Comput Stat (2012) 27:573–584
DOI 10.1007/s00180-011-0277-x

ORIGINAL PAPER

Smoothing splines using compactly supported, positive definite, radial basis functions

Guoyi Zhang

Received: 4 October 2008 / Accepted: 22 August 2011 / Published online: 10 September 2011

Abstract In this paper, we develop a fast algorithm for a smoothing spline estimator in multivariate regression. To accomplish this, we employ general concepts associated with roughness penalty methods in conjunction with the theory of radial basis functions and reproducing kernel Hilbert spaces. It is shown that through the use of compactly supported radial basis functions it becomes possible to recover the band structured matrix feature of univariate spline smoothing and thereby obtain a fast computational algorithm. Given n data points in R^2, the new algorithm has complexity $O(n^2)$ compared to $O(n^3)$, the order for the thin plate multivariate smoothing splines.

Keywords Computational complexity · Fourier transform · Generalized cross validation · Nonparametric regression · Reproducing kernel Hilbert space

1 Introduction

Spline smoothing is an important statistical tool for nonparametric function estimation. Whittaker (1923) first used smoothing splines for graduating data. Subsequently, there have been numerous papers and books on splines, many of them focusing on numerical computation rather than statistical properties (Ahlberg et al. 1967 and De Boor 1978). The recognition of spline smoothing as a statistical tool is largely from Wahba's efforts in the late 1980's: Wahba and Wendelberger (1980); Wahba (1981); Villalobos and Wahba (1983, 1987). Books on the spline smoothing include those by Eubank (1988), Wahba (1990), Green and Silverman (1994) and Gu (2002).

The statistical properties of splines are now fairly well understood and we have seen enormous applications of smoothing splines in myriad different disciplines during the

G. Zhang (✉)
Department of Mathematics and Statistics, University of New Mexico,
Albuquerque, NM, USA
e-mail: gzhang123@gmail.com

past twenty years. Smoothing splines have many attractive properties when compared with other nonparametric function estimation methods. In particular, smoothing splines are the most efficient method from a computational perspective. This property is a consequence of the band matrices that arise in the normal equation system for the smoother. Anselone and Laurent (1968), Reinsch (1971), Lyche and Schumaker (1973) and De Boor (1978) developed some different but equivalent (from a fit perspective) formulations that produce and lead to some fast algorithms for evaluation of smoothing spline fits.

There have been many attempts to extend smoothing splines into higher dimensional settings (Chui 1988, Berlinet and Thomas-Agnan 2004). Thin plate smoothing splines (see Duchon 1977 and Wahba and Wendelberger 1980) are arguably the most popular such method. However, the currently popular generalized smoothing splines including thin plate splines that are used in high dimensional settings have complexity $O(n^3)$ for samples of size n, which makes them computationally slow and difficult to use with large data sets. The research reported in this paper is to obtain a smoothing spline for use in higher dimensional settings with comparable estimation accuracy while being more computationally efficient.

Our solution to the multivariate computation problem relies on introducing band structure back into the normal equations for smoothing spline estimators. This feature is missing from the thin plate spline formulation because the associated basis functions have global support. Consequently, we propose an alternative formulation for the smoothing problem using a different penalty that leads to basis functions having local support. The resulting basis functions are the compactly supported, positive definite, radial basis functions that give rise to the title.

This paper is organized as follows. In Sects. 2 and 3, we review compactly supported, positive definite, radial basis functions, their reproducing kernel Hilbert spaces (RKHS) and multivariate smoothing splines using compactly supported, positive definite, radial basis functions. A corresponding fast computational algorithm is proposed in Sect. 4 and results from two computational studies are reported in Sect. 5. Finally, Sect. 6 gives the conclusion and future research.

2 Compactly supported, positive definite, radial basis functions and their reproducing kernel Hilbert space

In general, a function $\Phi : R^d \to R$ is said to be radial if there exists a function $\psi : [0, \infty) \to R$ such that $\Phi(\mathbf{t}) = \psi(||\mathbf{t}||)$ for all $\mathbf{t} \in R^d$ with $|| \cdot ||$ the Euclidian norm on R^d. The most important properties associated with our problem concern whether radial functions are positive definite and compactly supported. In this latter regard, a function is compactly supported if it vanishes outside of some compact subset of R^d.

A function $\Phi : R^d \to R$ is called positive definite if for all $n \in N$, the set of all natural numbers, all sets of pairwise distinct centers $\{\mathbf{t}_1, \dots, \mathbf{t}_n\} \subset R^d$ and all $\boldsymbol{\alpha} \in R^n \setminus \{0\}$, the quadratic form $\sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k \Phi(\mathbf{t}_j - \mathbf{t}_k)$ is positive. We also call a univariate function $\psi : [0, \infty) \to R$ positive definite on R^d if the corresponding multivariate function $\Phi(\mathbf{t}) = \psi(||\mathbf{t}||)$ for all $\mathbf{t} \in R^d$, is positive definite.

In the case of a radial function $\Phi(\cdot) = \psi(||\cdot||)$, positive definiteness has the consequence that the interpolation matrix $\mathbf{K} = \{\Phi(\mathbf{t}_j - \mathbf{t}_k)\}_{1\leq j,\ k\leq n}$ is positive definite if the function Φ is positive definite. Furthermore, if we choose a set of pairwise distinct centers $\{\mathbf{t}_1, \ldots, \mathbf{t}_n\} \subset R^d$, we will have a set of functions $\{\Phi(\cdot - \mathbf{t}_1), \Phi(\cdot - \mathbf{t}_2), \ldots, \Phi(\cdot - \mathbf{t}_n)\}$ which are linearly independent with dimension n.

Some useful lemmas concerning radial functions are provided by the following. Lemma 1 can be found in Cressie (1993). Lemma 2 can be found in Schoenberg (1938). Lemma 3 can be found in Thomas-Agnan (1991) as a special case.

Lemma 1 *If the function Φ is positive definite on R^d, it is also positive definite on R^k with $k \leq d$.*

Lemma 2 *If the function Φ is positive definite in all dimensions, it must be completely monotonic and, hence, nonvanishing.*

We are most interested in functions that are compactly supported, positive definite and radial for R^d. Such functions do, in fact, exist and a framework for producing compactly supported radial basis functions has been developed by Wu (1995), Wendland (1995) and Buhmann (1998, 2000). In this paper we focus on Wendland functions.

Wendland (2002) shows how to obtain compactly supported, positive definite, radial basis functions $\psi_{d,k}$ that are of minimal degree k with respect to a given dimension d. He shows the functions $\psi_{d,k}$ are positive definite on R^d and are of the form

$$\psi_{d,k}(r) = \begin{cases} p_{d,k}(r), & \text{if } 0 \leq r \leq 1, \\ 0, & \text{if } 1 < r, \end{cases} \tag{1}$$

with a univariate polynomial $p_{d,k}$ of degree $\lfloor \frac{d}{2} \rfloor + 3k + 1$, where $\lfloor . \rfloor$ is the largest integer function. If we write $p_{d,k}(r) = \sum_{j=0}^{l+2k} d_{j,k}^{(l)} r^j$ with $l = \lfloor \frac{d}{2} \rfloor + k + 1$, the coefficients can be computed recursively for $0 \leq s \leq k-1$ via

$$\begin{aligned} d_{j,0}^{(l)} &= (-1)^j, \quad 0 \leq j \leq l, \\ d_{0,s+1}^{(l)} &= \sum_{j=0}^{l+2s} \frac{d_{j,s}^{(l)}}{j+2}, \quad d_{1,s+1}^{(l)=0} s \geq 0, \\ d_{j,s+1}^{(l)} &= -\frac{d_{j-2,s}^{(l)}}{j} s \geq 0, \quad 2 \leq j \leq l+2s+2. \end{aligned}$$

Wendland (2002) shows that $\Phi_{d,k}(\mathbf{t}) = \psi_{d,k}(||\mathbf{t}||)$ has strictly positive Fourier transform except when $d = 1$ and $k = 0$, in which case the Fourier transform is nonnegative and not identically zero. This property has the following consequence:

Lemma 3 *Let Φ be a radial basis function with strictly positive Fourier transform, then there is an associated reproducing kernel Hilbert space with the inner product $(f, g) = J(f, g) = \int_{R^d} W(|\mathbf{k}|)^2 \hat{f}(\mathbf{k}) \overline{\hat{g}(\mathbf{k})} d\mathbf{k}$, where $\hat{f}(\mathbf{k})$, $\hat{g}(\mathbf{k})$ and $\hat{\Phi}_{d,k}$ are the Fourier transform of f, g and $\Phi_{d,k}$ respectively and $W(|\mathbf{k}|)^2 = 1/\hat{\Phi}_{d,k}(\mathbf{k})$.*

3 Smoothing splines in higher dimensions

Consider the multivariate regression model

$$y_i = \mu(\mathbf{t}_i) + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

where $\{\varepsilon_i\}$ is a sequence of uncorrelated random variables with $\mathrm{E}(\varepsilon_i) = 0$ and $\mathrm{E}(\varepsilon_i^2) = \sigma^2$, $\mu(\cdot)$ is an unknown, smooth, regression curve and $\mathbf{t}_i \in R^d$ for all i. Data was fitted with a variant of the spline smoothing paradigm that results in an estimator of the form

$$\mu_\lambda(\mathbf{t}) = \sum_{i=1}^{n} c_i \psi(||\mathbf{t} - \mathbf{t}_i||), \tag{3}$$

where $\psi(r)$ is a compactly supported, positive definite, radial basis function. Specifically, Wendland's functions discussed in Sect. 2 have the requisite properties and are the ones we employ in our numerical work.

Let Φ be a radial basis function with strictly positive Fourier transform, the smoothing spline in (3) is obtained by minimizing

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(\mathbf{t}_i))^2 + \lambda \int\limits_{R^d} W(|\mathbf{k}|)^2 |\hat{f}(\mathbf{k})|^2 d\mathbf{k} \tag{4}$$

over all function f in the reproducing kernel Hilbert space generated by Φ from Lemma 3. The first term in (4) is the residual sum of squares which is a standard measure of goodness-of-fit for the data. The second term in (4) is a measure of smoothness as discussed in Sect. 2.

The following result characterizes the form of the minimizer of (4). Its proof can be obtained by arguments similar to those in Wahba (1990) or found in Berlinet and Thomas-Agnan (2004).

Let $\mathbf{K} = \{\Phi(\mathbf{t}_j - \mathbf{t}_k)\}_{1 \le j,k \le n} = \{\psi(||\mathbf{t}_j - \mathbf{t}_k||)\}_{1 \le j,k \le n}$. The unique minimizer of (4) is of the form (3) with $\mathbf{c} = (c_1, \ldots, c_n)^T$ being the unique solution of the following system

$$(\mathbf{K} + \lambda\mathbf{I})\mathbf{d} = \mathbf{y}, \tag{5}$$

where $\mathbf{y} = (y_1, \ldots, y_n)$ is the response vector and $\mathbf{I}$ is the identity matrix. The smoothing spline fitted values at the $\mathbf{t}$ ordinates are

$$\mu_\lambda = \mathbf{K}(\mathbf{K} + \lambda\mathbf{I})^{-1}\mathbf{y}. \tag{6}$$

4 A fast computational algorithm

In this section, we discuss how the compact support property of radial basis functions such as those developed by Wendland (2002) can be exploited to develop an efficient method for solving (5). Thus, we assume in what follows that $\psi(r) = 0$ if $|r| > D$ for some constant D.

4.1 Ordered knots and band structure

In the univariate case it is trivial to order the $\{t_i\}_{i=1}^n$ when they are distinct. However, in higher dimensional settings this step could be complicated. The goal of ordering is to organize the $\mathbf{t}$ ordinates in such a way that the interpolation matrix is banded. The approach we take is motivated by the work of Baxter et al. (1994). Alternative formulations are possible using results from sparse matrix theory by Gibbs et al. (1976).

Theorem 1 *Consider n points $(x_{i_1,i_2,\ldots,i_d})$ on a grid in R^d, where $i_1, i_2, \ldots, i_d = 1, 2, \ldots, \tilde{n}$ with $n = \tilde{n}^d$. Now take $(\mathbf{t}_k)_{k=1}^n$ to be the set of ordered points in R^d defined by*

$$
\begin{aligned}
&\mathbf{t}_1 = x_{1,1,\ldots,1}, \quad \mathbf{t}_2 = x_{1,1,\ldots,2}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}} = x_{1,1,\ldots,\tilde{n}},\\
&\mathbf{t}_{\tilde{n}+1} = x_{1,1,\ldots,2,1}, \quad \mathbf{t}_{\tilde{n}+2} = x_{1,1,\ldots,2,2}, \quad \ldots, \quad \mathbf{t}_{2\tilde{n}} = x_{1,1,\ldots,2,\tilde{n}},\\
&\qquad\qquad\qquad\qquad \cdots\\
&\mathbf{t}_{(\tilde{n}-1)*\tilde{n}+1} = x_{1,1,\ldots,\tilde{n},1}, \quad \mathbf{t}_{(\tilde{n}-1)*\tilde{n}+2} = x_{1,1,\ldots,\tilde{n},2}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}*\tilde{n}} = x_{1,1,\ldots,\tilde{n},\tilde{n}},\\
&\mathbf{t}_{\tilde{n}*\tilde{n}+1} = x_{1,1,\ldots,2,1,1}, \quad \mathbf{t}_{\tilde{n}*\tilde{n}+2} = x_{1,1,\ldots,2,1,2}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}*\tilde{n}+\tilde{n}} = x_{1,1,\ldots,2,1,\tilde{n}},\\
&\qquad\qquad\qquad\qquad \cdots\\
&\mathbf{t}_{\tilde{n}*(\tilde{n}^{d-1}-1)+1} = x_{\tilde{n},\tilde{n},\ldots,\tilde{n},1}, \quad \mathbf{t}_{\tilde{n}*(\tilde{n}^{d-1}-1)+2} = x_{\tilde{n},\tilde{n},\ldots,\tilde{n},2}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}^d} = x_{\tilde{n},\tilde{n},\ldots,\tilde{n},\tilde{n}}.
\end{aligned}
$$

Then, $(\mathbf{t}_k)_{k=1}^n$ is ordered in such a way that $\mathbf{K} = \left(\Phi(\mathbf{t}_i - \mathbf{t}_j)\right)_{i,j=1}^n$ is a banded, positive definite, symmetric matrix if Φ is a positive definite function with compact support. Let $C = \lfloor D \rfloor$, i be an integer and $0 \le i < \tilde{n}$. Then the bandwidth is $2 * (C * \tilde{n}^{d-1} + i * \tilde{n}^{d-2}) + 1$, if $\sqrt{C^2 + i^2} \le D < \sqrt{C^2 + (i+1)^2}$ with $C = \lfloor D \rfloor$.

Proof Consider the two dimensional case. Given $(x_{ij})_{i,j=1}^{\tilde{n}}$ in R^2 with $x_{ij} = (i, j)$ for all $i, j = 1, 2, \ldots, \tilde{n}$, let $n = (\tilde{n})^2$ and $(\mathbf{t}_k)_{k=1}^n$ be the set of ordered points in R^2 defined by

$$
\begin{aligned}
&\mathbf{t}_1 = x_{11}, \quad \mathbf{t}_2 = x_{12}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}} = x_{1\tilde{n}};\\
&\mathbf{t}_{\tilde{n}+1} = x_{21}, \quad \mathbf{t}_{\tilde{n}+2} = x_{22}, \quad \ldots, \quad \mathbf{t}_{2\tilde{n}} = x_{2\tilde{n}};\\
&\qquad\qquad\qquad \cdots\\
&\mathbf{t}_{(\tilde{n}-1)*\tilde{n}+1} = x_{\tilde{n}1}, \quad \mathbf{t}_{(\tilde{n}-1)*\tilde{n}+2} = x_{\tilde{n}2}, \quad \ldots, \quad \mathbf{t}_{\tilde{n}*\tilde{n}} = x_{\tilde{n}\tilde{n}}.
\end{aligned}
$$

If Φ is a positive definite function with compact support, $\mathbf{K} = \left(\Phi(\mathbf{t}_i - \mathbf{t}_j)\right)_{i,j=1}^n$ is a banded, positive definite and symmetric matrix. The bandwidth is $2 * (C * \tilde{n} + i) + 1$, if $\sqrt{C^2 + i^2} \le D < \sqrt{C^2 + (i+1)^2}$, where i is an integer, $0 \le i < \tilde{n}$ and $C = \lfloor D \rfloor$.

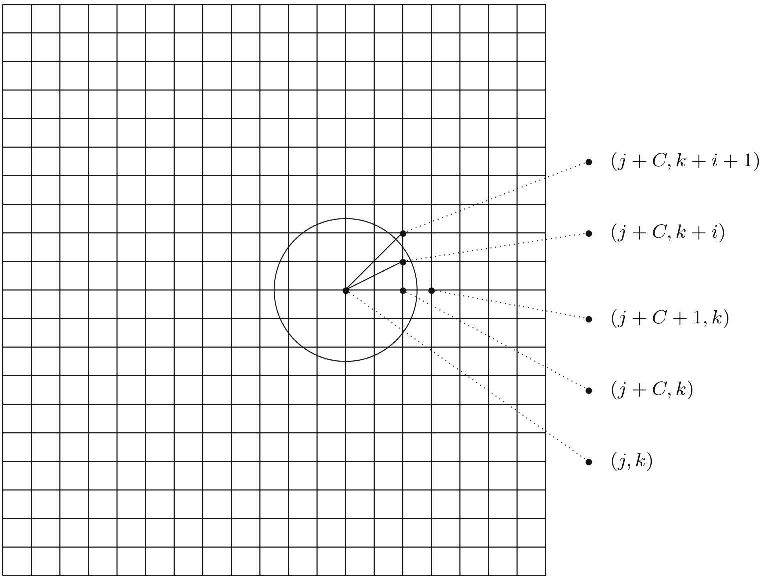

Fig. 1 Compact support in R^2 with equally spaced grid

In terms of the above formulation we have points

$$\mathbf{t}_{(j-1)*\tilde{n}+k} = (j,k) \quad \text{for } j = 1,\ldots,\tilde{n};\; k = 1,\ldots,\tilde{n}.$$

So, $\Phi(\mathbf{t}_{(j-1)*\tilde{n}+k} - \mathbf{t}_{(j'-1)*\tilde{n}+k'})$ will vanish if

$$(j-j')^2 + (k-k')^2 > D^2.$$

Figure 1 illustrates the general situation. For the fixed center $\mathbf{t}_{(j-1)*\tilde{n}+k} = (j,k)$, $\mathbf{t}_{(j-1+C)*\tilde{n}+k} = (j+C,k)$ is within the support and $\mathbf{t}_{(j+C)*\tilde{n}+k} = (j+C+1,k)$ is not within the support because $C \le D < C+1$. If the support radius is bounded by $C^2+i^2 \le D^2 < C^2+(i+1)^2$, $\mathbf{t}_{(j-1+C)*\tilde{n}+k+i} = (j+C,k+i)$ is within the support and $\mathbf{t}_{(j-1+C)*\tilde{n}+k+i+1} = (j+C,k+i+1)$ is outside the support. Furthermore, $\Phi(\mathbf{t}_{(j-1)*\tilde{n}+k} - \mathbf{t}_l) = 0$ for all $l > (j-1+C)*\tilde{n}+k+i$. There are $C*\tilde{n}+i+1$ points from the fixed center $\mathbf{t}_{(j-1)*\tilde{n}+k} = (j,k)$ up to and including $\mathbf{t}_{(j-1+C)*\tilde{n}+k+i} = (j+C,k+i)$ and symmetry gives the bandwidth of the interpolation matrix as $2*(C*\tilde{n}+i)+1$.

Proof of Theorem 1 can be obtained similarly. □

4.2 Fast algorithm for solving c and selecting λ

The band structured matrix feature derived in Theorem 1 leads to a fast algorithm for solving **c** and selecting smoothing parameter λ. We use a banded cholesky to solve (5) and Generalized cross validation method (GCV) to select λ.

The GCV was first proposed by Craven and Wahba (1979) for use in the context of nonparametric regression. In the 1980s there were numerous theoretical and practical studies which demonstrated that GCV had a variety of statistical applications (Wahba 1990). The GCV criterion can be viewed as a weighted version of the cross validation method. It can be shown that in our case the generalized cross validation criterion may be expressed as

$$\mathrm{GCV}(\lambda) = \frac{n \sum_{i=1}^{i=n} c_i^2}{(\sum_{i=1}^{i=n} h_{ii})^2}, \tag{7}$$

where h_{ii} is the ith diagonal element of $(\mathbf{K} + \lambda\mathbf{I})^{-1}$ and c_i is the ith coefficient of the estimator. Formula (7) shows that $GCV(\lambda)$ relies on c_i's and the diagonal elements of $(\mathbf{K} + \lambda\mathbf{I})^{-1}$. The fast algorithm we use for obtaining h_{ii}'s is from Hutchinson and deHoog (1985).

Now, we discuss the computational complexity of our algorithm. If the bandwidth k is fixed, the matrix is banded and we have the banded cholesky factorization, to solve (5), we only need $O(n)$ operations; to solve h_{ii}'s, we need $O(n)$ operations (Hutchinson and deHoog 1985). Note, in our case, if the $\mathbf{t}$ ordinates are on an $\tilde{n}^d$ grid (or approximately so), the bandwidth is a function of $n^{(d-1)/d}$, hence the computation is in $O(n^{3-2/d})$ operations. For the R^2 case we only need $O(n^2)$ operations compared to $O(n^3)$, the order for the thin plate multivariate smoothing splines.

5 Numerical study

In this section, we provide two computational studies from which we are able to draw some limited conclusions regarding the performance of our estimator and algorithm. The first study concerns run time comparisons in which we want to evaluate the speed of our algorithm relative to thin plate splines. The second experiment assesses the effectiveness of our estimator. We will evaluate how it performs using different sizes of compact supports and noise to signal ratios. The algorithm is implemented in C++ in both serial and parallel computing environments. Source code is available from the author upon request. R is used to generate all the figures. The test function used and reported is from Hickernell and Hon (1999). Other test functions or other radial basis functions can be implemented using similar code. The author tried a couple of other test functions and several radial basis functions but didn't find significant difference from the simulation study.

5.1 Run time comparisons

To obtain some insight into the computational gains that can be realized from exploiting band structure we want to considered two types of estimators:

– New splines proposed in this paper (NS): Splines on R^2 using positive definite, compactly supported, radial basis functions
– Thin plate splines (TPS) on R^2

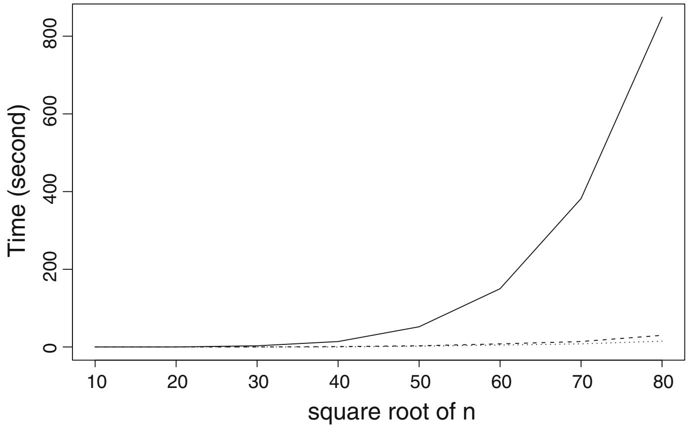

Fig. 2 Elapsed times of three estimators; *solid line* for TPS, *dotted line* for NS(25) and *dashed line* for NS(100)

To accomplish this we proceed as follows. We compute our estimator using Wendland's compactly supported, positive definite, radial basis function

$$\psi(r) = \begin{cases} (1-r)^4(4r+1), & \text{if } 0 \le r \le 1, \\ 0, & \text{if } 1 < r. \end{cases}$$

The support radius is chosen in such a way that there are approximately 100 observations within the support of the radial basis function. The simulation study in Sect. 5.2 shows that 100 observations are needed to obtain comparable estimation accuracy when using Wendland's function. From our experience with univariate cubic smoothing splines where five observations are within the support of the B-spline basis functions, we might guess that 5^2 observations within the support could be a satisfactory choice in R^2. So we also investigate the case with 25 observations within the support of the radial basis function. We use NS(100) and NS(25) to denote our estimators using different sizes of support. The elapsed time is measured in cpu seconds at various numbers of observations (equally spaced design within a square domain). Figure 2 explains the computational gains of our estimators compared to thin plate splines quite well.

5.2 Performance of the smoother

To evaluate the effectiveness of the estimators in a particular case we considered the following test function on R^2 that was used by Hickernell and Hon (1999):

$$\mu(\mathbf{t}) = \exp(-15|\mathbf{t}-(0.5,0)|^2) + 0.5 * \exp(-20|\mathbf{t}-(-0.5,0.25)|^2) \\ -0.75 * \exp(-8|\mathbf{t}-(-0.5,-0.5)|^2). \quad (8)$$

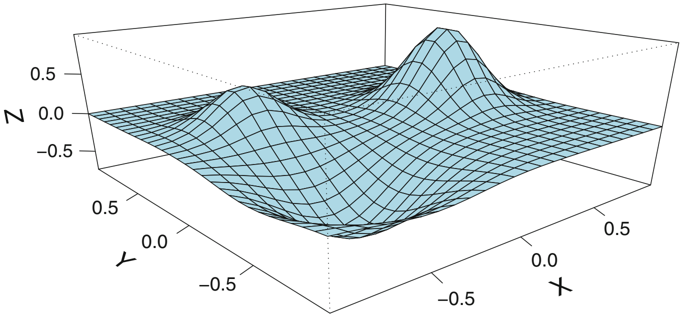

Fig. 3 Test function

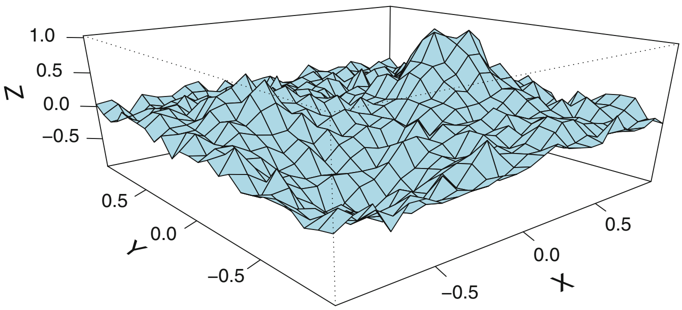

Fig. 4 Simulated Data using the test function (8) with NSR $= 0.1$ and 625 observations

The function is plotted in Fig. 3 on the square $[-1, 1] \times [-1, 1]$. We employed this function in some small scale empirical experiments described below. Note that our theory has been developed for an arbitrary positive definite, radial basis function. But when we implement this algorithm, we must make a choice for the support of the basis function. In this regard, we must be careful about the size of the support because a small support will produce an estimator with minimal smoothness regardless of the choice of λ. Another important factor that will affect the performance of the smoother is the level of noise. In this regard, we define the noise to signal ratio (NSR) as NSR $= \sigma/A$, where A is the maximum value of the true function and σ is the standard deviation of the random error terms that are added to the regression function to produce our data. Figure 4 is the picture of a simulated data set using the test function (8) with NSR $= 0.1$. This simulated data has 625 observations and the domain is square shaped. Figure 5 shows the new spline using Wendland's function and support radius 0.5 that was fitted to the data in Fig. 4 with λ selected via GCV.

To measure the performance of the smoothers, we calculated the squared error loss in estimating μ by

$$L(\lambda) = \frac{\sum_{i=1}^{n} (\mu(\mathbf{t}_i) - \mu_\lambda(\mathbf{t}_i))^2}{n}.$$

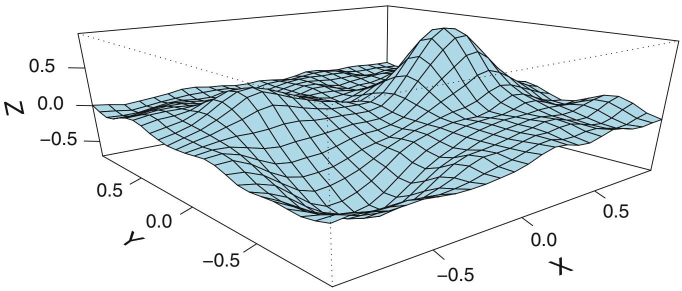

Fig. 5 Smoothing spline fit for data in Fig. 4

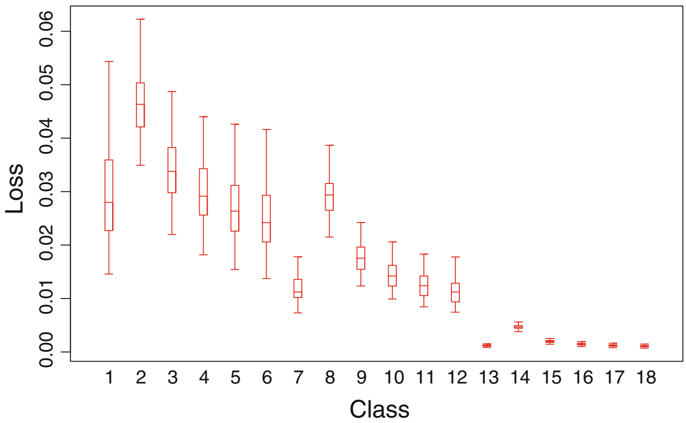

Fig. 6 Boxplot of losses (*class 1–6* for NSR $= 1$, *class 7–12* for NSR $= 0.5$, *class 13–18* for NSR $= 0.1$, *class 1, 7, 13* are losses of TPS, others are losses of NS estimators)

We generate 100 replicate random samples of size 625 for each of three NSR levels: 1, 0.5, 0.1. For each sample, we then apply a thin plate spline and a new spline with supports $D = 0.24, 0.4, 0.5, 0.6$ and 0.7 respectively. Figure 6 gives a boxplot of the losses. Class 1 to 6 are for NSR $= 1$, class 7 to 12 are for NSR $= 0.5$, class 13 to 18 are for NSR $= 0.1$. Class 1, 7 and 13 are losses for the thin plate spline and the others are losses for the new spline estimators. From the plot, we see that class 1 and 4, class 7 and 10 and class 13 and 16 perform similarly. Notice, class 4, 7 and 16 are for $D = 0.5$, where the number of observations are around 100. Also notice that as the support radius D increases, the average and sample standard deviation of the loss of new spline estimators decreases toward the levels of those of thin plate splines. This gives us some idea that the new spline could do as well from an estimation perspective as a thin plate spline. As expected, the bigger the NSR, the bigger the variation of losses. It is interesting to observe that for any NSR level the loss is very big when $D = 0.24$ (class 2, 8, 14) compared with all other cases. We believe this is the situation we mentioned previously when the support has become too small for the estimator to actually provide a smooth fit to the data.

6 Conclusions

In this paper, we develop a fast algorithm for a smoothing spline estimator in multivariate regression. Given n data points in R^2, the new algorithm has complexity $O(n^2)$ compared to $O(n^3)$, the order for the thin plate multivariate smoothing splines. Two simulation studies show that our estimator is as good as thin plate splines. Future research may consider a complete study of the effect of different type of test functions together with different type of radial basis functions on the performance of the estimator and algorithm.

Acknowledgments The author thanks Dr. Randy Eubank for his valuable guidance and assistance during the research. The author also thanks the editor and referees for their insightful and helpful comments.

References

Ahlberg JH, Nilson EN, Walsh JL (1967) The theory of splines and their applications. Academic Press, New York

Anselone PM, Laurent PJ (1968) A general method for the construction of interpolating or smoothing spline functions. Numerische Mathematik 12: 66–82

Baxter BJC, Sivakumar N, Ward JD (1994) Regarding the p-norms of radial basis interpolation matrices. Constr Approx 10: 451–468

Berlinet A, Thomas-Agnan C (2004) Reproducing kernel Hilbert spaces in probability and statistics. Kluwer, New York

Buhmann MD (1998) Radial functions on compact support. Proc Edinb Math Soc 41: 33–46

Buhmann MD (2000) A new class of radial functions with compact support. Math Comput 70: 307–318

Chui CK (1988) Multivariate splines. Society for Industrial and Applied Mathematics, Philadelphia

Craven P, Wahba G (1979) Smoothing noisy data with spline function: estimating the correct degree of smoothing by the method of cross-validation. Numerische Mathematik 3 : 377–403

Cressie N (1993) Statistics for spatial data. Wiley, Hoboken

De Boor C (1978) A practical guide to splines. Springer, New York

Duchon J (1977) Splines minimizing rotation-invariant semi-norms in sobolev spaces. Constr Theory Funct Several Var 1: 85–100

Eubank RL (1988) Spline smoothing and nonparametric regression. Marcel Dekker, New York

Gibbs NE, Poole WG, Stockmeyer PK (1976) An algorithm for reducing the bandwidth and profile of a sparse matrix. SIAM J Numer Anal 13 : 236–250

Green PJ, Silverman BW (1994) Nonparametric regression and generalized linear models. Chapman & Hall, London

Gu C (2002) Smoothing spline anova models. Springer, New York

Hickernell FJ, Hon YC (1999) Radial basis function approximations as smoothing splines. Appl Math Comput 102 : 1–24

Hutchinson MF, deHoog FR (1985) Smoothing noisy data with spline functions. Numerische Mathematik 47: 99–106

Lyche T, Schumaker LL (1973) Computation of smoothing and interpolating natural splines via local bases. SIAM J Numer Anal 10 : 1027–1038

Reinsch CH (1971) Smoothing by spline functions. II. Numerische Mathematik 16: 451–454

Schoenberg IJ (1938) Metric spaces and completely monotone functions. Ann Math 39: 811–841

Thomas-Agnan C (1991) Spline functions and stochastic filtering. Ann Stat 19: 1512–1527

Villalobos MA, Wahba G (1983) Multivariate thin plate spline estimates for the posterior probabilities in the classification problem. Commun Stat Theory Methods 12: 1449–1479

Villalobos MA, Wahba G (1987) Inequality-constrained multivariate smoothing splines with application to the estimation of posterior probabilities. J Am Stat Assoc 82: 239–248

Wahba G (1981) Spline interpolation and smoothing on the sphere. SIAM J Sci Stat Comput 2: 5–16

Wahba G (1990) Spline models for observational data. Society for Industrial and Applied Mathematics, Pennsylvania

Wahba G, Wendelberger J (1980) Some new mathematical methods for variational objective analysis using splines and cross validation. Mon Weather Rev 108: 1122–1143

Wendland H (1995) Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. Adv Comput Math 4: 389–396

Wendland H (2002) Scattered data modelling by radial and related functions. Habilitation Thesis, Mathematics, University of Gottingen, Göttingen, Germany

Whittaker ET (1923) On a new method of graduation. Proc Edinb Math Soc 41: 63–75

Wu Z (1995) Compactly supported positive definite radial functions. Adv Comput Math 4: 283–292