is for wavelet regression with symmetric boundary correction.

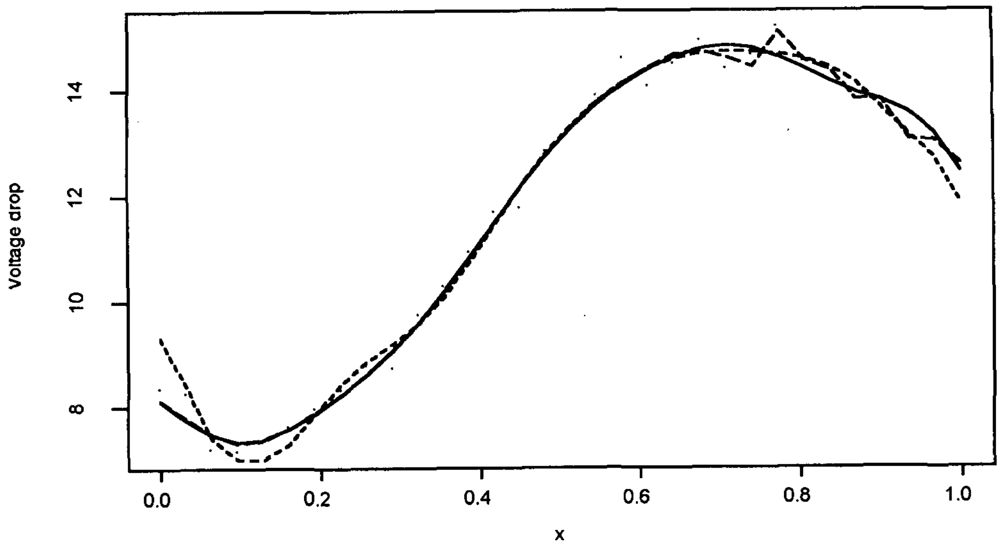

Figure 4 The fits of the voltage drop data. The x variable represents time. Polynomial-wavelet regression is shown by the solid line, wavelet regression with periodic boundary by the short dashed line, and wavelet regression with symmetric boundary correction by the long dashed line.

Acknowledgement

The authors would like to thank Guy Nason and Doug Nychka for helpful comments.

References

Abramovich, F., Sapatinas, T. and Silverman, B. W. (1998). Wavelet thresholding via a Bayesian approach. *J. R. Statist. Soc. B 60, 725-49.*

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM

Donoho, D. and Johnstone, I. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81, 425-55.

Donoho, D. & Johnstone, I. (1995). Adapting to unknown smoothing via wavelet shrinkage. *J. Am. Statist. Assoc.* 90, 1200-24.

Eubank, R. L. and Speckman, P. (1990). Curve fitting by polynomial-trigonometric regression. Biometrika 77, 1-9.

Eubank, R. L. and Speckman, P. (1991a). Convergence rates for trigonometric and polynomial-trigonometric regression estimators. *Statist. Prob. Lett.* 11, 119-24.

Eubank, R. L. and Speckman, P. (1991b). A bias reduction theorem with applications in nonparametric regression. *Scand. J. Statist.* 18, 211-22.

Fan, J. and Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation. *J. R. Statist. Soc. B* 57, 371-94.

Nason, G. P. (1996). Wavelet shrinkage using cross-validation. *J. R. Statist. Soc. B* 58, 463-79.