Proceedings of the Edinburgh Mathematical Society: page 1 of 21
doi:10.1017/S0013091524000415

FRACTAL INTERPOLANT CURVE FITTING AND REPRODUCING KERNEL HILBERT SPACES

DAH-CHIN LUOR[1] AND CHIAO-WEN LIU[2]

[1]*Department of Data Science and Analytics, I-Shou University, Kaohsiung City, Taiwan, R.O.C.*
[2]*Department of Applied Science, School of Academic Studies, R.O.C. Naval Academy, Kaohsiung City, Taiwan, R.O.C.*
Corresponding author: Dah-Chin Luor, email dclour@isu.edu.tw

(Received 8 June 2022)

Abstract In this paper, the linear space $\mathcal{F}$ of a special type of fractal interpolation functions (FIFs) on an interval I is considered. Each FIF in $\mathcal{F}$ is established from a continuous function on I. We show that, for a finite set of linearly independent continuous functions on I, we get linearly independent FIFs. Then we study a finite-dimensional reproducing kernel Hilbert space (RKHS) $\mathcal{F}_{\mathcal{B}} \subset \mathcal{F}$, and the reproducing kernel $\mathbf{k}$ for $\mathcal{F}_{\mathcal{B}}$ is defined by a basis of $\mathcal{F}_{\mathcal{B}}$. For a given data set $\mathcal{D} = \{(t_k, y_k) : k = 0, 1, \ldots, N\}$, we apply our results to curve fitting problems of minimizing the regularized empirical error based on functions of the form $f_{\mathcal{V}} + f_{\mathcal{B}}$, where $f_{\mathcal{V}} \in C_{\mathcal{V}}$ and $f_{\mathcal{B}} \in \mathcal{F}_{\mathcal{B}}$. Here $C_{\mathcal{V}}$ is another finite-dimensional RKHS of some classes of regular continuous functions with the reproducing kernel $\mathbf{k}^*$. We show that the solution function can be written in the form $f_{\mathcal{V}} + f_{\mathcal{B}} = \sum_{m=0}^{N} \gamma_m \mathbf{k}^*_{t_m} + \sum_{j=0}^{N} \alpha_j \mathbf{k}_{t_j}$, where $\mathbf{k}^*_{t_m}(\cdot) = \mathbf{k}^*(\cdot, t_m)$ and $\mathbf{k}_{t_j}(\cdot) = \mathbf{k}(\cdot, t_j)$, and the coefficients γ_m and α_j can be solved by a system of linear equations.

Keywords: fractal interpolation; reproducing kernel; reproducing kernel Hilbert space; curve fitting

2020 Mathematics subject classification: Primary 28A80
Secondary 46E22; 65D10

1. Introduction

Approximation theory is concerned with approximating complex or unknown functions by other simpler functions. The problems of approximation of functions by polynomials, splines, rational functions, trigonometric functions and wavelets have been well studied. Similarly, one of the main tasks in the problems of learning, curve fitting and pattern recognition is to develop suitable models for given data sets. In particular, curve fitting is a process of constructing a curve that has the best fit to a given data set. The theory of non-parametric curve estimations has been developed well, and many researchers have established several types of estimators, see [12–14, 16, 30] and references given in these books.

In many real-world applications, data arise from unknown functions, and a function that interpolates these data is required to be generated. The interpolation problem is finding a function f in some class of functions and interpolating those data in a given data set $\mathcal{D}$. Polynomials, splines and rational functions have been applied in interpolation methods. However, in many practical problems, sampled signals are of irregular forms, and fractal theory can provide new technologies for making complicated curves and fitting experimental data. A fractal function is a function whose graph is the attractor of an iterated function system. A fractal interpolation function (FIF) is a continuous fractal function interpolating points in a given data set. The theory of FIFs is developed for the interpolation problem with a class of fractal functions. It generalizes traditional interpolation techniques through the property of self-similarity. The concept of FIFs defined through an iterated function system was introduced by Barnsley [2, 3]. It is known that the theory of FIFs can be applied to model discrete sequences (see [19, 22, 23]). Various types of FIFs and their approximation properties have been discussed in [4, 5, 8–10, 15, 17, 20, 21, 24–27, 32–35], see also the references given in the literature.

The theory of reproducing kernel Hilbert spaces (RKHSs) has been proven to be a powerful tool in functional analysis, integral equations and learning theory. The notion of positive definite functions plays a role in reproducing kernels in RKHSs, see the excellent monographs [1, 6, 11, 28, 31]. In [7], Bouboulis and Mavroforakis constructed fractal-type reproducing kernels. They showed that the spaces of some types of FIFs are RKHSs, and the connection between FIFs and RKHSs was established.

This paper aims to discuss the RKHS consisting of FIFs further and apply such RKHS to curve fitting problems. Curve-fitting aims to obtain a suitable function that has a good approximation to the given data set. Such problems have been well studied in non-parametric regression and machine learning. Although FIFs are constructed to be interpolation functions, the theory of FIFs has many applications in approximation theory. In [19, 22, 23], FIFs were applied to model discrete sequences. In [24–26], fractal function spaces with linearly independent sets of α-fractal functions were studied. The author also discussed the role that these fractal function spaces play in approximation theory. Since the theory of RKHSs is a useful tool in approximation theory and machine learning, we are interested in RKHSs that consist of FIFs and their applications to curve fitting problems.

For a given data set, we aim to fit the data by a linear combination of linearly independent FIFs rather than a single FIF constructed from the data set directly. Moreover, we consider functions of the form $f_{\mathcal{V}} + f_{\mathcal{B}}$, where $f_{\mathcal{V}}$ belongs to an RKHS of regular continuous functions and $f_{\mathcal{B}}$ belongs to an RKHS of FIFs. Combining these two types of functions can make solutions to curve-fitting problems more general and flexible.

In § 2, the construction of a particular type of FIFs which are applied in this paper is given. Each FIF is established from a continuous function and can be treated as a fractal perturbation of that continuous function. In § 3, we prove that, for fixed parameters, the set $\mathcal{F}$ of these FIFs is a linear space, and there is a one-to-one correspondence between $C[I]$ and $\mathcal{F}$. Here, $C[I]$, defined below, is the space of continuous functions on the interval I. We also show that, for a finite set of linearly independent functions in $C[I]$, we get linearly independent FIFs. Then, for applications in curve fitting problems, a finite-dimensional RKHS $\mathcal{F}_{\mathcal{B}} \subset \mathcal{F}$ is established, and the reproducing kernel $\mathbf{k}$ for $\mathcal{F}_{\mathcal{B}}$ is defined by a basis of $\mathcal{F}_{\mathcal{B}}$. In § 4, suppose a data set $\mathcal{D} = \{(t_k, y_k) : k = 0, 1, \ldots, N\}$ is

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

given, and we aim to find a function to fit the data in $\mathcal{D}$. The type of functions considered here is the sum of a regular continuous function and an FIF. Therefore, we also consider a finite-dimensional RKHS $C_{\mathcal{V}}$ of some classes of regular continuous functions with the reproducing kernel $\mathbf{k}^*$, and we study the problem of learning a function in $C_{\mathcal{V}} \oplus \mathcal{F}_{\mathcal{B}}$ for $\mathcal{D}$ by minimizing the regularized empirical error

$$\frac{1}{N+1}\sum_{k=0}^{N}(y_k - (f_{\mathcal{V}}(t_k) + f_{\mathcal{B}}(t_k)))^2 + \lambda_1\|f_{\mathcal{V}}\|_C^2 + \lambda_2\|f_{\mathcal{B}}\|_{\mathcal{F}}^2 \tag{1.1}$$

for $f_{\mathcal{V}} \in C_{\mathcal{V}}$ and $f_{\mathcal{B}} \in \mathcal{F}_{\mathcal{B}}$ with fixed non-negative regularization parameters λ_1 and λ_2. Here $\|\cdot\|_C$ and $\|\cdot\|_{\mathcal{F}}$ are norms on $C_{\mathcal{V}}$ and $\mathcal{F}_{\mathcal{B}}$, respectively. We show that the solution of Equation (1.1) can be written in the form

$$f = f_{\mathcal{V}} + f_{\mathcal{B}} = \sum_{m=0}^{N}\gamma_m \mathbf{k}^*_{t_m} + \sum_{j=0}^{N}\alpha_j \mathbf{k}_{t_j}, \tag{1.2}$$

where $\mathbf{k}^*_{t_m}(\cdot) = \mathbf{k}^*(\cdot, t_m)$ and $\mathbf{k}_{t_j}(\cdot) = \mathbf{k}(\cdot, t_j)$, and the coefficients γ_m and α_j can be solved by a system of linear equations.

Throughout this paper, let $t_0 < t_1 < t_2 < \cdots < t_N$ and $I = [t_0, t_N]$, where N is a positive integer and $N \geq 2$. For each $k = 1, \ldots, N$, let $I_k = [t_{k-1}, t_k]$ and $J_k = [t_{j(k)}, t_{l(k)}]$. Here $j(k), l(k) \in \{0, 1, \ldots, N\}$ and $j(k) < l(k)$. To avoid trivial cases, we assume $J_k \neq I_k$. We will denote by $C[I]$ the set of all real-valued continuous functions defined on I. Define $\|f\|_\infty = \max_{t\in I}|f(t)|$ for $f \in C[I]$. For a given set of points $\mathcal{D} = \{(t_k, y_k) : k = 0, 1, \ldots, N\}$, let $C_{\mathcal{D}}[I]$ be the set of functions in $C[I]$ that interpolate all points in $\mathcal{D}$. It is known that $(C[I], \|\cdot\|_\infty)$ is a Banach space and $C_{\mathcal{D}}[I]$ is a complete metric space, where the metric is induced by $\|\cdot\|_\infty$.

2. Construction of FIFs

The approach to constructing FIFs in this section has been treated in [18]. We show the details here for readers' convenience.

Let $u \in C[I]$ and $\mathcal{D} = \{(t_k, y_k) : y_k = u(t_k), k = 0, 1, \ldots, N\}$. For $k = 1, \ldots, N$, let $L_k : J_k \to I_k$ be a homeomorphism such that $L_k(t_{j(k)}) = t_{k-1}$ and $L_k(t_{l(k)}) = t_k$, and define $M_k : J_k \times \mathbb{R} \to \mathbb{R}$ by

$$M_k(t, y) = s_k y + u(L_k(t)) - s_k p_k(t), \tag{2.1}$$

where $-1 < s_k < 1$ and p_k is a polynomial on J_k such that $p_k(t_{j(k)}) = y_{j(k)}$ and $p_k(t_{l(k)}) = y_{l(k)}$. Then $M_k(t_{j(k)}, y_{j(k)}) = y_{k-1}$, $M_k(t_{l(k)}, y_{l(k)}) = y_k$, and

$$|M_k(t, y) - M_k(t, y^*)| \leq |s_k||y - y^*| \text{ for all } t \in J_k \text{ and } y, y^* \in \mathbb{R}. \tag{2.2}$$

Define $W_k : J_k \times \mathbb{R} \to I_k \times \mathbb{R}$ by $W_k(t, y) = (L_k(t), M_k(t, y))$. For $h \in C_{\mathcal{D}}[I]$ and for each $k = 1, \ldots, N$, let $A_k = \{(t, h(t)) : t \in J_k\}$. Then $W_k(A_k) = \{(L_k(t), M_k(t, h(t))) : t \in J_k\}$. Since $L_k : J_k \to I_k$ is a homeomorphism, $W_k(A_k)$ can be written as

$$W_k(A_k) = \{(t, M_k(L_k^{-1}(t), h(L_k^{-1}(t)))) : t \in I_k\}.$$

Hence $W_k(A_k)$ is the graph of the continuous function $h_k : I_k \to \mathbb{R}$ defined by

$$h_k(t) = M_k(L_k^{-1}(t), h(L_k^{-1}(t))).$$

It is easy to see that $h_k(t_{k-1}) = y_{k-1}$ and $h_k(t_k) = y_k$. Define a mapping $T : C_{\mathcal{D}}[I] \to C_{\mathcal{D}}[I]$ by $T(h)(t) = h_k(t)$ for $t \in I_k$, that is, for $h \in C_{\mathcal{D}}[I]$ and $t \in I_k$,

$$T(h)(t) = s_k h(L_k^{-1}(t)) + u(t) - s_k p_k(L_k^{-1}(t)). \tag{2.3}$$

For $h_1, h_2 \in C_{\mathcal{D}}[I]$, we have

$$\|T(h_1) - T(h_2)\|_\infty \le s\|h_1 - h_2\|_\infty, \quad s = \max\{|s_1|, \dots, |s_N|\}.$$

Since $0 \le s < 1$, we see that T is a contraction mapping on $C_{\mathcal{D}}[I]$.

Theorem 2.1. *(Luor [18, Theorem 2.1]). The operator T given by Equation (2.3) is a contraction mapping on $C_{\mathcal{D}}[I]$.*

Definition 2.2. *The fixed point $f_{[u]}$ of T in $C_{\mathcal{D}}[I]$ is called an FIF on I corresponding to the continuous function u.*

The FIF $f_{[u]}$ given in Definition 2.2 satisfies the equation for $k = 1, \dots, N$:

$$f_{[u]}(t) = s_k \left\{ f_{[u]}(L_k^{-1}(t)) - p_k(L_k^{-1}(t)) \right\} + u(t), \quad t \in I_k. \tag{2.4}$$

If $s_k = 0$ for all k, then $f_{[u]} = u$. Therefore, $f_{[u]}$ can be treated as a fractal perturbation of u.

3. RKHSs of FIFs

3.1. Introduction to RKHSs

We give a brief introduction to RKHSs. We refer the readers to [6] and [28] for more details. Recall that a $m \times m$ real matrix $\mathbf{A} = [a_{i,j}]$ is positive semi-definite if and only if for every $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ we have that $\sum_{i,j=1}^m \alpha_i \alpha_j a_{i,j} \ge 0$. We call $\mathbf{A}$ positive definite if and only if for every $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ with $\alpha_1^2 + \dots + \alpha_m^2 \ne 0$, we have $\sum_{i,j=1}^m \alpha_i \alpha_j a_{i,j} > 0$.

Let Ω be a set. The function $\mathbf{k} : \Omega \times \Omega \to \mathbb{R}$ is positive semi-definite (definite) if for every positive integer m and every choice of distinct points $t_1, \dots, t_m$ in Ω, the matrix $[\mathbf{k}(t_i, t_j)]$ is positive semi-definite (definite). Here, we call $\mathbf{k}$ a kernel if it is symmetric and positive semi-definite. By Moore's theorem (Paulsen and Raghupathi [28, Theorem 2.14]), there exists an RKHS $\mathcal{H}$ of functions defined on Ω with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ such that $\mathbf{k}$ is the reproducing kernel for $\mathcal{H}$. For each $t \in \Omega$, define $\mathbf{k}_t(z) = \mathbf{k}(z, t)$, $z \in \Omega$. Then $\mathbf{k}_t \in \mathcal{H}$ and for $f \in \mathcal{H}$, we have $f(t) = \langle f, \mathbf{k}_t \rangle_{\mathcal{H}}$. Moreover, $\mathbf{k}(z, t) = \mathbf{k}_t(z) = \langle \mathbf{k}_t, \mathbf{k}_z \rangle_{\mathcal{H}}$ for $t, z \in \Omega$, and the set $\operatorname{span}\{\mathbf{k}_t : t \in \Omega\}$ is dense in $\mathcal{H}$.

Throughout the following subsections, we suppose that N, $t_0, \ldots, t_N$, $t_{j(1)}, \ldots, t_{j(N)}$, $t_{l(1)}, \ldots, t_{l(N)}$ and $s_1, \ldots, s_N$ are all fixed numbers, and $L_1, \ldots, L_N$ given in § 2 are fixed functions. Let $\mathcal{F}$ be the subset of $C[I]$ such that each f in $\mathcal{F}$ is an FIF corresponding to some function $u \in C[I]$ and is constructed by the approach given in § 2 with linear polynomials p_k for $k = 1, \ldots, N$.

3.2. $\mathcal{F}$ is a linear space

If $u \equiv 0$, the interpolated data set is $\{(t_k, 0) : k = 0, 1, \ldots, N\}$ and each p_k is the zero polynomial on J_k. The mapping T defined by Equation (2.3) is reduced to $T(h)(t) = s_k h(L_k^{-1}(t))$ for $h \in C_{\mathcal{D}}[I]$. The zero function is the fixed point of T and hence $f_{[u]} \equiv 0 \in \mathcal{F}$.

Suppose that $f, g \in \mathcal{F}$ and $a, b \in \mathbb{R}$. Then f and g are FIFs on I corresponding to some u and v in $C[I]$, respectively. For $t \in I_k$, we have

$$f(t) = s_k f(L_k^{-1}(t)) - s_k p_k(L_k^{-1}(t)) + u(t), \tag{3.1}$$

$$g(t) = s_k g(L_k^{-1}(t)) - s_k q_k(L_k^{-1}(t)) + v(t), \tag{3.2}$$

where p_k is a linear polynomial on J_k such that $p_k(t_{j(k)}) = u(t_{j(k)})$, $p_k(t_{l(k)}) = u(t_{l(k)})$, and q_k is a linear polynomial on J_k such that $q_k(t_{j(k)}) = v(t_{j(k)})$, $q_k(t_{l(k)}) = v(t_{l(k)})$. Then

$$(af + bg)(t) = s_k(af + bg)(L_k^{-1}(t)) - s_k(ap_k + bq_k)(L_k^{-1}(t)) + (au + bv)(t), \quad t \in I_k.$$

Since $ap_k + bq_k$ is a linear polynomial that satisfies

$$(ap_k + bq_k)(t_{j(k)}) = (au + bv)(t_{j(k)}),\ (ap_k + bq_k)(t_{l(k)}) = (au + bv)(t_{l(k)})$$

for $k = 1, \ldots, N$, we see that $af + bg$ satisfies Equation (2.4) with $f_{[u]}$, p_k, and u being replaced by $af + bg$, $ap_k + bq_k$ and $au + bv$, respectively. This shows that $af + bg$ is an FIF in $\mathcal{F}$ corresponding to the function $au + bv$, and hence $\mathcal{F}$ is a linear space.

3.3. One-to-one correspondence between $C[I]$ and $\mathcal{F}$

Note that functions in $\mathcal{F}$ only depend on functions in $C[I]$. When $u \in C[I]$ is given, the data set for interpolation, $\mathcal{D} = \{(t_k, y_k) : y_k = u(t_k), k = 0, 1, \ldots, N\}$, and all linear polynomials p_k are determined. Then, the unique FIF $f_{[u]}$ in $\mathcal{F}$ can be obtained by the approach in § 2.

Theorem 3.1. *The mapping $\Phi : C[I] \to \mathcal{F}$ defined by $\Phi(u) = f_{[u]}$ is a one-to-one and onto bounded linear mapping.*

Proof. We first show that Φ is bounded. For $u \in C[I]$ and $t \in I_k$,

$$\begin{aligned}|f_{[u]}(t) - u(t)| &= |s_k||f_{[u]}(L_k^{-1}(t)) - p_k(L_k^{-1}(t))| \\ &\leq |s_k|\left(\sup_{z\in J_k}|f_{[u]}(z)| + \sup_{z\in J_k}|p_k(z)|\right) \\ &\leq |s_k|\left(\|f_{[u]}\|_\infty + \max\{|u(t_{j(k)})|, |u(t_{l(k)})|\}\right).\end{aligned}$$

This implies that $\|f_{[u]} - u\|_\infty \leq s\|f_{[u]}\|_\infty + s\|u\|_\infty$, where $s = \max\{|s_1|, \ldots, |s_N|\}$. Then

$$\|f_{[u]}\|_\infty \leq \|f_{[u]} - u\|_\infty + \|u\|_\infty \leq s\|f_{[u]}\|_\infty + (s+1)\|u\|_\infty$$

and hence

$$\|\Phi(u)\|_\infty = \|f_{[u]}\|_\infty \leq \left(\frac{1+s}{1-s}\right)\|u\|_\infty, \qquad u \in C[I]. \tag{3.3}$$

The boundedness of Φ is obtained by Equation (3.3).

Suppose $u, v \in C[I]$ and $\Phi(u) = f_{[u]}$, $\Phi(v) = f_{[v]}$. Then $f_{[u]}$ and $f_{[v]}$ satisfy the Equations (3.1) and (3.2) with f and g being replaced by $f_{[u]}$ and $f_{[v]}$, respectively. Then, for $a, b \in \mathbb{R}$, $af_{[u]} + bf_{[v]}$ is in $\mathcal{F}$ and is constructed from the function $au + bv$. This shows that $\Phi(au + bv) = af_{[u]} + bf_{[v]} = a\Phi(u) + b\Phi(v)$, and Φ is linear.

The mapping Φ is onto since every f in $\mathcal{F}$ is constructed from a function u in $C[I]$.

In the following, we show that Φ is one-to-one. Since Φ is linear, we prove that $f_{[u]} \equiv 0$ only when $u \equiv 0$. If $f_{[u]}(t) = 0$ for all $t \in I$, then the interpolated data set is $\{(t_k, 0) : k = 0, 1, \ldots, N\}$ and then each p_k is the zero polynomial on J_k. Since $f_{[u]}$ satisfies Equation (2.4), we have $u(t) = 0$ for all $t \in I$. □

Note that $\mathcal{F}$ is a subset of $C[I]$ and each function in $\mathcal{F}$ is an FIF constructed by the approach given in §2. For fixed numbers N, $t_0, \ldots, t_N$, $t_{j(1)}, \ldots, t_{j(N)}$, $t_{l(1)}, \ldots, t_{l(N)}$, $s_1, \ldots, s_N$ and for fixed functions $L_1, \ldots, L_N$, Theorem 3.1 shows that each function $f \in \mathcal{F}$ is an FIF corresponding to a function $u \in C[I]$, and for different functions in $C[I]$, we get different FIFs.

Corollary 3.2. *If $u_1, \ldots, u_n$ are linearly independent functions in $C[I]$, then $f_{[u_1]}, \ldots, f_{[u_n]}$ are linearly independent.*

Proof. Let $\sum_{i=1}^n a_i f_{[u_i]} \equiv 0$. Since $f_{[u_i]} = \Phi(u_i)$ for each i and Φ is linear, we have $\Phi(\sum_{i=1}^n a_i u_i) \equiv 0$. This implies $\sum_{i=1}^n a_i u_i \equiv 0$ by the one-to-one property of Φ. Since $u_1, \ldots, u_n$ are linearly independent, we have $a_1 = \cdots = a_n = 0$. □

3.4. Finite-dimensional RKHSs of fractal interpolants

Suppose that an inner product $\langle\cdot,\cdot\rangle_{\mathcal{F}}$ on $\mathcal{F}$ is defined. Let $\mathcal{B} = \{\phi_0, \phi_1, \ldots, \phi_\eta\}$ be a linearly independent set of functions in $\mathcal{F}$ and let $\mathcal{F}_{\mathcal{B}}$ be the subspace $\mathcal{F}_{\mathcal{B}} = \operatorname{span}\{\phi_0, \phi_1, \ldots, \phi_\eta\}$. Then $\mathcal{F}_{\mathcal{B}}$ is a finite-dimensional Hilbert space with a basis $\mathcal{B}$.

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

Let $\mathbf{A} = [A_{i,j}]$, where $A_{i,j} = \langle \phi_i, \phi_j \rangle_{\mathcal{F}}$. By [28, Proposition 2.23], $\mathbf{A}$ is a positive definite matrix; hence, $\mathbf{A}$ is invertible. Define

$$\mathbf{k}(t', t) = \sum_{j=0}^{\eta} \sum_{m=0}^{\eta} \phi_j(t')\phi_m(t)B_{j,m}, \quad t', t \in I, \tag{3.4}$$

where the matrix $\mathbf{B} = [B_{j,m}]$ is the inverse of $\mathbf{A}$. Here, we show that $\mathbf{k}$ is positive semi-definite. Since $\mathbf{B}$ is symmetric, let $\mathbf{B}^{1/2}$ be the matrix such that $\mathbf{B}^{1/2}\mathbf{B}^{1/2} = \mathbf{B}$. Let ℓ be any positive integer and let $z_1, \ldots, z_\ell$ be any choice of distinct points in I. Let $\Psi = [\phi_i(z_j)]$. Then for any column vector $\mathbf{d} = [d_1, \ldots, d_\ell]^T$ in $\mathbb{R}^\ell$,

$$\sum_{p=1}^{\ell} \sum_{q=1}^{\ell} d_p d_q \mathbf{k}(z_p, z_q) = \mathbf{d}^T \Psi^T \mathbf{B} \Psi \mathbf{d} = (\mathbf{B}^{1/2}\Psi\mathbf{d})^T (\mathbf{B}^{1/2}\Psi\mathbf{d}) \geq 0.$$

Let $\mathbf{k}_t(\cdot) = \mathbf{k}(\cdot, t)$ and we write $\mathbf{k}_t$ in the form

$$\mathbf{k}_t(\cdot) = \sum_{j=0}^{\eta} \left(\sum_{m=0}^{\eta} \phi_m(t)B_{j,m} \right) \phi_j(\cdot). \tag{3.5}$$

Then for $f = \sum_{k=0}^{\eta} a_k\phi_k \in \mathcal{F}_{\mathcal{B}}$ and $t \in I$,

$$\begin{aligned}\langle f, \mathbf{k}_t \rangle_{\mathcal{F}} &= \sum_{k=0}^{\eta} \sum_{j=0}^{\eta} a_k \left(\sum_{m=0}^{\eta} \phi_m(t)B_{j,m} \right) \langle \phi_k, \phi_j \rangle_{\mathcal{F}} \\ &= \sum_{k=0}^{\eta} \sum_{m=0}^{\eta} a_k \phi_m(t) \left(\sum_{j=0}^{\eta} A_{k,j}B_{j,m} \right) = \sum_{k=0}^{\eta} a_k\phi_k(t) = f(t).\end{aligned} \tag{3.6}$$

We also have $\mathbf{k}(t', t) = \mathbf{k}_t(t') = \langle \mathbf{k}_t, \mathbf{k}_{t'} \rangle_{\mathcal{F}}$ for $t, t' \in I$.

Theorem 3.3. *The space $\mathcal{F}_{\mathcal{B}}$ is a finite-dimensional RKHS with the reproducing kernel $\mathbf{k}$ defined by Equation (3.4).*

By Equation (3.5), we have

$$\mathbf{k}_{t_i} = \sum_{j=0}^{\eta} \left(\sum_{m=0}^{\eta} \phi_m(t_i)B_{j,m} \right) \phi_j, \qquad i = 0, 1, \ldots, N. \tag{3.7}$$

In general, $\mathbf{k}_{t_0}, \ldots, \mathbf{k}_{t_N}$ may not be linearly independent. Since each $\mathbf{k}_{t_i}$ is a function in $\mathcal{F}_{\mathcal{B}}$, $\mathbf{k}_{t_0}, \ldots, \mathbf{k}_{t_N}$ are linearly dependent when $\eta < N$.

Proposition 3.4. *Let $\mathbb{K} = [\mathbf{k}(t_i, t_j)]$ and $\Psi_i = [\phi_0(t_i), \phi_1(t_i), \ldots, \phi_\eta(t_i)]^T$ for $i, j = 0, \ldots, N$. Then, the following three statements are equivalent.*

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

(1) $\mathbf{k}_{t_0}, \ldots, \mathbf{k}_{t_N}$ *are linearly independent.*
(2) *The column vectors* $\Psi_0, \Psi_1, \ldots, \Psi_N$ *are linearly independent.*
(3) *The matrix* $\mathbb{K}$ *is positive definite.*

Proof. We first show that statements (1) and (2) are equivalent. Suppose that

$$c_0\mathbf{k}_{t_0} + \cdots + c_N\mathbf{k}_{t_N} = \sum_{j=0}^{\eta}\left(\sum_{i=0}^{N}\sum_{m=0}^{\eta} c_i B_{j,m}\phi_m(t_i)\right)\phi_j \equiv 0.$$

Since $\phi_0, \ldots, \phi_\eta$ are linearly independent, we have $\mathbf{B}\Psi\mathbf{C} = \mathbf{0}$, where $\mathbf{B} = [B_{j,m}]$, $\mathbf{C} = [c_0, \ldots, c_N]^T$, $\mathbf{0}$ is the zero column vector, and $\Psi = [\phi_m(t_i)]$ is the matrix with column vectors Ψ_i, $i = 0, \ldots, N$. Since $\mathbf{B}$ is invertible, we have $\Psi\mathbf{C} = \mathbf{0}$. Therefore, $\mathbf{k}_{t_0}, \ldots, \mathbf{k}_{t_N}$ are linearly independent if and only if the equation $\Psi\mathbf{C} = \mathbf{0}$ has only one solution, $c_i = 0$ for $i = 0, \ldots, N$, if and only if $\Psi_0, \Psi_1, \ldots, \Psi_N$ are linearly independent.

The following shows that statements (1) and (3) are equivalent. Since $\mathbf{k}(t_i, t_j) = \langle \mathbf{k}_{t_j}, \mathbf{k}_{t_i}\rangle_{\mathcal{F}}$ for $i, j = 0, \ldots, N$, we see that, for $\alpha_0, \ldots, \alpha_N \in \mathbb{R}$,

$$\sum_{i=0}^{N}\sum_{j=0}^{N}\alpha_i\alpha_j\mathbf{k}(t_i, t_j) = \left\langle \sum_{j=0}^{N}\alpha_j\mathbf{k}_{t_j}, \sum_{i=0}^{N}\alpha_i\mathbf{k}_{t_i}\right\rangle_{\mathcal{F}} = \left\|\sum_{j=0}^{N}\alpha_j\mathbf{k}_{t_j}\right\|_{\mathcal{F}}^2 \geq 0. \tag{3.8}$$

The matrix $\mathbb{K}$ is positive definite if and only if for every $\alpha_0, \ldots, \alpha_N \in \mathbb{R}$ with $\alpha_0^2 + \cdots + \alpha_N^2 \neq 0$, we have $\sum_{i=0}^{N}\sum_{j=0}^{N}\alpha_i\alpha_j\mathbf{k}(t_i, t_j) > 0$. Then statements (1) and (3) are equivalent and can be obtained by Equation (3.8). $\square$

If N is large, functions $\mathbf{k}_{t_0}, \ldots, \mathbf{k}_{t_N}$ are usually linearly dependent. In the following, we investigate the dependence of these functions. Let $u_j = \Phi^{-1}(\phi_j)$ for $j = 0, 1, \ldots, \eta$ and $\mathcal{U}_i = [u_0(t_i), u_1(t_i), \ldots, u_\eta(t_i)]^T$ for $i = 0, 1, \ldots, N$. Since $u_j(t_i) = \phi_j(t_i)$, we have $\mathcal{U}_i = \Psi_i$ for each i, where Ψ_i is given in Proposition 3.4.

Proposition 3.5. *If* $\mathbf{k}_{t_\delta} = \sum_{i=1}^{s}\beta_i\mathbf{k}_{t_{r(i)}}$, *where* $\beta_i \neq 0$, $\delta, r(i) \in \{0, 1, \ldots, N\}$ *and* $r(i) \neq \delta$ *for* $i = 1, \ldots, s$, *then* $\mathcal{U}_\delta = \sum_{i=1}^{s}\beta_i\mathcal{U}_{r(i)}$. *The converse is also true.*

Proof. By Equation (3.5), we have

$$\begin{aligned}\mathbf{k}_{t_\delta} &= \sum_{j=0}^{\eta}\left(\sum_{m=0}^{\eta}\phi_m(t_\delta)B_{j,m}\right)\phi_j = \sum_{i=1}^{s}\beta_i\left\{\sum_{j=0}^{\eta}\left(\sum_{m=0}^{\eta}\phi_m(t_{r(i)})B_{j,m}\right)\phi_j\right\}\\ &= \sum_{j=0}^{\eta}\left(\sum_{m=0}^{\eta}\sum_{i=1}^{s}\beta_i\phi_m(t_{r(i)})B_{j,m}\right)\phi_j.\end{aligned}$$

This implies

$$\sum_{j=0}^{\eta}\left\{\sum_{m=0}^{\eta}\phi_m(t_\delta)B_{j,m} - \sum_{m=0}^{\eta}\sum_{i=1}^{s}\beta_i\phi_m(t_{r(i)})B_{j,m}\right\}\phi_j = 0. \tag{3.9}$$

Since $\phi_0, \ldots, \phi_\eta$ are linearly independent, we have

$$\sum_{m=0}^{\eta}\left(\phi_m(t_\delta) - \sum_{i=1}^{s}\beta_i\phi_m(t_{r(i)})\right)B_{j,m} = 0, \quad j = 0, 1, \ldots, \eta. \tag{3.10}$$

Then

$$\mathbf{B}(\Psi_\delta - [\Psi_{r(1)}, \ldots, \Psi_{r(s)}]\beta) = \mathbf{0},$$

where $\mathbf{B} = [B_{j,m}]$, $\beta = [\beta_1, \ldots, \beta_s]^T$, and $\mathbf{0}$ is the zero column vector. Here $\Psi_\ell = [\phi_0(t_\ell), \phi_1(t_\ell), \ldots, \phi_\eta(t_\ell)]^T$ for $\ell = \delta, r(1), \ldots, r(s)$. Since $\mathbf{B}$ is invertible, we have $\Psi_\delta = \sum_{i=1}^{s}\beta_i\Psi_{r(i)}$. The equalities $\mathcal{U}_i = \Psi_i$ for each i show that $\mathcal{U}_\delta = \sum_{i=1}^{s}\beta_i\mathcal{U}_{r(i)}$.

Conversely, if $\mathcal{U}_\delta = \sum_{i=1}^{s}\beta_i\mathcal{U}_{r(i)}$, then Equation (3.10) holds and we have Equation (3.9). This implies $\mathbf{k}_{t_\delta} = \sum_{i=1}^{s}\beta_i\mathbf{k}_{t_{r(i)}}$. □

4. Curve fitting problems

Let $\mathcal{D} = \{(t_k, y_k) : k = 0, 1, \ldots, N\}$ be a given data set. Suppose that $t_{j(1)}, \ldots, t_{j(N)}$, $t_{l(1)}, \ldots, t_{l(N)}$ and $s_1, \ldots, s_N$ are all fixed numbers, and $L_1, \ldots, L_N$ given in § 2 are fixed functions. Let $\mathcal{F}$ be the space of FIFs that are constructed by the approach given in § 2 with a function $u \in C[I]$, where $I = [t_0, t_N]$, and linear polynomials p_k on $J_k = [t_{j(k)}, t_{l(k)}]$ such that $p_k(t_{j(k)}) = u(t_{j(k)})$ and $p_k(t_{l(k)}) = u(t_{l(k)})$ for $k = 1, \ldots, N$. Let $\mathcal{B} = \{\phi_0, \phi_1, \ldots, \phi_\eta\}$ be a linearly independent set of functions in $\mathcal{F}$ and let $\mathcal{F}_\mathcal{B} = \text{span}\{\phi_0, \phi_1, \ldots, \phi_\eta\}$. Suppose that an inner product $\langle\cdot,\cdot\rangle_\mathcal{F}$ on $\mathcal{F}$ is defined. Theorem 3.3 shows that $\mathcal{F}_\mathcal{B}$ is a finite-dimensional RKHS with a basis $\mathcal{B}$, and the reproducing kernel $\mathbf{k}$ is given by Equation (3.4).

Let $\mathcal{V} = \{v_0, v_1, \ldots, v_\xi\}$ be a linearly independent set of functions in $C[I]$ such that $\mathcal{B} \cup \mathcal{V}$ is also linearly independent. Let $C_\mathcal{V} = \text{span}\{v_0, v_1, \ldots, v_\xi\}$. Suppose that an inner product $\langle\cdot,\cdot\rangle_C$ on $C[I]$ is defined. By a similar approach given in § 3.4, we see that $C_\mathcal{V}$ is also a finite-dimensional RKHS with the kernel $\mathbf{k}^*$ defined by

$$\mathbf{k}^*(t', t) = \sum_{j=0}^{\xi}\sum_{m=0}^{\xi} v_j(t')v_m(t)B^*_{j,m}, \quad t', t \in I, \tag{4.1}$$

where the matrix $[B^*_{j,m}]$ is the inverse of $A^* = [\langle v_i, v_j\rangle_C]$.

Consider the problem of learning a function in $C_\mathcal{V} \oplus \mathcal{F}_\mathcal{B}$ from $\mathcal{D}$ by minimizing the regularized empirical error

$$\frac{1}{N+1}\sum_{k=0}^{N}(y_k - (f_\mathcal{V}(t_k) + f_\mathcal{B}(t_k)))^2 + \lambda_1\|f_\mathcal{V}\|_C^2 + \lambda_2\|f_\mathcal{B}\|_\mathcal{F}^2 \tag{4.2}$$

for $f_\mathcal{V} \in C_\mathcal{V}$ and $f_\mathcal{B} \in \mathcal{F}_\mathcal{B}$ with fixed non-negative regularization parameters λ_1 and λ_2. Here $\|f_\mathcal{V}\|_C^2 = \langle f_\mathcal{V}, f_\mathcal{V}\rangle_C$ and $\|f_\mathcal{B}\|_\mathcal{F}^2 = \langle f_\mathcal{B}, f_\mathcal{B}\rangle_\mathcal{F}$. The function $f_\mathcal{V} + f_\mathcal{B}$ is in $C_\mathcal{V} \oplus \mathcal{F}_\mathcal{B}$, and the first item of Equation (4.2) is the empirical mean squared error. It is often to add a complexity penalty item to the objective function to avoid overfitting. In the RKHS approach, we may choose the squared norm of functions in RKHS as the penalty item. In this paper, the penalty item is given by $\lambda_1\|f_{\mathcal{V}}\|_C^2+\lambda_2\|f_{\mathcal{B}}\|_{\mathcal{F}}^2$. $\|f_{\mathcal{V}}\|_C^2$ and $\|f_{\mathcal{B}}\|_{\mathcal{F}}^2$ measure the complexity of the functions, and λ_1 and λ_2 control the strength of the complexity penalty. If $\lambda_1=\lambda_2=0$, minimizing Equation (4.2) is reduced to the least mean squared error problem. If there exists an interpolation function for $\mathcal{D}$ in $C_{\mathcal{V}}\oplus\mathcal{F}_{\mathcal{B}}$, it is a solution for the least mean squared error problem, and the empirical mean squared error is equal to 0.

Let $\mathbf{k}$ and $\mathbf{k}^*$ be the kernels defined by Equations (3.4) and (4.1), respectively. Let $\mathcal{F}_{\mathcal{D}}=\operatorname{span}\{\mathbf{k}_{t_0},\mathbf{k}_{t_1},\ldots,\mathbf{k}_{t_N}\}$ and $C_{\mathcal{D}}^*=\operatorname{span}\{\mathbf{k}^*_{t_0},\mathbf{k}^*_{t_1},\ldots,\mathbf{k}^*_{t_N}\}$, where $\mathbf{k}_{t_i}$ is given by Equation (3.7) and

$$\mathbf{k}^*_{t_i}=\sum_{j=0}^{\xi}\left(\sum_{m=0}^{\xi}v_m(t_i)B^*_{j,m}\right)v_j,\qquad i=0,1,\ldots,N. \tag{4.3}$$

We see that $\mathcal{F}_{\mathcal{D}}$ is a subspace of $\mathcal{F}_{\mathcal{B}}$ and $C^*_{\mathcal{D}}$ is a subspace of $C_{\mathcal{V}}$.

For $f_{\mathcal{B}}\in\mathcal{F}_{\mathcal{B}}$, let $\mathbf{P}_{\mathcal{F}_{\mathcal{D}}}(f_{\mathcal{B}})$ be the orthogonal projection of $f_{\mathcal{B}}$ on $\mathcal{F}_{\mathcal{D}}$. Then

$$f_{\mathcal{B}}(t_i)-\mathbf{P}_{\mathcal{F}_{\mathcal{D}}}(f_{\mathcal{B}})(t_i)=\langle f_{\mathcal{B}}-\mathbf{P}_{\mathcal{F}_{\mathcal{D}}}(f_{\mathcal{B}}),\mathbf{k}_{t_i}\rangle_{\mathcal{F}}=0,\quad i=0,\ldots,N,$$

and hence $f_{\mathcal{B}}(t_i)=\mathbf{P}_{\mathcal{F}_{\mathcal{D}}}(f_{\mathcal{B}})(t_i)$ for $i=0,\ldots,N$. Similarly, for $f_{\mathcal{V}}\in C_{\mathcal{V}}$, $f_{\mathcal{V}}(t_i)=\mathbf{P}_{C^*_{\mathcal{D}}}(f_{\mathcal{V}})(t_i)$ for each i, where $\mathbf{P}_{C^*_{\mathcal{D}}}(f_{\mathcal{V}})$ is the orthogonal projection of $f_{\mathcal{V}}$ on $C^*_{\mathcal{D}}$. Since $\|\mathbf{P}_{\mathcal{F}_{\mathcal{D}}}(f_{\mathcal{B}})\|_{\mathcal{F}}\le\|f_{\mathcal{B}}\|_{\mathcal{F}}$ and $\|\mathbf{P}_{C^*_{\mathcal{D}}}(f_{\mathcal{V}})\|_C\le\|f_{\mathcal{V}}\|_C$, we see that if a function $f_{\mathcal{V}}+f_{\mathcal{B}}$ minimizes the regularized empirical error given in Equation (4.2), where $f_{\mathcal{V}}\in\mathcal{C}_{\mathcal{V}}$ and $f_{\mathcal{B}}\in\mathcal{F}_{\mathcal{B}}$, then $f_{\mathcal{V}}\in\mathcal{C}^*_{\mathcal{D}}$ and $f_{\mathcal{B}}\in\mathcal{F}_{\mathcal{D}}$. Therefore, a solution of Equation (4.2) can be written in the form

$$f=f_{\mathcal{V}}+f_{\mathcal{B}}=\sum_{m=0}^{N}\gamma_m\mathbf{k}^*_{tm}+\sum_{j=0}^{N}\alpha_j\mathbf{k}_{t_j}. \tag{4.4}$$

This implies

$$f(t_i)=\sum_{j=0}^{N}\gamma_j\mathbf{k}^*(t_i,t_j)+\sum_{j=0}^{N}\alpha_j\mathbf{k}(t_i,t_j)\qquad i=0,1,\ldots,N,$$

and

$$\|f_{\mathcal{V}}\|_C^2=\langle f_{\mathcal{V}},f_{\mathcal{V}}\rangle_C=\left\langle\sum_{j=0}^{N}\gamma_j\mathbf{k}^*_{t_j},\sum_{i=0}^{N}\gamma_i\mathbf{k}^*_{t_i}\right\rangle_C=\sum_{j=0}^{N}\sum_{i=0}^{N}\gamma_j\gamma_i\mathbf{k}^*(t_i,t_j).$$

$$\|f_{\mathcal{B}}\|_{\mathcal{F}}^2=\langle f_{\mathcal{B}},f_{\mathcal{B}}\rangle_{\mathcal{F}}=\left\langle\sum_{j=0}^{N}\alpha_j\mathbf{k}_{t_j},\sum_{i=0}^{N}\alpha_i\mathbf{k}_{t_i}\right\rangle_{\mathcal{F}}=\sum_{j=0}^{N}\sum_{i=0}^{N}\alpha_j\alpha_i\mathbf{k}(t_i,t_j).$$

Then, Equation (4.2) can be reduced to

$$\frac{1}{N+1}\sum_{i=0}^{N}\Big(y_i-\sum_{j=0}^{N}\gamma_j\mathbf{k}^*(t_i,t_j)-\sum_{j=0}^{N}\alpha_j\mathbf{k}(t_i,t_j)\Big)^2 \\ +\lambda_1\sum_{j=0}^{N}\sum_{i=0}^{N}\gamma_j\gamma_i\mathbf{k}^*(t_i,t_j)+\lambda_2\sum_{j=0}^{N}\sum_{i=0}^{N}\alpha_j\alpha_i\mathbf{k}(t_i,t_j). \tag{4.5}$$

It is not hard to see that the solutions $\{\gamma_j\}$ and $\{\alpha_j\}$ that minimize Equation (4.5) satisfy the following equations

$$\sum_{i=0}^{N}\Big(\sum_{j=0}^{N}\gamma_j\mathbf{k}^*(t_i,t_j)+\sum_{j=0}^{N}\alpha_j\mathbf{k}(t_i,t_j)-y_i\Big)\mathbf{k}^*(t_i,t_\ell) \\ +\lambda_1(N+1)\sum_{i=0}^{N}\gamma_i\mathbf{k}^*(t_i,t_\ell)=0,$$

$$\sum_{i=0}^{N}\Big(\sum_{j=0}^{N}\gamma_j\mathbf{k}^*(t_i,t_j)+\sum_{j=0}^{N}\alpha_j\mathbf{k}(t_i,t_j)-y_i\Big)\mathbf{k}(t_i,t_n) \\ +\lambda_2(N+1)\sum_{i=0}^{N}\alpha_i\mathbf{k}(t_i,t_n)=0,$$

for $\ell, n = 0,1,\ldots,N$. Let $\mathbf{D}=[\gamma_0,\gamma_1,\ldots,\gamma_N]^T$, $\mathbf{C}=[\alpha_0,\alpha_1,\ldots,\alpha_N]^T$, $\mathbb{Y}=[y_0,y_1,\ldots,y_N]^T$, $\mathbb{K}=[\mathbf{k}(t_i,t_j)]$, $\mathbb{K}^*=[\mathbf{k}^*(t_i,t_j)]$, and let $\mathbf{0}$ be the zero column matrix. We can write the equations in the matrix forms

$$\mathbb{K}^*(\mathbb{K}^*\mathbf{D}+\mathbb{K}\mathbf{C}-\mathbb{Y})+\lambda_1(N+1)\mathbb{K}^*\mathbf{D}=\mathbf{0}, \tag{4.6}$$

$$\mathbb{K}(\mathbb{K}^*\mathbf{D}+\mathbb{K}\mathbf{C}-\mathbb{Y})+\lambda_2(N+1)\mathbb{K}\mathbf{C}=\mathbf{0}. \tag{4.7}$$

Putting the two matrix equations together, we have

$$\begin{bmatrix}\mathbb{K}^* & \mathbf{0}\\ \mathbf{0} & \mathbb{K}\end{bmatrix}\left(\begin{bmatrix}\mathbb{K}^* & \mathbb{K}\\ \mathbb{K}^* & \mathbb{K}\end{bmatrix}+(N+1)\begin{bmatrix}\lambda_1\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \lambda_2\mathbf{I}\end{bmatrix}\right)\begin{bmatrix}\mathbf{D}\\ \mathbf{C}\end{bmatrix}=\begin{bmatrix}\mathbb{K}^* & \mathbf{0}\\ \mathbf{0} & \mathbb{K}\end{bmatrix}\begin{bmatrix}\mathbb{Y}\\ \mathbb{Y}\end{bmatrix}. \tag{4.8}$$

Here $\mathbf{I}$ is the $(N+1)\times(N+1)$ identity matrix and $\mathbf{0}$ is the $(N+1)\times(N+1)$ zero matrix. If $[\mathbf{D}^T\,\mathbf{C}^T]^T$ is a column matrix that satisfies Equation (4.8), then

$$f=[\mathbf{k}^*_{t_0},\ldots,\mathbf{k}^*_{t_N}]\mathbf{D}+[\mathbf{k}_{t_0},\ldots,\mathbf{k}_{t_N}]\mathbf{C} \tag{4.9}$$

is a function in $C_{\mathcal{V}}\oplus\mathcal{F}_{\mathcal{B}}$ that minimizes Equation (4.2).

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

If $\mathbb{K}^*$ and $\mathbb{K}$ are invertible, then Equations (4.6) and (4.7) imply $\lambda_1\mathbf{D} = \lambda_2\mathbf{C}$. In the case that $\lambda_2 = 0$ and $\lambda_1 \neq 0$, we have $\gamma_m = 0$ for $m = 0, \ldots, N$, and $f \in \mathcal{F}_\mathcal{B}$. Similarly, in the case that $\lambda_1 = 0$ and $\lambda_2 \neq 0$, we have $\alpha_j = 0$ for $j = 0, \ldots, N$, and $f \in C_\mathcal{V}$. If $\lambda_1 = \lambda_2 \neq 0$, then $\mathbf{D} = \mathbf{C}$.

The empirical error defined in Equation (4.5) is equal to

$$\begin{aligned}
&\frac{1}{N+1}\|\mathbb{Y} - \mathbb{K}^*\mathbf{D} - \mathbb{K}\mathbf{C}\|_2^2 + \lambda_1\mathbf{D}^T\mathbb{K}^*\mathbf{D} + \lambda_2\mathbf{C}^T\mathbb{K}\mathbf{C} \\
&\quad = \frac{1}{N+1}(\mathbb{Y}^T - \mathbf{D}^T\mathbb{K}^* - \mathbf{C}^T\mathbb{K})(\mathbb{Y} - \mathbb{K}^*\mathbf{D} - \mathbb{K}\mathbf{C}) + \lambda_1\mathbf{D}^T\mathbb{K}^*\mathbf{D} + \lambda_2\mathbf{C}^T\mathbb{K}\mathbf{C} \\
&\quad = \frac{1}{N+1}\mathbb{Y}^T(\mathbb{Y} - \mathbb{K}^*\mathbf{D} - \mathbb{K}\mathbf{C}).
\end{aligned} \tag{4.10}$$

The last equality is based on Equations (4.6) and (4.7). If $\mathbb{K}^*$ is invertible, then Equation (4.6) implies that $\mathbb{Y} - \mathbb{K}^*\mathbf{D} - \mathbb{K}\mathbf{C} = \lambda_1(N+1)\mathbf{D}$, and Equation (4.10) can be reduced to $\lambda_1\mathbb{Y}^T\mathbf{D}$. Similarly, if $\mathbb{K}$ is invertible, then Equation (4.7) implies that $\mathbb{Y} - \mathbb{K}^*\mathbf{D} - \mathbb{K}\mathbf{C} = \lambda_2(N+1)\mathbf{C}$, and Equation (4.10) can be reduced to $\lambda_2\mathbb{Y}^T\mathbf{C}$.

Let Φ be the operator given in Theorem 3.1. For the basis $\mathcal{B} = \{\phi_0, \phi_1, \ldots, \phi_\eta\}$ of the RKHS $\mathcal{F}_\mathcal{B}$, let $\mathcal{U} = \{u_i : u_i = \Phi^{-1}(\phi_i), i = 0, 1, \ldots, \eta\}$. Let $C_\mathcal{U} = \operatorname{span}\{u_0, u_1, \ldots, u_\eta\}$. If we choose $\langle\cdot,\cdot\rangle_C$ to be an inner product on $C_\mathcal{U}$, then $C_\mathcal{U}$ is a finite-dimensional RKHS with the kernel $\tilde{\mathbf{k}}$ defined by

$$\tilde{\mathbf{k}}(t', t) = \sum_{m=0}^{\eta}\sum_{j=0}^{\eta} u_m(t)u_j(t')\tilde{B}_{j,m}, \quad t, t' \in I, \tag{4.11}$$

where the matrix $[\tilde{B}_{j,m}]$ is the inverse of $\tilde{A} = [\langle u_i, u_j\rangle_C]$. Consider the problem of learning a function in $C_\mathcal{V} \oplus C_\mathcal{U}$ by minimizing the regularized empirical error

$$\frac{1}{N+1}\sum_{k=0}^{N}(y_k - (f_\mathcal{V}(t_k) + u(t_k)))^2 + \lambda_1\|f_\mathcal{V}\|_C^2 + \lambda_2\|u\|_C^2, \tag{4.12}$$

where $f_\mathcal{V} \in \mathcal{C}_\mathcal{V}$ and $u \in C_\mathcal{U}$. In the following, we show that if we define $\langle f, g\rangle_\mathcal{F} = \langle\Phi^{-1}(f), \Phi^{-1}(g)\rangle_C$ for $f, g \in \mathcal{F}$, and if $f_\mathcal{V} + u$ minimizes Equation (4.12), then $f_\mathcal{V} + \Phi(u)$ minimizes Equation (4.2). Suppose that $f_\mathcal{V} + u$ minimizes Equation (4.12). Then by a similar approach, u can be written in the form $u = \sum_{i=0}^{N}\tilde{\alpha}_i\tilde{\mathbf{k}}_{t_i}$, where

$$\tilde{\mathbf{k}}_{t_i} = \sum_{j=0}^{\eta}\left(\sum_{m=0}^{\eta} u_m(t_i)\tilde{B}_{j,m}\right)u_j, \qquad i = 0, 1, \ldots, N, \tag{4.13}$$

and the vectors $\mathbf{D}$ and $\tilde{\mathbf{C}} = [\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_N]^T$ satisfy

$$\begin{bmatrix}\mathbb{K}^* & \mathbf{0}\\ \mathbf{0} & \tilde{\mathbb{K}}\end{bmatrix}\left(\begin{bmatrix}\mathbb{K}^* & \tilde{\mathbb{K}}\\ \mathbb{K}^* & \tilde{\mathbb{K}}\end{bmatrix} + (N+1)\begin{bmatrix}\lambda_1\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \lambda_2\mathbf{I}\end{bmatrix}\right)\begin{bmatrix}\mathbf{D}\\ \tilde{\mathbf{C}}\end{bmatrix} = \begin{bmatrix}\mathbb{K}^* & \mathbf{0}\\ \mathbf{0} & \tilde{\mathbb{K}}\end{bmatrix}\begin{bmatrix}\mathbb{Y}\\ \mathbb{Y}\end{bmatrix}, \tag{4.14}$$

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

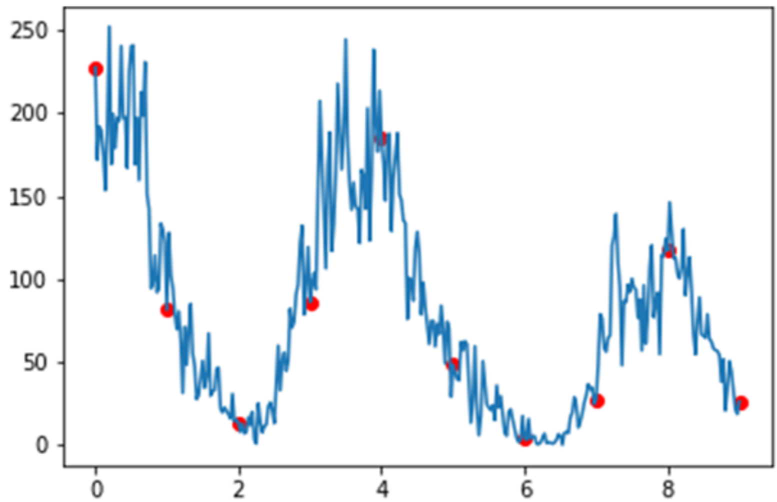

Figure 1. Monthly mean total sunspot numbers.

where $\tilde{\mathbb{K}} = [\tilde{\mathbf{k}}(t_i, t_j)]$. Since $A_{m,j} = \langle \phi_m, \phi_j \rangle_{\mathcal{F}} = \langle u_m, u_j \rangle_C = \tilde{A}_{m,j}$, $A = \tilde{A}$ and then $[B_{j,m}] = [\tilde{B}_{j,m}]$. By $u_m(t_i) = \phi_m(t_i)$, $m = 0, 1, \ldots, \eta$ and $i = 0, 1, \ldots, N$, we have $\Phi(\tilde{\mathbf{k}}_{t_i}) = \mathbf{k}_{t_i}$ for each i. By Equations (3.4) and (4.11), $\mathbf{k}(t_i, t_j) = \tilde{\mathbf{k}}(t_i, t_j)$ for each i, j and hence $\mathbb{K} = \tilde{\mathbb{K}}$. This implies that Equations (4.14) and (4.8) have the same solutions. Note that $\|\Phi(u)\|_{\mathcal{F}}^2 = \|u\|_C^2$ and $\Phi(u)(t_k) = u(t_k)$ for $k = 0, 1, \ldots, N$. Therefore, if $f_{\mathcal{V}} + u$ minimizes Equation (4.12), then $f_{\mathcal{V}} + \Phi(u)$ minimizes Equation (4.2). In general, this conclusion may be false if we define $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ in another way.

Remark 4.1. Gaussian process regression is a Bayesian-based machine learning model that produces a posterior distribution for an unknown regression function [29]. The positive semi-definite kernel $\mathbf{k}$ defined by Equation (3.4) can be applied to Gaussian process regression as the covariance function. Applications of FIFs and $\mathbf{k}$ to the Gaussian process are interesting and valuable directions for future research.

5. An example

5.1. Data description

The monthly mean total sunspot number is obtained by taking the arithmetic mean of the daily total sunspot numbers over all days of each calendar month. The data set we used in our example is the series of monthly mean total sunspot numbers from 1990/01 *to* 2017/01*. There are* 325 *data in total. These data are open and available on the webpage https://www.sidc.be/SILSO/datafiles.*

We choose the monthly mean total sunspot numbers for 1990/01, 1993/01,..., 2014/01, and 2017/01 as our data in $\mathcal{D}$. To simplify our example, we set $\mathcal{D} = \{(k, y_k) : k = 0, 1, \ldots, 9\}$. The data curve of the monthly mean total sunspot numbers and data points in $\mathcal{D}$ are shown in Figure 1.

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

Table 1. Values of parameters.

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
-0.6	0.2	0.2	-0.8	-0.5	0.1	0.1	-0.6	0.3

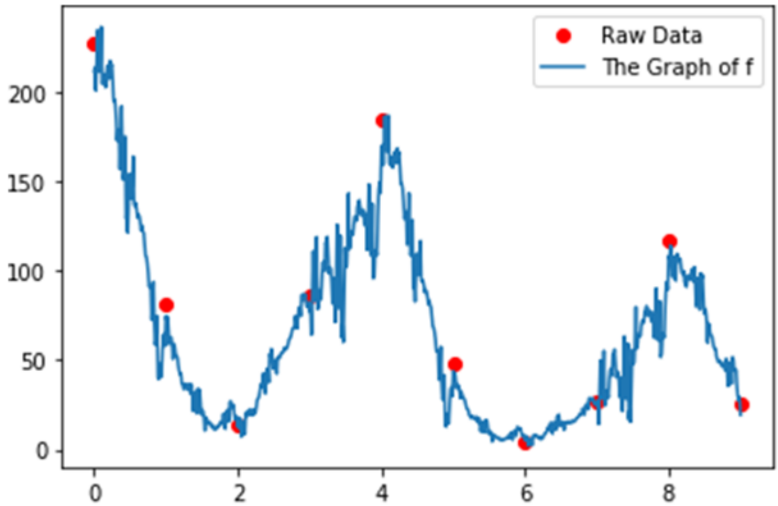

Figure 2. Fractal curve of $f_V + f_B(\lambda_1 = 0.02, \lambda_2 = 0.03)$.

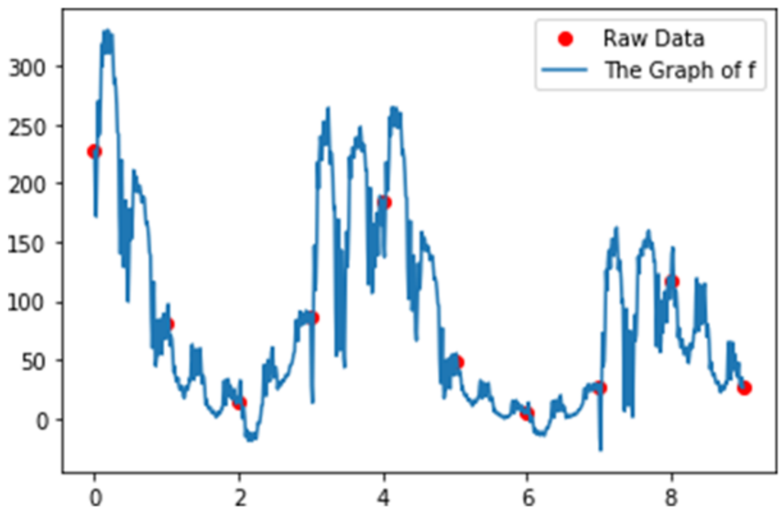

Figure 3. Fractal curve of $f_V + f_B(\lambda_1 = 0.02, \lambda_2 = 0)$.

Define $\langle f, g\rangle_{\mathcal{F}} = \int_I f(t)g(t)dt$ for $f, g \in \mathcal{F}$. For $k = 1, \ldots, 9$, let $J_k = I = [0, 9]$ and let L_k be the linear polynomial that satisfies the conditions $L_k(0) = k - 1$ and $L_k(9) = k$. For $j = 0, 1, \ldots, 9$, let u_j be the Gaussian function defined by

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

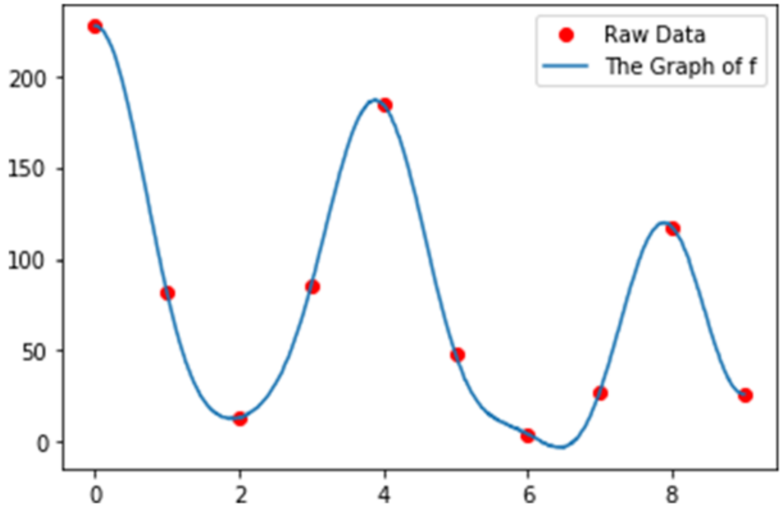

Figure 4. Fractal curve of $f_V + f_B(\lambda_1 = 0, \lambda_2 = 0.03)$.

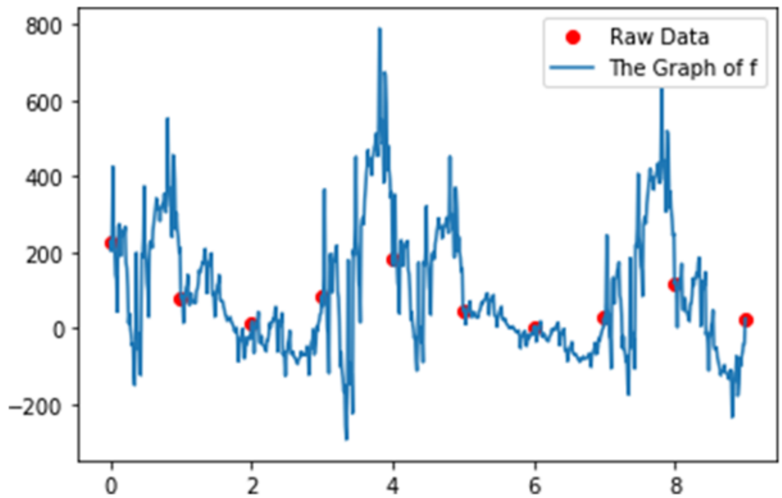

Figure 5. Fractal curve of $f_V + f_B(\lambda_1 = 0, \lambda_2 = 0.0)$.

$$u_j(t) = \frac{1}{h} \exp\left[-\frac{(t-t_j)^2}{h^2}\right], \qquad h > 0. \tag{5.1}$$

In this example, $t_j = j$, and we set $h = 0.7$. We construct $\phi_j = \Phi(u_j)$ by the approach given in § 2 with linear polynomials p_k and parameters s_k given in Table 1. We also define $\langle f, g\rangle_{\mathcal{C}} = \int_I f(t)g(t)dt$ for $f, g \in \mathcal{C}$, and let

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

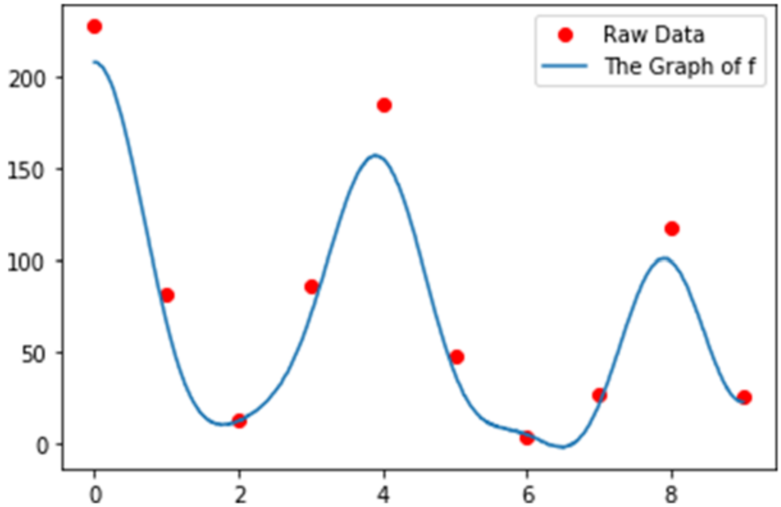

Figure 6. Fractal curve of $f_V(\lambda_1 = 0.02)$.

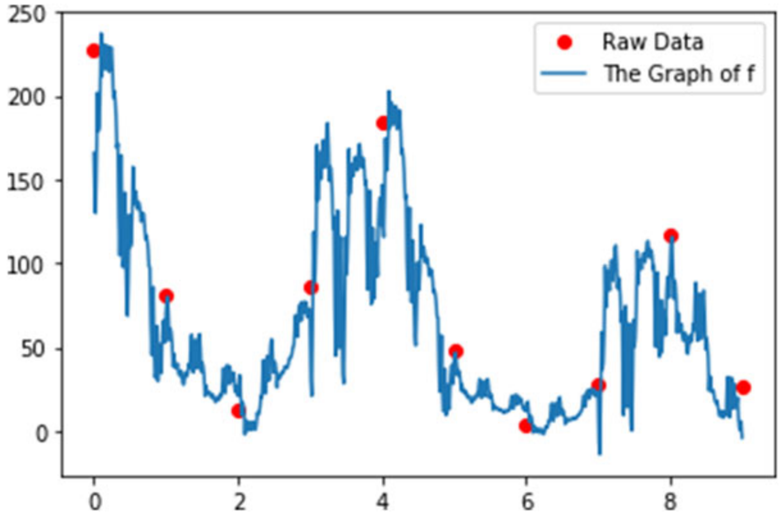

Figure 7. Fractal curve of $f_B(\lambda_2 = 0.03)$.

$$v_j(t) = \cos\left(\frac{\pi j t}{|I|}\right), \quad j = 0, 1, \ldots, \xi.$$

In this example, $|I| = 9$ and we choose $\xi = 9$.

Let $\mathbf{k}^*$ and $\mathbf{k}$ be defined by Equations (4.1) and (3.4), respectively. By choosing $\lambda_1 = 0.02$, $\lambda_2 = 0.03$, we compute the coefficients γ_k and α_k by Equation (4.8) and then establish

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

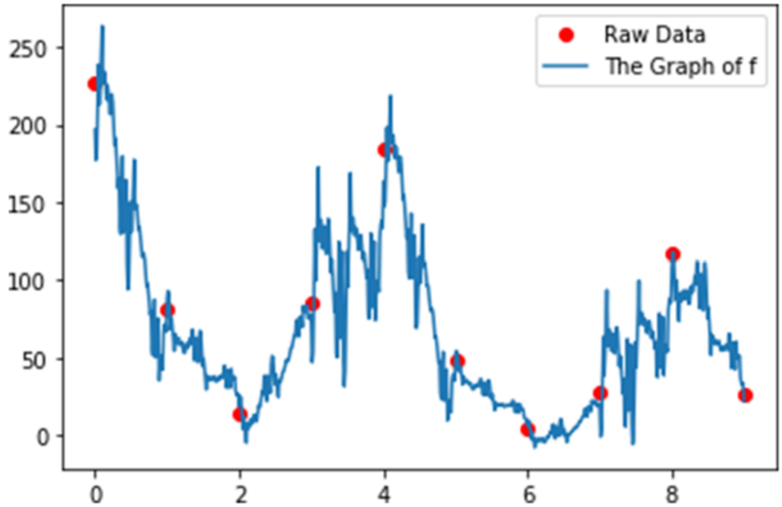

Figure 8. Fractal curve of $f_V + f_B(\lambda_1 = 0.02, \lambda_2 = 0.03)$.

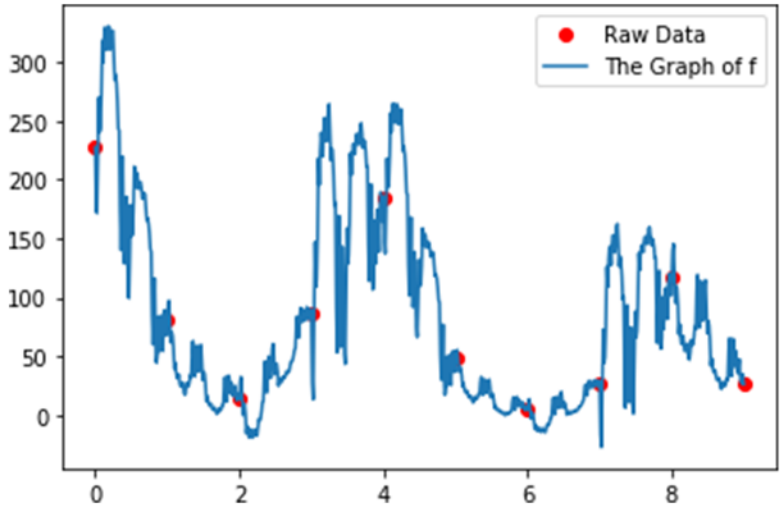

Figure 9. Fractal curve of $f_V + f_B(\lambda_1 = 0.02, \lambda_2 = 0.0)$.

$$f = f_{\mathcal{V}} + f_{\mathcal{B}} = \sum_{m=0}^{9} \gamma_m \mathbf{k}^*_{tm} + \sum_{j=0}^{9} \alpha_j \mathbf{k}_{t_j}, \tag{5.2}$$

where each $\mathbf{k}^*_{tm}$ is given by Equation (4.3) with $\xi = 9$ and each $\mathbf{k}_{t_j}$ is given by Equation (3.7) with $\eta = 9$. The graph of f is shown in Figure 2. The fractal curve with $\lambda_1 = 0.02$ and $\lambda_2 = 0$ is shown in Figure 3. The fractal curve with $\lambda_1 = 0$ and $\lambda_2 = 0.03$ is demonstrated in Figure 4. If we choose $\lambda_1 = \lambda_2 = 0$, we obtain the fractal interpolation curve which is shown in Figure 5.

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

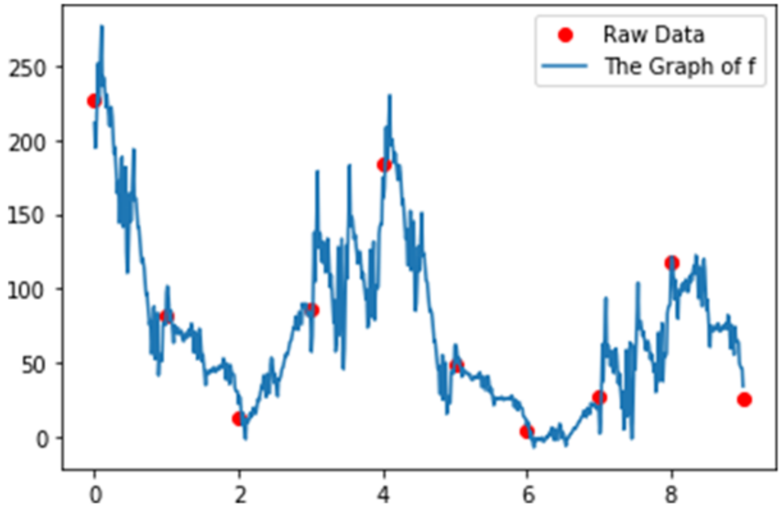

Figure 10. Fractal curve of $f_V + f_B(\lambda_1 = 0.0, \lambda_2 = 0.03)$.

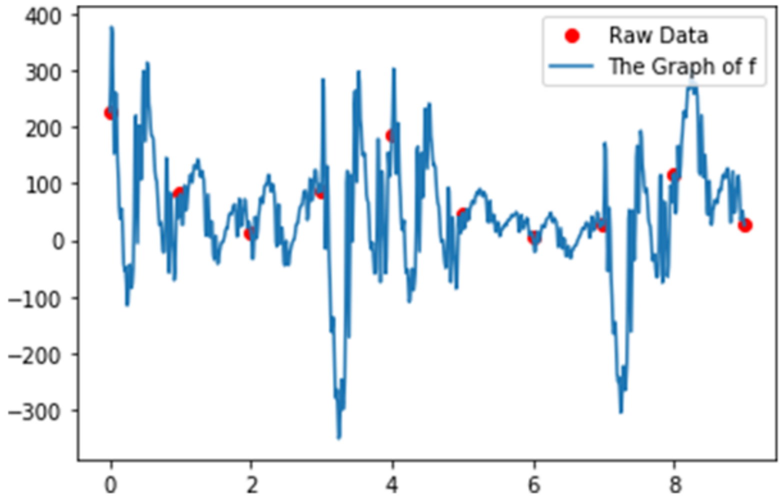

Figure 11. Fractal curve of $f_V + f_B(\lambda_1 = 0.0, \lambda_2 = 0.0)$.

A function in $\mathcal{C}_{\mathcal{V}}$ that minimizes

$$\frac{1}{N+1}\sum_{k=0}^{N}(y_k - (f_{\mathcal{V}}(t_k)))^2 + \lambda_1 \|f_{\mathcal{V}}\|_{\mathcal{C}}^2 \tag{5.3}$$

with $N = 9$ is given by $f_{\mathcal{V}} = \sum_{m=0}^{9} \gamma_m \mathbf{k}_{t_m}^*$, where $\mathbf{D} = [\gamma_0, \gamma_1, \ldots, \gamma_9]^T$ satisfies

$$\mathbb{K}^*(\mathbb{K}^* + \lambda_1(N+1)\mathbf{I})\mathbf{D} = \mathbb{K}^*\mathbf{Y}. \tag{5.4}$$

https://doi.org/10.1017/S0013091524000415 Published online by Cambridge University Press

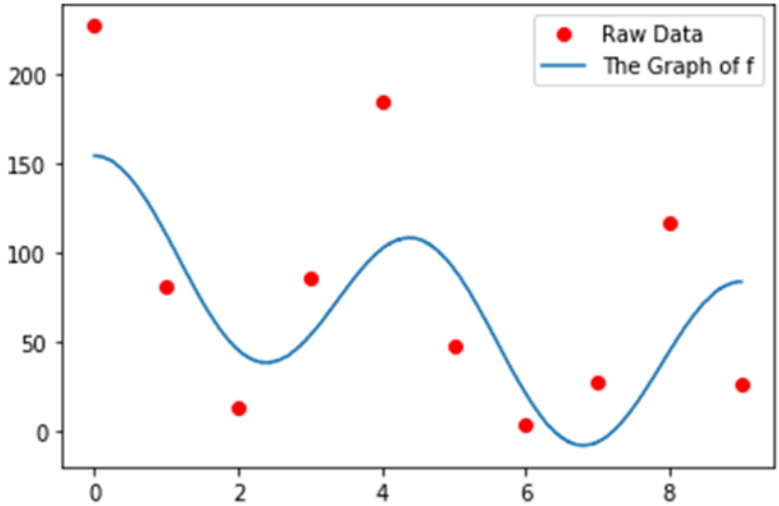

Figure 12. Fractal curve of $f_V(\lambda_1 = 0.02)$.

If we set $\lambda_1 = 0.02$, the graph of $f_{\mathcal{V}}$ is shown in Figure 6.

A function in $\mathcal{F}_{\mathcal{B}}$ that minimizes

$$\frac{1}{N+1}\sum_{k=0}^{N}(y_k - (f_{\mathcal{B}}(t_k)))^2 + \lambda_2\|f_{\mathcal{B}}\|_{\mathcal{F}}^2 \tag{5.5}$$

with $N = 9$ is given by $f_{\mathcal{B}} = \sum_{j=0}^{9}\alpha_j \mathbf{k}_{t_j}$, where $\mathbf{C} = [\alpha_0, \alpha_1, \ldots, \alpha_9]^T$ satisfies

$$\mathbb{K}(\mathbb{K} + \lambda_2(N+1)\mathbf{I})\mathbf{C} = \mathbb{K}\mathbf{Y}. \tag{5.6}$$

If we set $\lambda_2 = 0.03$, the graph of $f_{\mathcal{B}}$ is shown in Figure 7.

Similar graphs of functions for the case $\xi = 4$ are given in Figures 8–12.

Remark 5.1. In this paper, the parameters $\{s_k\}$ given in Table 1 are just an example of the graphs of fractal functions constructed by our approach. These parameters are not good choices for the cases shown in Figures 5 and 11. Determining parameters $\{s_k\}$ plays an essential role in the theory of fractal functions and applications on curve fitting problems. The study of finding the optimal values of $\{s_k\}$ for curve fitting problems is one of our future research directions.

Funding Statement. This work was supported by the Ministry of Science and Technology, R.O.C., under Grant MOST 110-2115-M-214-002.

References

(1) N. Aronszajn, Theory of reproducing kernels, *Trans. Amer. Math. Soc.* **68**(3) (1950), 337–404.

(2) M. F. Barnsley, Fractal functions and interpolation, *Constr. Approx.* **2** (1986), 303–329.

(3) M. F. Barnsley, *Fractals everywhere* (Academic Press, Orlando, FL, 1988).

(4) M. F. Barnsley, J. Elton, D. Hardin and P. Massopust, Hidden variable fractal interpolation functions, *SIAM J. Math. Anal.* **20**(5) (1989), 1218–1242.

(5) M. F. Barnsley and P. R. Massopust, Bilinear fractal interpolation and box dimension, *J. Approx. Theory* **192**(C) (2015), 362–378.

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

(7) P. Bouboulis and M. Mavroforakis, Reproducing kernel Hilbert spaces and fractal interpolation, *J. Comput. Appl. Math.* **235**(12) (2011), 3425–3434.

(8) A. K. B. Chand and G. P. Kapoor, Generalized cubic spline fractal interpolation functions, *SIAM J. Numer. Anal.* **44**(2) (2006), 655–676.

(9) A. K. B. Chand and M. A. Navascués, Generalized Hermite fractal interpolation, *Rev. R. Acad. Cienc. Zaragoza* **64**(2) (2009), 107–120.

(10) A. K. B. Chand and P. Viswanathan, A constructive approach to cubic Hermite fractal interpolation function and its constrained aspects, *BIT* **53**(4) (2013), 841–865.

(11) F. Cucker and D. X. Zhou, *Learning theory: an approximation theory viewpoint* (Cambridge University Press, New York, NY, 2007).

(12) L. Györfi, M. Kohler, A. Krzyzak and H. Walk, *A distribution-free theory of nonparametric regression* (Springer, New York, 2002).

(13) W. Härdle, M. Müller, S. Sperlich and A. Werwatz, *Nonparametric and semiparametric models* (Springer, New York, 2004).

(14) J. D. Hart, *Nonparametric smoothing and lack-of-fit tests* (Springer-Verlag, New York, 1997).

(15) S. K. Katiyar, A. K. B. Chand and G. S. Kumar, A new class of rational cubic spline fractal interpolation function and its constrained aspects, *Appl. Math. Comput.* **346** (2019), 319–335.

(16) Q. Li and J. S. Racine, *Nonparametric econometrics* (Princeton University Press, Princeton, NJ, 2007).

(17) D.-C. Luor, Fractal interpolation functions with partial self similarity, *J. Math. Anal. Appl.* **464**(1) (2018), 911–923.

(18) D.-C. Luor, Fractal interpolation functions for random data sets, *Chaos Solitons Fractals* **114** (2018), 256–263.

(19) M. A. Marvasti and W. C. Strahle, Fractal geometry analysis of turbulent data, *Signal Process.* **41**(2) (1995), 191–201.

(20) P. R. Massopust, *Fractal functions, fractal surfaces, and wavelets* (Academic Press, San Diego, 1994).

(21) P. R. Massopust, *Interpolation and approximation with splines and fractals* (Oxford University Press, New York, 2010).

(22) D. S. Mazel and M. H. Hayes, Using iterated function systems to model discrete sequences, *IEEE Trans. Signal Process.* **40**(7) (1992), 1724–1734.

(23) D. S. Mazel, Representation of discrete sequences with three-dimensional iterated function systems, *IEEE Trans. Signal Process.* **42**(11) (1994), 3269–3271.

(24) M. A. Navascués, A fractal approximation to periodicity, *Fractals* **14**(4) (2006), 315–325.

(25) M. A. Navascués, Fractal approximation, *Complex Anal. Oper. Theory* **4** (2010), 953–974.

(26) M. A. Navascués, Fractal bases of L_p spaces, *Fractals* **20**(2) (2012), 141–148.

(27) M. A. Navascués and M. V. Sebastián, Generalization of Hermite functions by fractal interpolation, *J. Approx. Theory* **131**(1) (2004), 19–29.

(28) V.I. Paulsen and M. Raghupathi, *An introduction to the theory of reproducing kernel Hilbert spaces* (Cambridge University Press, Cambridge, UK, 2016).

(29) C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning* (The MIT Press, Cambridge, MA, 2006).

(30) A. B. Tsybakov, *Introduction to nonparametric estimation* (Springer, New York, 2009).

(31) V. N. Vapnik, *Statistical learning theory* (Wileys, New York, 1998).

(32) P. Viswanathan and A. K. B. Chand, Fractal rational functions and their approximation properties, *J. Approx. Theory* **185** (2014), 31–50.

(33) P. Viswanathan and A. K. B. Chand, α-fractal rational splines for constrained interpolation, *Electron. Trans. Numer. Anal.* **41** (2014), 420–442.

(34) P. Viswanathan, M. A. Navascués and A. K. B. Chand, Associate fractal functions in $\mathcal{L}^p$-spaces and in one-sided uniform approximation, *J. Math. Anal. Appl.* **433**(2) (2016), 862–876.

(35) H.-Y. Wang and J.-S. Yu, Fractal interpolation functions with variable parameters and their analytical properties, *J. Approx. Theory* **175** (2013), 1–18.