

## Confidence Band for Additive Regression Model

Lijian Yang  
*Michigan State University*

*Abstract:* Additive model is widely recognized as an effective tool for dimension reduction. Existing methods for estimation of additive regression function, including backfitting, marginal integration, projection and spline methods, do not provide any level of uniform confidence. In this paper a simple construction of confidence band is proposed for the additive regression function based on polynomial spline estimation and wild bootstrap. Monte Carlo results show three desirable properties of the proposed band: excellent coverage of the true function, width rapidly shrinking to zero with increasing sample size, and minimal computing time. These properties make the procedure highly recommended for nonparametric regression with confidence when additive modelling is appropriate.

*Key words:* Confidence interval, curse of dimensionality, inflation ratio, linear splines, wild bootstrap.

### 1. Introduction

Non- and semi- parametric smoothing estimation of unknown regression functions has found many applications in modelling nonlinearity which exists widely in various disciplines such as econometrics (Feng 2004, Racine and Li 2004), and biological sciences (Liang 2004, Liang, Wang, Robins and Carroll 2004). One common reservation about using nonparametric smoothing is the “curse of dimensionality”, which refers to the lack of accuracy in estimating functions of general nonparametric multivariate form. Additive modelling is a very effective technique to alleviate the curse of dimensionality without constraining the regression function to simple linear form, see for example, Hastie and Tibshirani (1990) and the more recent works of Chen and Tsay (1993), Tjøstheim and Auestad (1994), Linton and Nielsen (1995), Yang, Härdle and Nielsen (1999), Sperlich, Tjøstheim and Yang (2002), and Yang, Sperlich and Härdle (2003), which applied various kernel smoothing techniques. Huang and Yang (2004), on the other hand, employed spline smoothing methods.

As with any estimation procedure, confidence regions provide more information than a single estimate can. With nonparametric smoothing, however, confidence bands for the regression functions have proven to be one of the more challenging problems, and existing literature is rather limited and narrow in scope. The early works of Hall and Titterton (1988), Härdle (1989) established asymptotic confidence band of nonparametric regression function which are univariate, based on Nadaraya-Watson kernel smoothing. This line of research has been extended to the more efficient local polynomial kernel smoothing in Xia (1998), which has further developed into bootstrap procedure in Claeskens and Van Keilegom (2003). All of the aforementioned methods involve local averaging by kernel weights, and requires solving a linear least squares problem for every design point. In contrast, Wang and Yang (2006) propose new confidence bands based on the polynomial spline method, which obliges the user to solving merely one linear least squares problem, and is practically as easy to implement and fast as a simple linear regression with slowly increasing number of parameters. Such nonparametric confidence bands, however, does not exist for estimation of additive regression function. Techniques based on kernel smoothing reply on complicated backfitting or averaging algorithms [also known as projection in Tjøstheim and Auestad (1994) and marginal integration in Linton and Nielsen (1995)], from which it is extremely difficult to construct confidence bands. Spline technique is therefore chosen for tackling this difficult problem.

In this paper the spline confidence band for univariate regression in Wang and Yang (2006) is extended to confidence band of multivariate additive regression. In section 2, I propose an asymptotically conservative confidence band for the nonparametric autoregression function. Intuitive justifications are given for the procedure, which is based on wild bootstrap. In section 3, I examine the properties of the procedure through simulation study and discuss the issues in implementation. Section 4 concludes.

## 2. Additive Confidence Band

The additive regression model is of the form

$$Y = m(\mathbf{X}) + \varepsilon, \mathbf{X} = (X_1, \dots, X_d), m(\mathbf{X}) = c + \sum_{\alpha=1}^d m_{\alpha}(X_{\alpha}) \quad (2.1)$$

with  $E(\varepsilon) = 0$ ,  $\text{var}(\varepsilon) = \sigma^2$  and the identification condition that  $E m_{\alpha}(X_{\alpha}) \equiv 0$ ,  $\alpha = 1, \dots, d$ . Given a sample  $\{Y_i, \mathbf{X}_i\}_{i=1}^n = \{Y_i, X_{i1}, \dots, X_{id}\}_{i=1}^n$  generated from (2.1), estimators of the functions  $\{m_{\alpha}(x_{\alpha})\}_{\alpha=1}^d$  and  $m(\mathbf{x})$  can be constructed by employing either kernel or spline smoothing. In this section, I describe the spline estimator  $\hat{m}(\mathbf{x})$  of  $m(\mathbf{x})$  and the construction of a conservative confidence band

based on  $\hat{m}(\mathbf{x})$ .

One needs to first introduce the spline basis of order  $p$  ( $p \geq 1$ ) as the following set of piecewise polynomials

$$\begin{aligned} B_{1,\beta}(x_\beta) &= x_\beta, \dots, B_{p,\beta}(x_\beta) = x_\beta^p \\ B_{p+1,\beta}(x_\beta) &= \left(x_\beta - t_\beta^{(1)}\right)_+^p, \dots, B_{p+N,\beta}(x_\beta) = \left(x_\beta - t_\beta^{(N)}\right)_+^p \end{aligned}$$

in which  $\left\{t_\beta^{(J)}\right\}_{J=1}^N$  are equally spaced points on the range of variable  $X_\beta$ , called knots. To be precise, denote the compact support of the density  $f_\beta(x_\beta)$  of variable  $X_\beta$  as  $[a_\beta, b_\beta]$ , then

$$t_\beta^{(J)} = a_\beta + J(b_\beta - a_\beta)(N+1)^{-1}, J = 1, \dots, N$$

where the number of knots  $N = N_n = \lceil kn^{1/(2p+3)} \rceil$ , in which  $k$  is a tuning constant. One then define for any  $\mathbf{x} = (x_1, \dots, x_d)$

$$\hat{m}(\mathbf{x}) = \hat{\lambda}_{0,1} + \sum_{\alpha=1}^d \sum_{L=1}^{p+N} \hat{\lambda}_{L,\alpha} B_{L,\alpha}(x_\alpha) \quad (2.2)$$

where the coefficients  $(\hat{\lambda}_{0,1}, \hat{\lambda}_{1,1}, \dots, \hat{\lambda}_{p+N,d})$  are solutions of the following least squares problem

$$\begin{aligned} \hat{\lambda} &= \left\{ \hat{\lambda}_{0,1}, \hat{\lambda}_{1,1}, \dots, \hat{\lambda}_{p+N,d} \right\}^T \\ &= \underset{R^{1+dp+N}}{\operatorname{argmin}} \sum_{i=1}^n \left\{ Y_i - \lambda_{0,1} - \sum_{\alpha=1}^d \sum_{L=1}^{p+N} \lambda_{L,\alpha} B_{L,\alpha}(X_{i\alpha}) \right\}^2. \quad (2.3) \end{aligned}$$

Although splines of any order  $p \geq 1$  can be employed, the simplicity of linear splines (with  $p = 1$ ), which are piecewise linear functions, makes them especially appealing. Thus in this paper I have used only linear splines. The tuning constant is set to  $k = 1$ , which is not crucial based on observation from extensive simulation. In what follows, a confidence level  $1 - \alpha \in (0, 1)$  is fixed, whose default is 0.95. The main goal is to construct simultaneous confidence intervals

for the value of function  $m(\mathbf{x})$  at each point  $\mathbf{x} \in \mathcal{X} = \prod_{\beta=1}^d [a_\beta, b_\beta]$ , in other words, a confidence band for the function  $m(\mathbf{x})$  over  $\mathcal{X}$ .

For univariate spline regression ( $d = 1$ ), Huang (2003) had obtained asymptotic pointwise confidence intervals of level  $(1 - \alpha)$  in the form of  $\hat{m}(x) \pm \sigma_{\hat{m}}(x) \cdot z_{1-\alpha/2}$ , in which  $z_{1-\alpha/2}$  is the  $(1 - \alpha/2)$ -quantile of the standard normal distribution,  $\hat{m}(x)$  is as defined in (2.2) while  $\sigma_{\hat{m}}(x)$  is its standard error, whose explicit

formula is contained in Huang (2003) in the conditional form, and Wang and Yang (2006) in the unconditional form. The latter had employed the Brownian bridge strong approximation theorem of Tusnády (1977) to obtain the following simultaneous confidence intervals based on linear spline estimator

$$\hat{m}(x) \pm \sigma_{\hat{m}}(x) \{2 \log(N+1) - 2 \log \alpha\}^{1/2} = \hat{m}(x) \pm \sigma_{\hat{m}}(x) \sqrt{\chi_{2,1-\alpha/(N+1)}^2}$$

where I denote by  $\chi_{t,1-\alpha}^2$  the  $(1-\alpha)$ -quantile of the chi-square distribution with  $t$  degrees of freedom. Comparing the pointwise confidence interval of Huang (2003) and the simultaneous confidence intervals (i.e., a confidence band) of Wang and Yang (2006), one clearly see that both are centered at  $\hat{m}(x)$ , while simultaneity is achieved by inflating the pointwise interval at the ratio of  $z_{1-\alpha/2}^{-1} \sqrt{\chi_{2,1-\alpha/(N+1)}^2}$ , or  $\sqrt{\chi_{2,1-\alpha/(N+1)}^2 \chi_{1,1-\alpha}^{-2}}$ . It was established in both Huang (2003) and Wang and Yang (2006) that  $\sigma_{\hat{m}}(x)$  is of order  $n^{-2/5}$ , hence the confidence band has a bound of order  $n^{-2/5} \sqrt{\log(n)}$  for its width throughout the interval  $[a_1, b_1]$ .

In the case of additive model (2.1), the deviation of  $\hat{m}(\mathbf{x})$  from  $m(\mathbf{x})$  is to be bounded on a total of  $(N+1)^d$   $d$ -dimensional intervals. On each interval, the deviation comes from each of the  $d$  variables. Hence the degree of freedom for the chi-square quantile in the inflation ratio is  $2d$ , while the tail probability is  $1-\alpha/(N+1)^d$ . Hence the inflation ratio is naturally  $z_{1-\alpha/2}^{-1} \sqrt{\chi_{2d,1-\alpha/(N+1)^d}^2}$ . In the case of univariate regression (i.e.  $d=1$ ), the ratio is the exact same one proved in Wang and Yang (2006). In addition, the confidence band should have a bound of order  $n^{-2/5} \sqrt{\log(n)}$  for its width throughout the compact set  $\mathcal{X} = \prod_{\beta=1}^d [a_\beta, b_\beta]$ . A rigorous proof of this asymptotic result would be feasible if strong approximation result as sharp as Tusnády (1977) would be available for  $d+1$  dimensions.

Wang and Yang (2006) has used a plug-in procedure to estimate  $\sigma_{\hat{m}}(x)$  in the construction of the confidence band. The plug-in method, however, requires additional smoothing steps which may render the results less accurate, and hence has been replaced in this paper with wild bootstrap for improved performance. The steps are described here. In the following, I define the residuals  $\hat{\varepsilon}_i = Y_i - \hat{m}(\mathbf{X}_i)$ ,  $1 \leq i \leq n$ , and denote a predetermined large integer by  $n_B$ , whose default value is 400.

*Step 1* Let  $\{\delta_{i,b}\}_{\substack{1 \leq b \leq n_B \\ 1 \leq i \leq n}}$  be i.i.d. samples of the following discrete distribution

$$\delta_{i,b} = \begin{cases} 2^{-1} (1 - \sqrt{5}) & \text{with probability } 10^{-1} (5 + \sqrt{5}) \\ 2^{-1} (1 + \sqrt{5}) & \text{with probability } 10^{-1} (5 - \sqrt{5}) \end{cases},$$

$1 \leq i \leq n, 1 \leq b \leq n_B$ . One can verify that each  $\delta_{i,b}$  defined above satisfies  $E(\delta_{i,b}) \equiv 0, \text{var}(\delta_{i,b}) \equiv 1$ .

*Step 2* For any  $1 \leq b \leq n_B$ , define the  $b$ -th wild bootstrap sample

$$Y_{i,b} = \hat{m}(\mathbf{X}_i) + \delta_{i,b}\hat{\varepsilon}_i, 1 \leq i \leq n. \quad (2.4)$$

Next, replace the response  $Y_i$  in (2.3) with  $Y_{i,b}$  and then solve the following least squares problem

$$\hat{\lambda}_b = \underset{R^{1+d(p+N)}}{\text{argmin}} \sum_{i=1}^n \left\{ Y_{i,b} - \lambda_{0,1} - \sum_{\alpha=1}^d \sum_{L=1}^{p+N} \lambda_{L,\alpha} B_{L,\alpha}(X_{i\alpha}) \right\}^2$$

and obtain the  $b$ -th bootstrap estimator of  $m(\mathbf{x})$

$$\hat{m}^{(b)}(\mathbf{x}) = \hat{\lambda}_{0,1,b} + \sum_{\alpha=1}^d \sum_{L=1}^{p+N} \hat{\lambda}_{L,\alpha,b} B_{L,\alpha}(x_\alpha). \quad (2.5)$$

*Step 3* Denote by  $\hat{m}^{L,\alpha/2}(\mathbf{x})$  and  $\hat{m}^{U,\alpha/2}(\mathbf{x})$  respectively the lower and upper  $100(1 - \alpha/2)\%$  quantiles of the set  $\{\hat{m}^{(b)}(\mathbf{x})\}_{1 \leq b \leq n_B}$ , obtained from (2.5) for each of the bootstrap sample in (2.4). The wild bootstrap  $(1 - \alpha)$  pointwise confidence interval for function value  $m(\mathbf{x})$  at one point  $\mathbf{x}$  is  $[\hat{m}^{L,\alpha/2}(\mathbf{x}), \hat{m}^{U,\alpha/2}(\mathbf{x})]$ .

*Step 4* Denote the wild bootstrap  $(1 - \alpha)$  confidence band for the function  $m(\mathbf{x})$  over  $\mathcal{X}$  as  $[\hat{m}^{L,\alpha/2}(\mathbf{x}), \hat{m}^{U,\alpha/2}(\mathbf{x})], \mathbf{x} \in \mathcal{X}$ . As mentioned earlier, the uniform confidence band needs to be wider than the pointwise confidence interval by a factor of  $z_{1-\alpha/2}^{-1} \sqrt{\chi_{2d,1-\alpha/(N+1)}^2}$  when localized at any point, hence I define

$$\hat{m}^{L,\alpha/2}(\mathbf{x}) = \hat{m}(\mathbf{x}) + \left\{ \hat{m}^{L,\alpha/2}(\mathbf{x}) - \hat{m}(\mathbf{x}) \right\} K \quad (2.6)$$

$$\hat{m}^{U,\alpha/2}(\mathbf{x}) = \hat{m}(\mathbf{x}) + \left\{ \hat{m}^{U,\alpha/2}(\mathbf{x}) - \hat{m}(\mathbf{x}) \right\} K, \quad (2.7)$$

where  $K = z_{1-\alpha/2}^{-1} \sqrt{\chi_{2d,1-\alpha/(N+1)}^2}$ .

The justification for making use of the wild bootstrap sample (2.4) is the same as in Sperlich, Tjøstheim and Yang (2002), i.e., the sample  $\{Y_{i,b}, \mathbf{X}_i\}_{i=1}^n$  mimicks the dependence structure of sample  $\{Y_i, \mathbf{X}_i\}_{i=1}^n$  up to second order. Intuitively this is seen from the aforementioned fact that  $E(\delta_{i,b}) \equiv 0, \text{var}(\delta_{i,b}) \equiv 1$  and therefore each  $\delta_{i,b}\hat{\varepsilon}_i$  has the same mean and variance as  $\hat{\varepsilon}_i$ , approximately 0 and

$\sigma^2$ . In the next section, I will demonstrate the excellent practical performance of the proposed band by applying it to simulated examples.

### 3. Monte Carlo Study

To evaluate the performance of the confidence band defined in (2.6) and (2.7), the following Monte Carlo study has been carried out. The data generating model is an additive regression model of the form

$$Y_i = c + \sum_{\beta=1}^d m_{\beta}(X_{i\beta}) + \varepsilon_i, 1 \leq i \leq n, c = 2, m_{\beta}(x_{\beta}) = \sin(2\pi x_{\beta}), 1 \leq \beta \leq d \quad (3.1)$$

with predictor vector  $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,d})$  i.i.d.  $U[0, 1]^d$  and the error  $\varepsilon_i$  i.i.d.  $N(0, 1)$ . A total of 100 replications  $\{Y_i^{(k)}, \mathbf{X}_i^{(k)}\}_{i=1, k=1}^{n, 100}$  are generated according to model (3.1) for sample sizes  $n = 50, 100, 200, 400$  and dimensions  $d = 2, 4$ , and the 95% conservative confidence band computed for each replication. All computing has been carried out in the environment of XploRe [see, for example, Härdle, Hlavka and Klinke (2000)] or visit the link

<http://www.quantlet.com/mdstat/scripts/xag/html/xaghtml.html>

The simulation attempts to address three questions: First, does the conservative confidence band given in (2.6) and (2.7) contains the additive function with empirical frequencies higher than the predetermined confidence level? Second, is the conservative confidence band too wide? Third, how much computing is needed to produce the confidence band?

The answer to the first question is extremely positive: for all the experiments, it is observed that the empirical frequencies of coverage are 100% in all cases. Hence the constructed band is indeed conservative in terms of coverage.

To address the second and third questions, I have plotted the bands as well as the true additive function against the first variable  $X_{i,1}^{(20)}$ , in Figures 1 and 2, for the 20-th replication of each dimension  $d$ , sample size  $n$  combination. To be precise, the plots are made by connecting points

$$\left\{ X_{i,1}^{(20)}, \widehat{m}^{L,0.025}(\mathbf{X}_i^{(20)}) \right\}_{i=1}^n, \left\{ X_{i,1}^{(20)}, \widehat{m}^{U,0.025}(\mathbf{X}_i^{(20)}) \right\}_{i=1}^n$$

(thin lines) and

$$\left\{ X_{i,1}^{(20)}, m(\mathbf{X}_i^{(20)}) \right\}_{i=1}^n$$

(thick line) respectively. For each dimension  $d$ , sample size  $n$  combination, I have also recorded in Table 1 the average runtime of one replication, as well as

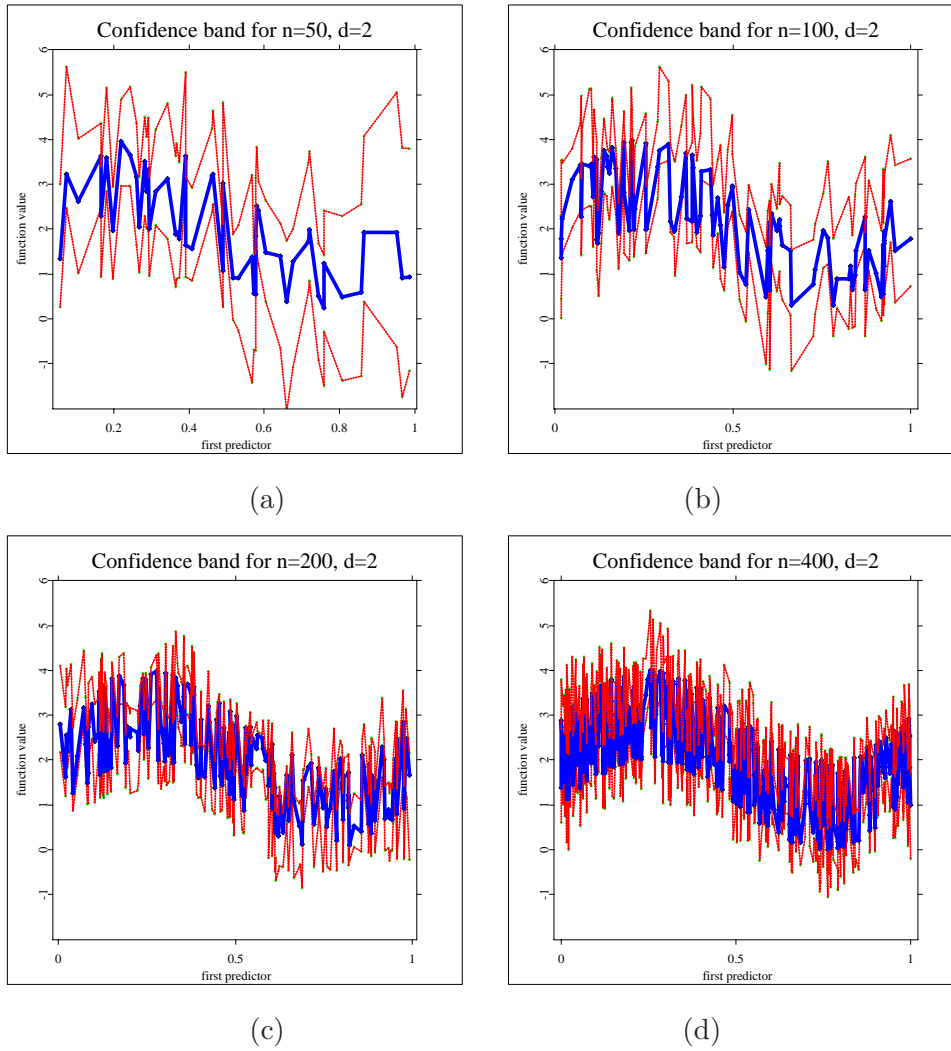


Figure 1: Confidence bands constructed from the 20-th data set generated by model (3.1). The thin lines are the band's lower and upper boundaries  $\left\{X_{i,1}^{(20)}, \widehat{m}^{L,0.025}\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$  and  $\left\{X_{i,1}^{(20)}, \widehat{m}^{U,0.025}\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$  respectively, while the thick line is plot of the true function  $\left\{X_{i,1}^{(20)}, m\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$ .

the average of the width of the confidence bands over all sample points, and all 100 replications of the same dimension  $d$ , sample size  $n$  combination. The runtime values are recorded for experiments carried out on a Dell notebook, while the

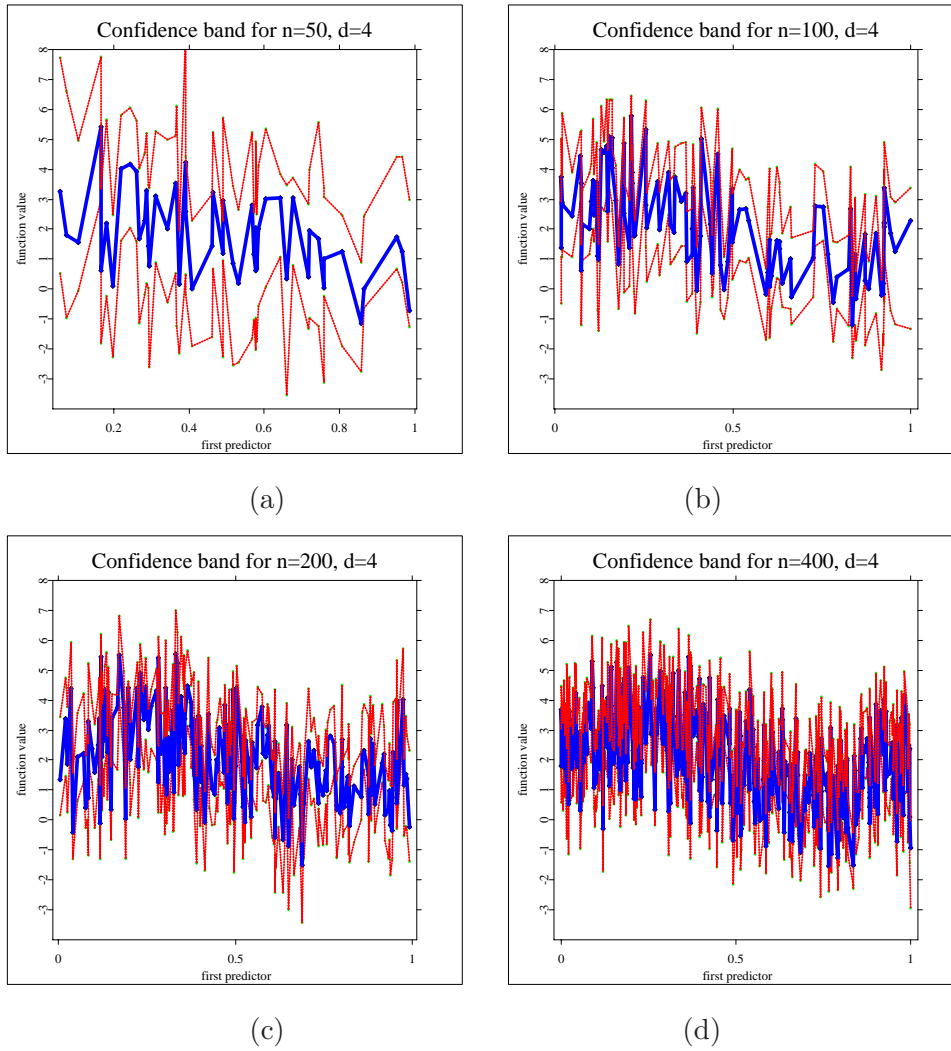


Figure 2: Confidence bands constructed from the 20-th data set generated by model (3.1). The thin lines are the band's lower and upper boundaries  $\left\{X_{i,1}^{(20)}, \widehat{m}^{L,0.025}\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$  and  $\left\{X_{i,1}^{(20)}, \widehat{m}^{U,0.025}\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$  respectively, while the thick line is plot of the true function  $\left\{X_{i,1}^{(20)}, m\left(\mathbf{X}_i^{(20)}\right)\right\}_{i=1}^n$ .

mean width of confidence band is defined as

$$\frac{1}{100} \sum_{k=1}^{100} \frac{1}{n} \sum_{i=1}^n \left\{ \widehat{m}^{U,\alpha/2}\left(\mathbf{X}_i^{(k)}\right) - \widehat{m}^{L,\alpha/2}\left(\mathbf{X}_i^{(k)}\right) \right\}.$$

From Figures 1 and 2, one can see that the band narrows for the 20-th sample as sample size  $n$  increases. This is corroborated by the confidence bands' mean



widths in Table 1, which decreases as sample size increases, thus the confidence bands are not "too conservative" so to speak. Table 1 also shows that the runtime for constructing one confidence band is less than one second when  $n = 400$ , hence the computing burden is extremely light.

Table 1: Asymptotic performance and computing cost of the additive confidence band

Sample size $n$	Mean widths of confidence band ( $d = 2$ )	Mean widths of confidence band ( $d = 4$ )	Mean runtime (seconds)
50	2.735	4.935	0.15
100	2.016	2.993	0.27
200	1.296	2.729	0.45
400	1.318	2.119	0.88

#### 4. Conclusions

In this paper, a method is proposed for constructing confidence band of additive nonparametric regression function, filling one significant current gap in statistical literature. The band covers the true regression function with frequency much higher than the nominal confidence level, can be computed within fractions of a second, and becomes narrower rather rapidly with increasing sample size. These three features make the proposed confidence band very appealing for application.

The idea of this paper could be easily extended to related regression models. For instance, the recent works of Xue and Yang (2006a,b) have made both kernel and spline estimation available for the so-called additive coefficient model, a class of model that includes as special cases additive model, functional coefficient model, and varying coefficient model. It is conceivable that similar confidence band could be constructed for this broad model class.

#### Acknowledgements

The research has been partially supported by NSF grants SES 0127722 and DMS 0405330, and carried out while the author was an ASA/NSF/BLS Research Fellow at the U.S. Bureau of Labor Statistics. The author gratefully acknowledges the helpful comments and suggestions he has received from Stuart Scott at the U.S. Bureau of Labor Statistics. An anonymous referee's comments are also appreciated, which have led to improvements in several places.

---

**References**

- Chen, R. and Tsay, R. S. (1993). Nonlinear additive ARX models. *Journal of the American Statistical Association* **88**, 955-967.
- Claeskens, G. and Van Keilegom, I. (2003). Bootstrap confidence bands for regression curves and their derivatives. *The Annals of Statistics* **31**, 1852-1884.
- Feng, Y. (2004). Simultaneously modeling conditional heteroskedasticity and scale change. *Econometric Theory* **20**, 563-596.
- Hall, P. and Titterton, D. M. (1988). On confidence bands in nonparametric density estimation and regression. *Journal of Multivariate Analysis* **27**, 228-254.
- Härdle, W. (1989). Asymptotic maximal deviation of M-smoothers. *Journal of Multivariate Analysis* **29**, 163-179.
- Härdle, W., Hlavka, Z. and Klinke, S. (2000). *XploRe Application Guide*. Springer.
- Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman & Hall.
- Huang, J. Z. (2003). Local asymptotics for polynomial spline regression. *The Annals of Statistics* **31**, 1600-1635.
- Huang, J. and Yang, L. (2004). Identification of nonlinear additive autoregressive models. *Journal of the Royal Statistical Society Series B* **66**, 463-477.
- Liang, H. (2004). Comparison of curves based on a Cramér-von Mises statistic. *Computational Statistics and Data Analysis* **45**, 805-812.
- Liang, H., Wang, S., Robins, J. M. and Carroll, R. J. (2004). Estimation in partially linear models with missing covariates. *Journal of the American Statistical Association* **99**, 357-367.
- Linton, O. B. and Nielsen, J. P., 1995. A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika* **82**, 93-100.
- Racine, J. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* **119**, 99-130.
- Sperlich, S., Tjøstheim, D. and Yang, L. (2002). Nonparametric estimation and testing of interaction in additive models. *Econometric Theory* **18**, 197-251.
- Tjøstheim, D. and Auestad, B. (1994). Nonparametric identification of nonlinear time series: projections. *Journal of the American Statistical Association* **89**, 1398-1409.
- Tusnády, G. (1977). A remark on the approximation of the sample DF in the multidimensional case. *Periodica Mathematica Hungarica* **8**, 53-55.
- Wang, J. and Yang, L. (2006). Polynomial spline confidence bands for regression curves. Manuscript.

- 
- Xia, Y. (1998). Bias-corrected confidence bands in nonparametric regression. *Journal of the Royal Statistical Society Series B* **60**, 797-811.
- Xue, L. and Yang, L. (2006a). Estimation of semiparametric additive coefficient model. *Journal of Statistical Planning and Inference* **136**, 2506-2534.
- Xue, L. and Yang, L. (2006b). Additive coefficient modeling via polynomial spline. *Statistica Sinica* **16**, 1423-1446.
- Yang, L., Härdle, W. and Nielsen, J. P. (1999). Nonparametric autoregression with multiplicative volatility and additive mean. *Journal of Time Series Analysis* **20**, 579-604.
- Yang, L., Sperlich, S. and Härdle, W. (2003). Derivative estimation and testing in generalized additive models. *Journal of Statistical Planning and Inference* **115**, 521-542.

February 25, 2007; accepted March 25, 2007.

Lijian Yang  
Department of Statistics and Probability  
Michigan State University  
East Lansing, MI 48824, USA  
yang@stt.msu.edu