Title
Stochastic dynamic models and Chebyshev splines

Journal
Canadian Journal of Statistics, 42(4)

ISSN
0319-5724

Authors
Fan, Ruzong
Zhu, Bin
Wang, Yuedong

Publication Date
2014-12-01

DOI
10.1002/cjs.11233

Peer reviewed

HHS Public Access
Author manuscript
Can J Stat. Author manuscript; available in PMC 2015 June 02.

Published in final edited form as:
Can J Stat. 2014 December 1; 42(4): 610–634. doi:10.1002/cjs.11233.

Stochastic dynamic models and Chebyshev splines

Ruzong Fan[1,*], **Bin Zhu**[2], and **Yuedong Wang**[3]

[1]Biostatistics and Bioinformatics Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Rockville, MD 20852, U.S.A

[2]Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD 20850, U.S.A

[3]Department of Statistics & Applied Probability, University of California—Santa Barbara, Santa Barbara, CA 93106, U.S.A

Abstract

In this article, we establish a connection between a stochastic dynamic model (SDM) driven by a linear stochastic differential equation (SDE) and a Chebyshev spline, which enables researchers to borrow strength across fields both theoretically and numerically. We construct a differential operator for the penalty function and develop a reproducing kernel Hilbert space (RKHS) induced by the SDM and the Chebyshev spline. The general form of the linear SDE allows us to extend the well-known connection between an integrated Brownian motion and a polynomial spline to a connection between more complex diffusion processes and Chebyshev splines. One interesting special case is connection between an integrated Ornstein–Uhlenbeck process and an exponential spline. We use two real data sets to illustrate the integrated Ornstein–Uhlenbeck process model and exponential spline model and show their estimates are almost identical.

Keywords

Brownian motion; Ornstein–Uhlenbeck process; reproducing kernel Hilbert space; smoothing splines; stochastic differential equations

1. INTRODUCTION

There exists a one-to-one correspondence between a reproducing kernel Hilbert space (RKHS) and Gaussian stochastic processes by the Kolmogorov consistency theorem (Cramér & Leadbetter, 1967; Wahba, 1990). This correspondence has led to the development of connections between general smoothing spline models and Gaussian stochastic processes which enables researchers to borrow strength across different fields. For example, stochastic models corresponding to smoothing splines have been used to derive the generalized maximum likelihood estimate of smoothing parameters and to construct

*Author to whom correspondence may be addressed. fanr@mail.nih.gov.

Bayesian confidence intervals (Wahba, 1990). A specific connection between an L-spline and an integrated Brownian motion has been explored to develop efficient O(n) algorithms (Wecker & Ansley, 1983).

Specific connections between more complex processes and related smoothing splines have not been studied extensively. The purpose of this article is to establish a connection between a stochastic dynamic model (SDM) driven by a linear stochastic differential equation (SDE) and a Chebyshev spline. We construct a differential operator for the penalty function and develop an RKHS corresponding to the SDM and the Chebyshev spline. The general form of the SDE allows us to establish a connection between more complex diffusion processes and Chebyshev splines. As an interesting special case, an integrated Ornstein–Uhlenbeck process will be connected to an exponential spline.

The extended connection between the SDM driven by a linear SDE and a Chebyshev spline can be explored to motivate new SDMs based on spline models, and vice versa. As an illustration, we will present a partial spline model motivated by an SDM and an SDM motivated by logistic spline. We differ from the smoothing spline literature where one usually builds SDMs that connect to spline models of interest; here we will build spline models that connect to SDMs. We will present the construction of penalties corresponding to SDMs, which may be regarded as priors for regression functions in non-parametric regression models.

Estimation and computational methods for Chebyshev splines may be used to compute the posterior means of the corresponding SDMs, and vice versa. We will use two real data sets to show that the posterior mean of the integrated Ornstein–Uhlenbeck process in an SDM matches the penalized least squares estimate of the corresponding exponential spline.

Gaussian stochastic models have been widely used in many fields including physics, engineering, finance and biology (Ansley & Kohn, 1986; Rue & Martino, 2009; Stuart, 2010; Griebel & Hegland, 2010; Lindgren, Rue, & Lindstrom, 2011; Papaspiliopoulis et al., 2012). Some of the interesting special cases are given in Bishwal (2008). On the other hand, much work has been done in the area of smoothing splines. For instance, Pintore, Speckman, & Holmes (2006) used an RKHS representation to derive the smoothing spline with a particular inhomogeneous differential operator penalty; Furrer & Nychka (2007) built a framework to understand the asymptotic properties of Kriging and splines; in this framework Kriging estimators are interpreted as generalized smoothing splines. The differential operator $L = D^q$ in Pintore, Speckman, & Holmes (2006) is a special case of that considered in this paper.

The remainder of this article is organized as follows. In Section 2, we introduce the SDM. In Section 3, we compute mean and covariance functions for a general class of Gaussian models driven by a linear SDE and construct corresponding Chebyshev splines. We extend the SDM in Section 4 with general differential operators and connect them to general Chebyshev splines. We use two real data sets to confirm the theoretical results in Section 5 and conclude with some remarks in Section 6.

2. STOCHASTIC DYNAMIC MODELS

We present an SDM driven by an SDE in the Section 2.1. Then we develop an equivalent stochastic integration equation and use this equivalent model to compute the mean and covariance functions of the dynamic system in Section 2.2.

2.1. Stochastic Dynamic Models

Consider a temporal SDM

$$Y(t,\omega)=U(t,\omega)+\varepsilon(t,\omega), \omega\in\Omega, t\in[0,T], \quad (1)$$

where $Y(t, \omega)$ is the observation at time t, $U(t, \omega)$ is a latent stochastic process of interest observed at time t on the path ω, $\varepsilon(t, \omega)$ is an error term, Ω is the sample path probability space, and T is a positive real number which may be the right end point of time in consideration. We assume that $U(\cdot)$ and $\varepsilon(\cdot)$ are independent and $\varepsilon(\cdot)$ is a Gaussian white noise process with variance σ^2.

In this paper, we will first consider the following stochastic dynamic system for the latent stochastic process $U(\cdot)$

$$\frac{d^q U(t)}{\mathrm{d}t^q}=a_0(t)U_q+b(t)V(t), q\geq 0, t\in[0,T], \quad (2)$$

where $U_k = d^kU(0)/\mathrm{d}t^k$ for $k = 0, \ldots, q$ are initial values of $U(\cdot)$ and its derivatives up to the order q, and $a_0(t)$ and $b(t)$ are integrable deterministic functions. We assume that $U_0, \ldots, U_q$ are mutually independent and square integrable random variables. A more general equation for $U(\cdot)$ that corresponds to the general Chebyshev splines will be considered in Section 4.

We assume that the stochastic process $V(\cdot)$ in (2) is independent of $U_0, \ldots, U_q$ and is driven by the following SDE

$$\mathrm{d}V(t,\omega)=\beta(t,V(t,\omega))\mathrm{d}t+\sigma(t,V(t,\omega))\mathrm{d}B(t,\omega), V(0,\omega)=0, t\in[0,T], \quad (3)$$

where $\beta(t, V(t, \omega))$ is a drift term, $\sigma^2(t, V(t, \omega))$ is a diffusion coefficient, and $B(t, \omega)$ is a standard Brownian motion observed at time t on the path ω. We assume that the SDE (3) has a unique, continuous and adapted strong solution, and $V(\cdot)$ has finite moments of any order $p \in [1,\infty)$. This assumption holds when the coefficients in (3) satisfy the Lipschitz condition, $|\sigma(t, x) - \sigma(t, y)|^2 + |\beta(t, x) - \beta(t, y)|^2 \leq c_1(T)|x - y|^2$, and the growth condition, $|\sigma(t, x)|^2 + |\beta(t, x)|^2 \leq c_2(T)(1 + x^2)$ for all $x, y \in R$ and $t \in [0, T]$, where $c_1(T)$ and $c_2(T)$ are positive constants depending on T only (Ikeda & Watanabe, 1989).

Together, Equations (1), (2) and (3) define the SDM we study in this paper. The SDE (3) is a parametric Itô SDE with initial value equal to zero. The assumption about the initial value may be removed (see Remark 3.1). Zhu, Song, & Taylor (2011) considered a similar SDM with $a_0(t) = 0$ and $b(t) = 1$ in the stochastic dynamic system (2) without the initial condition $V(0) = 0$, and used an Ornstein–Uhlenbeck process $V(\cdot)$ to model the rate of changes of prostate specific antigen profiles. We consider the more general Equation (2) for the connection to Chebyshev splines.

It is noteworthy that a special case of the SDM has been connected to polynomial smoothing splines. Specifically, let $a_0(t) = b(t) = 1$, $\beta(t, V(t)) = 0$ and $\sigma(t, V(t)) = \sigma_V$, then $U(\cdot)$ is the same as the random effects model (1.5.8) in Wahba (1990) for a polynomial spline of order $q + 1$ (refer to Section S.5 in the Supplementary Materials for details). When $q = 1$ and $V(\cdot)$ is a standard Brownian motion, then $U(\cdot)$ is an integrated Brownian motion. Alternatively, one may want to use an Ornstein–Uhlenbeck process $V(\cdot)$ to model exponentially decreasing correlation. In this case, $U(\cdot)$ is an integrated Ornstein–Uhlenbeck process.

2.2. An Equivalent Stochastic Integration Equation

In this section, we first introduce a stochastic integration equation that is equivalent to the stochastic dynamic system (2). We then use the stochastic integration equation to compute the mean and covariance functions of the stochastic process $U(\cdot)$.

When $q = 0$, the stochastic dynamic system (2) has the form

$$U(t)=a_0(t)U_0+b(t)V(t), \quad (4)$$

and it is straightforward to compute the mean and covariance functions of $U(\cdot)$ in this case.

In the following discussion, we assume that $q \geq 1$.We will show that the following stochastic integration equation is equivalent to the stochastic dynamic system (2)

$$U(t)=\sum_{i=0}^{q-1}\psi_i(t)U_i+A_q(t)U_q+\int_0^t\psi_q(t-s)\mathrm{d}[b(s)V(s)], \quad (5)$$

where $\psi_i(t) \triangleq t^i/i!$, $A_q(t) \triangleq \int_0^t\psi_{q-1}(t-s)a_0(s)\mathrm{d}s$, and $V(\cdot)$ is a stochastic process driven by the SDE defined in (3). We first show that (5) can be rewritten as

$$U(t)=\sum_{i=0}^{q-1}\psi_i(t)U_i+A_q(t)U_q+\int_0^t\psi_{q-1}(t-s)b(s)V(s)\mathrm{d}s. \quad (6)$$

To see that (5) can be rewritten as (6), applying Itô formula to $b(s)V(s)(t - s)^q$, $s \leq t$, leads to

$$\int_0^t(t-s)^q\mathrm{d}[b(s)V(s)]+\int_0^t b(s)V(s)\mathrm{d}(t-s)^q=b(s)V(s)(t-s)^q|_0^t=0.$$

Hence, we have

$$\int_0^t(t-s)^q\mathrm{d}[b(s)V(s)]=q\int_0^t b(s)V(s)(t-s)^{q-1}\mathrm{d}s. \quad (7)$$

Suppose that $U(\cdot)$ is given by the stochastic integration Equation (5). By Equation (7), we have the expression (6).

When $q = 1$, it is easy to see that taking the derivative on both sides of (6) with respect to t leads to (2). When $q \geq 2$, taking the derivative on both sides of (6) with respect to t leads to

$$\begin{aligned}\frac{\mathrm{d}U(t)}{\mathrm{d}t} &= \sum_{i=1}^{q-1}\psi_{i-1}(t)U_i + A_{q-1}(t)U_q + \int_0^t \psi_{q-2}(t-s)b(s)V(s)\mathrm{d}s + [\psi_{q-1}(t-s)b(s)V(s)]\Big|_{s=t} \\ &= \sum_{i=1}^{q-1}\psi_{i-1}(t)U_i + A_{q-1}(t)U_q + \int_0^t \psi_{q-2}(t-s)b(s)V(s)\mathrm{d}s.\end{aligned}$$

Continuing this process to calculate higher order derivatives, we get $d^qU(t)/\mathrm{d}t^q = a_0(t)U_q + b(t)V(t)$. Therefore, the stochastic integration Equation (5) or (6) implies the stochastic differential system (2).

Conversely, assuming that $U(\cdot)$ is given by the stochastic differential system (2) and reversing the process above, it is not difficult to show that $U(\cdot)$ satisfies the stochastic integration Equation (5) or (6). The following lemma summarizes the above discussion.

Lemma 2.1—*Suppose that the stochastic process $V(\cdot)$ is driven by the SDE (3). Then, the stochastic dynamic system (2) is equivalent to the stochastic integration Equation (5) or (6).*

Note that when $a_0(t) = b(t) = 1$ and $V(t) = \sigma_V B(t)$, we have $A_q(t) = \psi_q(t)$ and $U(t) = \sum_{i=0}^{q}\psi_i(t)U_i + \sigma_V\int_0^t\psi_q(t-s)\mathrm{d}B(s)$ which is the stochastic process corresponding to the polynomial spline of order $q + 1$ (Kimeldorf & Wahba, 1970a,b; Wahba, 1990).

The mean and covariance functions of $U(\cdot)$ listed in the following proposition can be computed directly based on (4) when $q = 0$ or by applying the Fubini's theorem to Equation (6) when $q \geq 1$.

Proposition 2.1—*For the stochastic process $U(\cdot)$ defined in (2) or equivalently in (5) or (6), when $q \geq 1$, we have*

$$\mathrm{E}U(t) = \sum_{i=0}^{q-1}\psi_i(t)\mu_i + A_q(t)\mu_q + \int_0^t \psi_{q-1}(t-s)b(s)\mathrm{E}V(s)\mathrm{d}s,$$

$$\mathrm{Cov}(U(s), U(t)) = \sum_{i=0}^{q-1}\psi_i(s)\psi_i(t)\sigma_i^2 + A_q(s)A_q(t)\sigma_q^2 + \int_0^t\int_0^s \psi_{q-1}(t-u)\psi_{q-1}(s-v)b(u)b(v)\mathrm{Cov}(V(u), V(v))\mathrm{d}u\mathrm{d}v,$$

where $\mu_i \triangleq \mathrm{E}\,U_i$ and $\sigma_i^2 \triangleq \mathrm{Var}(U_i)$ When $q = 0$, we have $\mathrm{E}\,U(t) = a_0(t)\mu_0 + b(t)\,\mathrm{E}\,V(t)$ and $\mathrm{Cov}(U(s), U(t)) = a_0(s)a_0(t)\sigma_0^2 + b(s)b(t)\mathrm{Cov}(V(s), V(t))$.

There exists a unique RKHS that is congruent to the Hilbert space generated by the stochastic process *U*(·) (Berlinet & Thomas-Agnan, 2004). In the next section, we will make specific connections between Gaussian processes driven by a linear SDE and an RKHS.

3. GAUSSIAN PROCESS DRIVEN BY A LINEAR SDE AND CONNECTION TO CHEBYSHEV SPLINES

The stochastic process *V*(·) driven by the SDE (3) is a Markovian process. However, it is not necessarily a Gaussian process. In this section, we consider a Gaussian process driven by the following linear SDE

$$\mathrm{d}V(t)=[\beta_0(t)+\beta_1(t)V(t)]\mathrm{d}t+\sigma(t)\mathrm{d}B(t), V(0)=0, t\in[0,T], \quad (8)$$

where $\beta_0(t)$, $\beta_1(t)$ and $\sigma(t)$ are deterministic, measurable, and bounded functions of time *t* (Karatzas & Shreve, 1988). The SDE (8) is a special case of (3) when $\beta(t, V(t, \omega))$ is linear in *V*(·) and $\sigma(t, V(t, \omega))$ is deterministic.

In this section, we establish a connection between the SDM where *V*(·) is driven by the linear SDE (8) and Chebyshev splines. In Section 3.1, we apply results in Section 2.2 to compute the mean and covariance functions for Gaussian processes driven by the linear SDE. In Section 3.2, we build connections between Gaussian processes and Chebyshev splines. In the subsections, we provide examples of some specific SDMs and their corresponding spline models.

3.1. Gaussian Process Driven by the Linear SDE (8)

Denote $\Psi(t) \triangleq \exp\left(\int_0^t \beta_1(s)\mathrm{d}s\right)$. By the Itô formula, one can show that (Karatzas & Shreve, 1988, pp. 354–355)

$$V(t)=\Psi(t)\left[\int_0^t \frac{\beta_0(s)}{\Psi(s)}\mathrm{d}s+\int_0^t \frac{\sigma(s)}{\Psi(s)}\mathrm{d}B(s)\right]. \quad (9)$$

Lemma 3.1—*Suppose that the stochastic process V*(·) *is a Gaussian process driven by the linear SDE (8). Then, we have*

$$\mathrm{E}V(t)=\Psi(t)\int_0^t \frac{\beta_0(s)}{\Psi(s)}\mathrm{d}s,$$

$$\mathrm{Cov}(V(s),V(t))=\Psi(s)\Psi(t)\int_0^{s\wedge t}\left(\frac{\sigma(u)}{\Psi(u)}\right)^2\mathrm{d}u.$$

The proof of Lemma 3.1 can be found in the Supplementary Materials. For $q \geq 1$, denote

$$F(t,u) \triangleq I(u \le t)\int_u^t \psi_{q-1}(t-\tau)b(\tau)\Psi(\tau)\mathrm{d}\tau,$$

where $I(u \le t) = 1$ when $u \le t$ and 0 otherwise. Combining results in Proposition 2.1 and Lemma 3.1, we have the following theorem.

Theorem3.1—*Suppose that the stochastic process* $V(\cdot)$ *is a Gaussian process driven by the linear SDE (8) and* $U_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ *for* $i = 0, \ldots, q$*. Then, the stochastic process* $U(\cdot)$ *is a Gaussian process with the following mean and covariance functions*

$$\mathrm{E}U(t) = \sum_{i=0}^{q-1}\psi_i(t)\mu_i + A_q(t)\mu_q + \int_0^t F(t,s)\frac{\beta_0(s)}{\Psi(s)}\mathrm{d}s,$$

$$\mathrm{Cov}(U(s),U(t)) = \sum_{i=0}^{q-1}\psi_i(s)\psi_i(t)\sigma_i^2 + A_q(s)A_q(t)\sigma_q^2 + \int_0^{s\wedge t}F(s,u)F(t,u)\left(\frac{\sigma(u)}{\Psi(u)}\right)^2\mathrm{d}u,$$

where, for $q = 0$*, we take* $\sum_{i=0}^{q-1} \triangleq 0$*,* $A_0(t) \triangleq a_0(t)$*, and* $F(t,u) \triangleq b(t)\Psi(t)I(u \le t)$*.*

The proof of Theorem 3.1 is provided in the Supplementary Materials.

Remark 3.1—*We have assumed that* $V(0) = 0$ *in the SDE (8). This assumption can be removed from the construction as follows. Let*

$$\overline{V}(t) = \Psi(t)\left[U_q + \int_0^t \frac{\beta_0(s)}{\Psi(s)}\mathrm{d}s + \int_0^t \frac{\sigma(s)}{\Psi(s)}\mathrm{d}B(s)\right].$$

Then, $\overline{V}(\cdot)$ *satisfies*

$$\mathrm{d}\overline{V}(t) = [\beta_0(t) + \beta_1(t)\overline{V}(t)]\mathrm{d}t + \sigma(t)\mathrm{d}B(t), \overline{V}(0) = U_q, t \in [0,T]. \quad (10)$$

When $a_0(t) = \Psi(t)$*,* $b(t) = 1$*, and* $V(\cdot)$ *is a solution of linear SDE (8), the stochastic differential system (2) is equivalent to*

$$\frac{d^q U(t)}{\mathrm{d}t^q} = \overline{V}(t). \quad (11)$$

This means that $d^q U(t)/\mathrm{d}t^q$ *is equal to the solution* $\overline{V}(t)$ *of the linear SDE (10) which starts from* U_q*. Both* $\overline{V}(\cdot)$ *and* $U(\cdot)$ *defined in (11) are Gaussian processes.*

3.2. Construction of a Chebyshev Spline Model

We now construct a Chebyshev spline model such that, up to a known function, the penalized least squares estimate equals the best linear unbiased estimator of the SDM defined by (1), (2) and (8).We note that the following development is in the opposite direction of the common approach employed in the spline literature where one constructs a stochastic process to connect with a smoothing spline model.

Define a differential operator

$$Lf(t) \triangleq \frac{d}{\mathrm{d}t}\left(\frac{f^{(q)}(t)}{\Psi(t)}\right) = \frac{f^{(q+1)}(t) - \beta_1(t) f^{(q)}(t)}{\Psi(t)}. \quad (12)$$

We may write L as $Lf(u) = D[D^q f(u)/\Psi(u)]$ where $D^q = d^q/\mathrm{d}t^q$ is the qth derivative operator. The differential operator L is a special case of the general differential operator L_{q+1} defined in Equation (4.64) of Gu (2013) for the Chebyshev spline with $w_i(t) = 1$ for $i = 1, \ldots, q$ and $w_{q+1}(t) = \Psi(t)$. Then, from Equation (4.65) in Gu (2013), the Chebyshev system on $[0, T]$ is $\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t), \tilde{\psi}_q(t)$ and they span the null space $\mathscr{H}_0 = \mathrm{span}\{\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t), \tilde{\psi}_q(t)\}$ of the differential operator L, where

$$\tilde{\psi}_q(t) = \int_0^t \psi_{q-1}(t-s)\Psi(s)\mathrm{d}s.$$

Under an inner product $\sum_{i=0}^{q} f^{(i)}(0) g^{(i)}(0)$, $\{\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t), \tilde{\psi}_q(t)\}$ forms an orthonormal basis of the null space $\mathscr{H}_0$. Furthermore, it can be shown that the Green's function associated with the differential operator L is

$$G(t,u) = I(u \le t)\int_u^t \psi_{q-1}(t-\tau)\Psi(\tau)\mathrm{d}\tau,$$

which is given by the relation (4.67) in Gu (2013).

Consider the model space

$$W_2^{q+1}[0,T] \triangleq \left\{ f : f, f', \ldots, f^{(q)} \text{are absolutely continuous and} \int_0^T (Lf(t))^2 h(t)\mathrm{d}t < \infty \right\} \quad (13)$$

with an inner product

$$(f,g) \triangleq \sum_{i=0}^{q} f^{(i)}(0) g^{(i)}(0) + \int_0^T Lf(u) Lg(u) h(u)\mathrm{d}u, \quad (14)$$

where $h(u) \triangleq (\Psi(u)/\sigma(u))^2$ is a weight function. Following the same arguments as in Section 4.5.2 of Gu (2013), we have the following result.

Theorem 3.2—*Assume that* $a_0(t) = \Psi(t)$ *and* $b(t) = 1$. *Then,* $W_2^{q+1}[0, T]$ *is an RKHS. Let us denote the kernel of L as* $\mathscr{H}_0 = \left\{f \in W_2^{q+1}[0, T]: Lf = 0\right\}$. *Then,* $\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t)$, $A_q(t) = \tilde{\psi}_q(t)$ *form an orthonormal basis of* $\mathscr{H}_0$, and $W_2^{q+1}[0, T]$ *can be decomposed into* $W_2^{q+1}[0, T] = \mathscr{H}_0 \oplus \mathscr{H}_1$, *where*

$$\mathscr{H}_0 = span\left\{\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t), \tilde{\psi}_q(t)\right\},$$

$$\mathscr{H}_1 = \left\{f \in W_2^{q+1}[0, T]: f(0) = f'(0) = \cdots = f^{(q)}(0) = 0\right\}.$$

The reproducing kernels of $\mathscr{H}_0$ *and* $\mathscr{H}_1$ *are respectively given by*

$$R_0(s, t) = \sum_{i=0}^{q-1} \psi_i(s)\psi_i(t) + \tilde{\psi}_q(s)\tilde{\psi}_q(t),$$

$$R_1(s, t) = \int_0^{s \wedge t} G(s, u)G(t, u)[h(u)]^{-1} \mathrm{d}u.$$

The proof of Theorem 3.2 is straightforward based on the fact that $A_q(t) = \tilde{\psi}_q(t)$ and $G(t, u) = F(t, u)$ when $a_0(t) = \Psi(t)$ and $b(t) = 1$.

Remark 3.2—*In Theorem 3.2 we assumed that* $a_0(t) = \Psi(t)$. *In general, assume that* $a_0(t)$ *is strictly positive and* $a_0(0) = 1$. *It is not difficult to check that* $D^i\psi_j(0) = 1$ *if* $i = j$ *and* 0 *otherwise for* $0 \le i, j \le q - 1$. *Furthermore,* $D^q\psi_j(0) = 0$ *for* $0 \le j \le q - 1$, $D^iA_q(0) = 0$ *for* $0 \le i \le q - 1$, *and* $D^qA_q(0) = a_0(0) = 1$. *Then* $\{\psi_0(t), \psi_1(t), \ldots, \psi_{q-1}(t), A_q(t)\}$ *forms an orthonormal basis of a subspace of* $W_2^{q+1}[0, T]$ *under the inner product* $\sum_{i=0}^{q} f^{(i)}(0)g^{(i)}(0)$. *In addition, we have* $L\psi_i(t) = 0$ *for* $i = 0, 1, \ldots, q-1$. *It is easy to check that* $D^qA^q(t) = a_0(t)$. *Therefore,* $LA_q(t) = 0$ *iff* $a_0(t) = \Psi(t)$. *Consequently,* $A_q(t)$ *does not belong to the space* $\mathscr{H}_0 = \left\{f \in W_2^{q+1}[0, T]: Lf = 0\right\}$ *when* $a_0(t) \ne \Psi(t)$. *Nevertheless, when* $a_0(t) \ne \Psi(t)$, *a partial spline model may be constructed. See Section 3.3.2 for an example.*

Consider the following nonparametric regression model

$$Y_i = f(t_i) + \varepsilon_i, i = 1, \ldots, n, t_i \in [0, T]. \quad (15)$$

Assume that $f \in W_2^{q+1}[0, T]$. A Chebyshev spline is the solution to the following penalized least Squares

$$\min_{f \in W_2^{q+1}[0,T]} \left\{ \frac{1}{n} \sum_{i=1}^{n} [Y_i - f(t_i)]^2 + \lambda \int_0^T (Lf(t))^2 h(t) \mathrm{d}t \right\}, \quad (16)$$

where λ is a smoothing parameter and L is given in (12). Let $\boldsymbol{y} = (Y_1, \ldots, Y_n)'$, $\Sigma = \{R_1(t_i, t_j)\}_{i,j=1}^{n}$, $S = \{(\psi_0(t_i), \ldots, \psi_{q-1}(t_i), A_q(t_i))\}_{i=1}^{n}$, and $M = \Sigma + n\lambda I_n$, where I_n is an $n \times n$ identity matrix. The solution to (16) can be represented as (Wang, 2011)

$$\hat{f}_\lambda(t) = \sum_{i=0}^{q-1} d_i \psi_i(t) + d_q A_q(t) + \sum_{k=1}^{n} c_i R_1(t, t_i), \quad (17)$$

Where

$$(d_0, \ldots, d_q)' = (S' M^{-1} S)^{-1} S' M^{-1} \mathbf{y},$$

$$(c_1, \ldots, c_n)' = M^{-1} [I_n - S(S' M^{-1} S)^{-1} S' M^{-1}] \mathbf{y}.$$

Now consider n observations based on the SDM (1)

$$Y(t_i) = U(t_i) + \varepsilon(t_i), i = 1, \ldots, n, t_i \in [0, T], \quad (18)$$

where the stochastic processes $U(\cdot)$ and $V(\cdot)$ are defined in (2) and (8), respectively. The best linear unbiased estimator of $U(t)$ is the posterior mean E $(U(t)|Y(t_i), i = 1, \ldots, n)$ (Wahba, 1990). Denote

$$\mu(t) \triangleq \int_0^t G(t,s) \beta_0(s) / \Psi(s) \mathrm{d}s. \quad (19)$$

In this article, we assume that $\mu(t)$ is known. Subtracting $\mu(t_i)$ on both sides of (18) and following the same arguments as in Wahba (1990) and Gu (2013), we have the following connection between the Chebyshev spline and the best linear unbiased estimator of $U(t)$.

Proposition 3.1—*Assume that* $a_0(t) = \Psi(t)$, $b(t) = 1$, *and* $U_0, \ldots, U_q \overset{iid}{\sim} \mathrm{N}(0, a)$. *Denote*

$$\hat{U}_a(t) = \mathrm{E}(U(t) | Y(t_i), i = 1, \ldots, n)$$

as the posterior mean where the dependence on the variance a is expressed explicitly. For any fixed $t \in [0, T]$, *when* $\lambda = \sigma^2/n$, *we have*

$$\lim_{a \to \infty} \hat{U}_a(t) = \mu(t) + \hat{f}_\lambda(t), \quad (20)$$

where $\hat{f}_\lambda(t)$ *in (20) is the penalized least squares solution to (16) with observations* $\boldsymbol{y} = (Y(t_1) - \mu(t_1), \ldots, Y(t_n) - \mu(t_n))'$.

Remark 3.3—*The penalty can be simplified as follows*

$$\int_0^T (Lf(t))^2 h(t)\mathrm{d}t = \int_0^T \left[\frac{f^{(q+1)}(t) - \beta_1(t) f^{(q)}(t)}{\sigma(t)}\right]^2 \mathrm{d}t. \quad (21)$$

It is clear that the construction of RKHS including inner product, basis function of the space $\mathcal{H}_0$ *and reproducing kernel of the space* $\mathcal{H}_1$ *is independent of the function* $\beta_0(t)$*, while whether there exists a drift term* $\mu(t)$ *defined in (19) depends on if* $\beta_0(t) = 0$*. The condition* $n\lambda = \sigma^2$ *does not depend on* $\sigma(t)$ *since* $\sigma(t)$ *is involved in the penalty. When* $\sigma(t)$ *is a constant, say* σ_V*, it may be absorbed into the smoothing parameter and then we have the standard condition* $n\lambda = \sigma^2/\sigma_V^2$.

Remark 3.4—*Assume that* $\beta_0(t) = 0$*. Then* $\mu(t) = 0$*. Proposition 3.1 states that the best linear unbiased estimator of* $U(t)$ *coincides with the smoothing spline estimate as* $a \to \infty$*. This link has been explored to derive the generalized maximum likelihood (restricted maximum likelihood) estimate of the smoothing parameter* λ*. The variance function of* $\hat{U}_a(t)$ *as* $a \to \infty$ *has been used to construct Bayesian confidence intervals. See Wang (2011) for details*.

Remark 3.5—*Proposition 3.1 extends existing results to the case when* $\mu(t) \neq 0$*. Denote*

$$\hat{\mathbf{U}} = lim_{a\to\infty}(U_a(t_1), \ldots, U_a(t_n))'$$

as the best linear unbiased estimates at design points, $\mathbf{Y} = (Y(t_1), \ldots, Y(t_n))'$*, and* $\mu = (\mu(t_1), \ldots, \mu(t_n))'$*. Then, it can be seen that*

$$\hat{\mathbf{U}} = \boldsymbol{\mu} + (\hat{f}_\lambda(t_1), \ldots, \hat{f}_\lambda(t_n))' = \boldsymbol{\mu} + H(\lambda)(\boldsymbol{Y} - \boldsymbol{\mu}) = (I_n - H(\lambda))\boldsymbol{\mu} + H(\lambda)\boldsymbol{Y},$$

Where $H(\lambda) = I_n - n\lambda M^{-1}[I_n - S(S'M^{-1}S)^{-1}S'M^{-1}]$ *is the smoothing matrix. Thus,* $\hat{\mathbf{U}}$ *has the typical form of a shrinkage estimator*.

3.3. Ornstein–Uhlenbeck Process and Exponential Spline

Consider an Ornstein–Uhlenbeck process $V(\cdot)$ that satisfies the SDE

$$\mathrm{d}V(t) = \theta(\mu - V(t))\mathrm{d}t + \sigma_V \mathrm{d}B(t), V(0) = 0, \quad (22)$$

where μ is the equilibrium value of the process, $\theta > 0$ is the speed of reversion, and σ_V is the volatility. It is easy to see that the Ornstein–Uhlenbeck process is a special case of (8) with $\beta_0(t) = \mu\theta$, $\beta_1(t) = -\theta$, and $\sigma(t) = \sigma_V$.

The mean and covariance functions of $V(\cdot)$ are

$$\mathrm{E}V(t) = \mu[1 - \exp(-\theta t)],$$

$$\mathrm{Cov}(V(s),V(t))=\frac{\sigma_V^2}{2\theta}\exp(-\theta|s-t|)[1-\exp(-2\theta s\wedge t)].$$

In this subsection we assume that $q \geq 1$. See Remark 3.7 for the exponential spline with $q = 0$. It is easy to check that $\Psi(t) = \exp(-\theta t)$, $h(t)=\mathrm{e}^{-2\theta t}\sigma_V^{-2}$ and $F(t,u)=I(u\leq t)\int_u^t \psi_{q-1}(t-\tau)\mathrm{e}^{-\theta\tau}b(\tau)\mathrm{d}\tau$. Assume that $U_i\sim \mathrm{N}(\mu_i,\sigma_i^2)$. Then the mean and covariance functions of $U(\cdot)$ are

$$\mathrm{E}U(t)=\sum_{i=0}^{q-1}\psi_i(t)\mu_i+A_q(t)\mu_q+\mu\int_0^t\psi_{q-1}(t-s)b(s)(1-\mathrm{e}^{-\theta s})\mathrm{d}s$$

$$\mathrm{Cov}(U(s),U(t))=\sum_{i=0}^{q-1}\psi_i(s)\psi_i(t)\sigma_i^2+A_q(s)A_q(t)\sigma_q^2+\sigma_V^2\int_0^{s\wedge t}F(s,u)F(t,u)\mathrm{e}^{2\theta u}\mathrm{d}u.$$

The differential operator in (12) reduces to $Lf(u) = [f^{(q+1)}(u) + \theta f^{(q)}(u)]\,\mathrm{e}^{\theta u}$. Therefore, the penalty term in (16) is equal to $(\lambda/\sigma_V^2)\int_0^T[f^{(q+1)}(t)+\theta f^{(q)}(t)]^2\mathrm{d}t$. Note that $f^{(q+1)} + \theta f^{(q)} = D^{q-1}(D^2 + \theta D)f$, where $D^2 + \theta D$ is differential operator for exponential spline (Wang, 2011).

3.3.1. Exponential spline—

Consider a special case of model (2) with $a_0(t) = \mathrm{e}^{-\theta t}$ and $b(t) = 1$. When $q = 1$, the stochastic process $U(\cdot)$ can be represented as a summation of drift terms and an integrated Ornstein–Uhlenbeck process

$$U(t)=U_0+U_1\frac{1-\mathrm{e}^{-\theta t}}{\theta}+\int_0^t(t-s)\mathrm{d}V(s)=U_0+U_1\frac{1-\mathrm{e}^{-\theta t}}{\theta}+\int_0^tV(s)\mathrm{d}s. \quad (23)$$

The penalty $\int_0^T[(D^2+\theta D)f]^2\mathrm{d}t$ is the same as that for an exponential spline. It is easy to check that the basis functions of the space $\mathscr{H}_0$ in Theorem 3.2 are $\psi_0(t) = 1$ and $A_1(t) = (1 - \mathrm{e}^{-\theta t})/\theta$, and the reproducing kernel of the space $\mathscr{H}_1$ is

$$R_1(s,t)=\frac{\sigma_V^2}{\theta^2}\left[s\wedge t-\frac{\mathrm{e}^{-\theta s}+\mathrm{e}^{-\theta t}}{\theta}\left(\mathrm{e}^{\theta s\wedge t}-1\right)+\frac{\mathrm{e}^{-\theta(s+t)}}{2\theta}\left(\mathrm{e}^{2\theta s\wedge t}-1\right)\right], \quad (24)$$

which is equivalent to that of an exponential spline (Wang, 2011, pp 41–44). Furthermore,

$$\mu(t)=\int_0^t\frac{F(t,u)\beta_0(u)}{\Psi(u)}\mathrm{d}u=\int_0^t\mu\theta\mathrm{e}^{\theta u}\mathrm{d}u\int_u^t\mathrm{e}^{-\theta\tau}\mathrm{d}\tau=\mu\theta^{-1}\left(\theta t-1+\mathrm{e}^{-\theta t}\right). \quad (25)$$

Therefore, when $\mu = 0$, the integrated Ornstein–Uhlenbeck process $U(\cdot)$ defined in (23) is the corresponding stochastic process for the exponential spline. The case when $\mu \neq 0$ provides an extension of the exponential spline. We note that the integrated Ornstein–Uhlenbeck stochastic process with drift $U(t)$ represents a major deviation from existing spline literature which involves the Brownian motion only.

The mean function of stochastic process $U(\cdot)$ is

$$\mathrm{E}[U(t)]=\mu_0+\mu_1\frac{1-\mathrm{e}^{-\theta t}}{\theta}+\mu\left(t+\frac{\mathrm{e}^{-\theta t}-1}{\theta}\right).$$

When μ = 0, $\mu=0, \overline{V}(t)=\Psi(t)\left[U_1+\int_0^t\sigma_V\mathrm{e}^{\theta s}dB(s)\right]$ in Remark 3.1 satisfies

$$\mathrm{d}\overline{V}(t)=-\theta\overline{V}(t)\mathrm{d}t+\sigma_V\mathrm{d}B(t), \overline{V}(0)=U_1, t\in[0,T].$$

Therefore, $U(t)=U_0+\int_0^t\overline{V}(s)\mathrm{d}s$ and $E[U(t)] = \mu_0 + \mu_1(1 - \mathrm{e}^{-\theta t})/\theta$.

When $q > 1$, the stochastic process $U(\cdot)$ can be represented as

$$U(t)=\sum_{i=1}^{q-1}\psi_i(t)U_i+A_q(t)U_q+\int_0^t\psi_{q-1}(t-s)V(s)\mathrm{d}s. \quad (26)$$

The penalty $\int_0^T[D^{(q-1)}(D^2+\theta D)f]^2\mathrm{d}t$ adds polynomials up to the order $q - 1$ to the kernel space $\mathscr{H}_0$. Specifically, the orthonormal basis functions of $\mathscr{H}_0$ are $\psi_0(t)$, $\psi_1(t)$, …, $\psi_{q-1}(t)$, and $A_q(t)$, where $A_q(t)=\int_0^t\psi_{q-1}(t-s)\mathrm{e}^{-\theta s}\mathrm{d}s$. The function $A_q(t)$ can be calculated recursively by $A_q(t) = \psi_{q-1}(t)/\theta - A_{q-1}(t)/\theta$ and $A_1(t) = (1 - \mathrm{e}^{-\theta t})/\theta$. The Chebyshev spline in this case may be called a polynomial-exponential spline.

3.3.2. Partial exponential spline—Consider another special case of model (2) with $a_0(t) = b(t) = 1$. For $q \geq 1$, the stochastic process $U(\cdot)$ can be represented as

$$U(t)=\sum_{i=0}^{q}\psi_i(t)U_i+\int_0^t\psi_{q-1}(t-s)V(s)\mathrm{d}s. \quad (27)$$

Replacing a Brownian motion by the Ornstein–Uhlenbeck process $V(\cdot)$, the stochastic process $U(\cdot)$ in (27) extends the stochastic model for polynomial splines (Wahba, 1990) (see Eq. (S.5) in the Supplementary Materials). The mean and covariance of the stochastic process $U(\cdot)$ are

$$\mathrm{E}U(t)=\sum_{i=0}^{q}\psi_i(t)\mu_i+\int_0^t\psi_{q-1}(t-s)\mu\left(1-\mathrm{e}^{-\theta s}\right)\mathrm{d}s,$$

$$\mathrm{Cov}(U(s),U(t))=\sum_{i=0}^{q}\psi_i(s)\psi_i(t)\sigma_i^2+\sigma_V^2\int_0^{s\wedge t}F(s,u)F(t,u)\mathrm{e}^{2\theta u}\mathrm{d}u,$$

where $F(t,u)=I(u\le t)\int_u^t\psi_{q-1}(t-\tau)\mathrm{e}^{-\theta\tau}\mathrm{d}\tau$.

For simplicity, we now consider the special case when $q = 1$. The stochastic process $U(\cdot)$ is given by

$$U(t)=U_0+U_1t+\int_0^t(t-s)\mathrm{d}V(s)=U_0+U_1t+\int_0^tV(s)\mathrm{d}s. \quad (28)$$

The integrated Ornstein–Uhlenbeck process in (28) extends the integrated Brownian motion [see (S.6) in Section S.5 of the Supplementary Materials]. The mean and covariance functions are

$$\mathrm{E}U(t)=\mu_0+\mu_1t+\mu\left(t+\frac{\mathrm{e}^{-\theta t}-1}{\theta}\right),$$

$$\mathrm{Cov}(U(s),U(t))=\sigma_0^2+st\sigma_1^2+R_1(s,t), \quad (29)$$

where R_1 is given by (24).

Note that Theorem 3.2 does not apply in this case since $a_0(t)\ne\Psi(t)$. Compared with the exponential spline model, the basis function $(1-\exp(-\theta t))/\theta$ in $\mathscr{H}_0$ has been replaced by the function $\psi_1(t) = t$ which is not orthogonal to the space $\mathscr{H}_1$ (Remark 3.2). Nevertheless, the stochastic process $U(\cdot)$ defined in (28) can be connected to the following partial spline model

$$Y_i=\alpha_1+\alpha_2t_i+f(t_i)+\varepsilon_i, i=1,\dots,n, t_i\in[0,T]. \quad (30)$$

Assume that $f\in\mathscr{M}\triangleq W_2^2[0,T]\ominus\{1,(1-\exp(-\theta t))/\theta\}$ under the inner product (14) with $Lf(u) = [D^2f(u) + \theta Df(u)]\,\mathrm{e}^{\theta u}$ and $h(u)=\mathrm{e}^{-2\theta u}\sigma_V^{-2}$. Let $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{f}_\lambda(t)$ be the solution to the following penalized least squares

$$\min_{\alpha_1\in R,\alpha_2\in R,f\in\mathscr{M}}\left\{\frac{1}{n}\sum_{i=1}^{n}[Y_i-(\alpha_1+\alpha_2t_i+f(t_i))]^2+\lambda\int_0^T(Lf(t))^2h(t)\mathrm{d}t\right\}. \quad (31)$$

Following similar arguments as in Section 3.2 one can show that $\lim_{a\to\infty}\hat{U}_a(t) = \mu(t)+\hat{\alpha}_1+\hat{\alpha}_2t+\hat{f}_\lambda(t)$, where $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{f}_\lambda(t)$ are the penalized least squares solutions to (31) with observations $Y_i = Y(t_i) - \mu(t_i)$ for $i = 1, \dots, n$ where $\mu(t)$ is given in (25).

3.4. Logistic Spline and Its Corresponding SDM

The logistic spline is a special case of the Chebyshev spline with $q = 0$ and penalty $\int_0^T [f'(t) - \beta_1(t) f(t)]^2 \mathrm{d}t$, where

$$\beta_1(t) = \theta\gamma \mathrm{e}^{-\theta t}/(1+\gamma \mathrm{e}^{-\theta t}), \theta>0, \gamma>0. \quad (32)$$

To construct the corresponding SDM, consider a process $V(\cdot)$ driven by the following linear SDE

$$\mathrm{d}V(t) = \beta_1(t) V(t) \mathrm{d}t + \sigma_V \mathrm{d}B(t), V(0) = 0. \quad (33)$$

Equation (33) is a special case of (8) with $\beta_0(t) = 0$, $\beta_1(t)$ is given in (32), and $\sigma(t) = \sigma_V$. It is not difficult to check that $\mu(t) = 0$, $\Psi(t) = (1 + \gamma)/(1 + \gamma e^{-\theta t})$, and

$$V(t) = \frac{\sigma_V B(t)}{1+\gamma \mathrm{e}^{-\theta t}} + \frac{\sigma_V \gamma}{1+\gamma \mathrm{e}^{-\theta t}} \int_0^t \mathrm{e}^{-\theta s} \mathrm{d}B(s).$$

Then, E $V(t) = 0$ and

$$\begin{aligned} \mathrm{Cov}(V(s), V(t)) &= \Psi(s)\Psi(t) \int_0^{s\wedge t} \left(\frac{\sigma_V}{\Psi(u)}\right)^2 \mathrm{d}u \\ &= \frac{\sigma_V^2}{(1+\gamma \mathrm{e}^{-\theta s})(1+\gamma \mathrm{e}^{-\theta t})} \left[s \wedge t + \frac{2\gamma}{\theta}\left(1 - \mathrm{e}^{-\theta(s\wedge t)}\right) + \frac{\gamma^2}{2\theta}\left(1 - \mathrm{e}^{-2\theta(s\wedge t)}\right)\right]. \end{aligned} \quad (34)$$

According to (21), the penalty $\int_0^T (Lf(t))^2 h(t) \mathrm{d}t = \sigma_V^{-2} \int_0^T [f'(t) - \beta_1(t) f(t)]^2 \mathrm{d}t$ is the same as that for the logistic spline up to a multiplying constant which can be absorbed into the smoothing parameter λ (Wang, 2011). Consider the special case of stochastic dynamic system (2) with $a_0(t) = \Psi(t)$ and $b(t) = 1$. The basis function of the space $\mathscr{H}_0$ is $\Psi(t)$ and the Green's function is $\Psi(t) I(u \le t)$. Then, the reproducing kernel for the space $\mathscr{H}_1$ is $R_1(s,t) = \Psi(s)\Psi(t) \int_0^{s\wedge t} [\sigma_V/\Psi(u)]^2 \mathrm{d}u$. R_1 has the form in (34) which is the same as the reproducing kernel for the logistic spline (Eq. (2.61) in Wang (2011)). Thus, the SDM (1), (2) and (33) consist of the stochastic model for the logistic spline.

Remark 3.6—*When $q = 1$, following similar argument it can be shown that the SDM (1), (2) and (33) is the stochastic model for the logistic spline discussed by Gu (2013, p. 161). The cases when $q > 1$ can be regarded as extensions of the logistic spline where basis functions $\psi_0(t), \ldots, \psi_{q-1}(t)$ are added to the null space $\mathscr{H}_0$.*

Remark 3.7—*The connection between an SDM driven by (33) and a Chebyshev spline discussed in this section holds for a general function $\beta_1(t)$. Specifically, consider the SDE (33) with an unspecified $\beta_1(t)$. Assume that $q = 0$, $a_0(t) = \Psi(t)$ and $b(t) = 1$. Then,* $V(t) = \sigma_V \Psi(t) \int_0^t \mathrm{d}B(s)/\Psi(s)$ and $U(t) = \Psi(t) U_0 + \sigma_V \Psi(t) \int_0^t \mathrm{d}B(s)/\Psi(s)$. *According to*

Theorem 3.2, the RKHS $W_2^1[0,T]=\{\Psi(t)\}\oplus\mathscr{H}_1$, *where* $\mathscr{H}_1$ *has the reproducing kernel* $R_1(s,t)=\Psi(s)\Psi(t)\int_0^{s\wedge t}[\sigma_V/\Psi(u)]^2\mathrm{d}u$. *The penalty of the corresponding Chebyshev spline is* $\sigma_V^{-2}\int_0^T[f'(t)-\beta_1(t)f(t)]^2\mathrm{d}t$. *The logistic spline is a special case with* $\beta_1(t)=\theta\gamma e^{-\theta t}/(1+\gamma e^{-\theta t})$. *It is not difficult to check that the function* $\beta_1(t)=-\theta$ *corresponds to the exponential spline with* $\beta_0(t)=0$. *Other functions may be considered for* $\beta_1(t)$. *For example,* $\beta_1(t)=-\theta t$ *corresponds to a Chebyshev spline with the kernel space spanned by* $\Psi(t)=e^{-\theta t^2/2}$, *the Gaussian function*.

4. EXTENDED STOCHASTIC DYNAMIC MODELS AND THEIR CONNECTIONS TO CHEBYSHEV SPLINES

Motivated by the Chebyshev splines (Karlin & Ziegler, 1966; Kimeldorf & Wahba, 1971; Gu, 2013), we now consider a more general dynamic system by replacing D^q in (2) with the following differential operator

$$\Pi_q=D\frac{1}{a_1}D\frac{1}{a_2}\cdots D\frac{1}{a_q}, \quad (35)$$

where $a_1, \ldots, a_q$ are strictly positive and $(i+1)$th differentiable functions with $a_i(0)=1$. Specifically, we assume the following stochastic dynamic system for $U(\cdot)$

$$\Pi_qU(t)=a_0(t)\Pi_qU(0)+b(t)V(t), q\ge 0, t\in[0,T]. \quad (36)$$

As in previous sections, we will consider two types of models for $V(\cdot)$: the general diffusion process determined by the SDE (3) and the Gaussian model driven by the linear SDE (8).

The following stochastic differential system has been considered by Kimeldorf & Wahba (1971):

$$\Pi_{q+1}U(t)=\frac{\mathrm{d}B(t)}{\mathrm{d}t}, \quad (37)$$

where $\Pi_{q+1}=D\frac{1}{a_0}D\frac{1}{a_1}D\frac{1}{a_2}\cdots D\frac{1}{a_q}$ is a $(q+1)$-order differential operator of Chebyshev splines (Wahba, 1978, 1990). The stochastic dynamic system (37) is equivalent to $\Pi_qU(t)=a_0(t)[\Pi_qU(0)+B(t)]$. The stochastic dynamic system (36) extends (37) in two aspects: (1) Brownian motion $B(t)$ is replaced by a general solution $V(\cdot)$ of an SDE, and (2) the coefficient $b(t)$ of $V(\cdot)$ does not necessarily equal to $a_0(t)$.

4.1. General Stochastic Dynamic Models

Following Kimeldorf & Wahba (1971) and Wahba (1990), denote

$$\Pi_0=I,$$

$$\Pi_1 = D\frac{1}{a_q},$$

$$\Pi_2 = D\frac{1}{a_{q-1}}D\frac{1}{a_q},$$
$$\vdots$$
$$\Pi_{q-1} = D\frac{1}{a_2}D\frac{1}{a_3}\cdots D\frac{1}{a_q},$$

where *I* is the identity operator. In addition, define the following functions:

$$\omega_0(t) = a_q(t),$$

$$\omega_1(t) = a_q(t)\int_0^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1},$$
$$\vdots$$
$$\omega_{q-1}(t) = a_q(t)\int_0^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}\int_0^{t_{q-1}} a_{q-2}(t_{q-2})\mathrm{d}t_{q-2}\cdots\int_0^{t_2} a_1(t_1)\mathrm{d}t_1,$$

$$\omega_q(t) = a_q(t)\int_0^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}\int_0^{t_{q-1}} a_{q-2}(t_{q-2})\mathrm{d}t_{q-2}\cdots\int_0^{t_2} a_1(t_1)\mathrm{d}t_1\int_0^{t_1} a_0(t_0)\mathrm{d}t_0.$$

Note that for *i*, *j* = 0, 1, 2, …, *q*, we have

$$(\Pi_i\omega_j)(0) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{else.} \end{cases} \quad (38)$$

Define

$$\begin{aligned} X(t) &= a_q(t)\int_0^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}\cdots\int_0^{t_2} a_1(t_1)\mathrm{d}t_1\int_0^{t_1} b(t_0)\mathrm{d}t_0\int_0^{t_0}\mathrm{d}V(u) \\ &= a_q(t)\int_0^t b(t_0)V(t_0)\mathrm{d}t_0\int_{t_0}^t a_1(t_1)\mathrm{d}t_1\cdots\int_{t_{q-2}}^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}. \end{aligned} \quad (39)$$

We note a typo in Equation (7.3) of Kimeldorf & Wahba (1971) where … $\int_0^{t_2} a_1(t_1)\mathrm{d}W(t_1)$ should be … $\int_0^{t_2} a_1(t_1)W(t_1)\mathrm{d}t_1$.

Following similar arguments as in the Section 2.2, the stochastic differential system (36) is equivalent to the following stochastic integration equation

$$U(t) = \sum_{i=0}^{q}\omega_i(t)U_i + X(t), \quad (40)$$

where $U_i = \Pi_i U(0)$. Let $\mu_i = \mathrm{E}U_i$ and $\sigma_i^2=\mathrm{Var}(U_i)$. Applying the Fubini's Theorem to the stochastic integration equation (40), we have the following results which extends Proposition 2.1.

Proposition 4.1—*Suppose that U*(·) *is given by the stochastic integration equation (40), where V*(·) *is driven by the SDE (3). Then,*

$$\mathrm{E}U(t)=\sum_{i=0}^{q}\omega_i(t)\mu_i+a_q(t)\int_0^t b(\tau)\mathrm{E}V(\tau)\mathrm{d}\tau\int_\tau^t a_1(t_1)\mathrm{d}t_1\cdots\int_{t_{q-2}}^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1},$$

$$\mathrm{Cov}(U(s),U(t))=\sum_{i=0}^{q}\omega_i(s)\omega_i(t)\sigma_i^2+\mathrm{Cov}(X(s),X(t)),$$

$$\mathrm{Cov}(X(s),X(t))=a_q(s)a_q(t)\int_0^s\int_0^t b(u)b(\upsilon)\mathrm{Cov}(V(u),V(\upsilon))\mathrm{d}u\mathrm{d}\upsilon\int_u^s a_1(s_1)\mathrm{d}s_1\int_\upsilon^t a_1(t_1)\mathrm{d}t_1 \cdots\int_{s_{q-2}}^s a_{q-1}(s_{q-1})\mathrm{d}s_{q-1}\int_{t_{q-2}}^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}.$$

4.2. Gaussian Models Driven by a Linear SDE

Denote

$$\tilde{F}(t,u)=a_q(t)\int_u^t b(\tau)\Psi(\tau)\mathrm{d}\tau\int_\tau^t a_1(t_1)\mathrm{d}t_1\int_{t_1}^t a_2(t_2)\mathrm{d}t_2\cdots\int_{t_{q-2}}^t a_{q-1}(t_{q-1})\mathrm{d}t_{q-1}.$$

The following results extend Theorem 3.1 when the SDM is driven by the linear SDE (8).

Theorem 4.1—*Suppose that the stochastic process V*(·) *is a Gaussian process driven by the linear SDE (8) and* $U_i\sim \mathrm{N}(\mu_i,\sigma_i^2)$ *for i* = 0, …, *q. Then, the stochastic process U*(·) *given by (36) or (40) is a Gaussian process with the following mean and covariance functions*

$$\mathrm{E}U(t)=\sum_{i=0}^{q}\omega_i(t)\mu_i+\int_0^t\tilde{F}(t,u)\frac{\beta_0(u)}{\Psi(u)}\mathrm{d}u,$$

$$\mathrm{Cov}(U(s),U(t))=\sum_{i=0}^{q}\omega_i(s)\omega_i(t)\sigma_i^2+\int_0^{s\wedge t}\tilde{F}(s,u)\tilde{F}(t,u)\left(\frac{\sigma(u)}{\Psi(u)}\right)^2\mathrm{d}u.$$

The proof of Theorem 4.1 is provided in the Supplementary Materials.

4.3. Connection to Chebyshev Splines

Again, let $\Psi(t) \triangleq \exp\left(\int_0^t \beta_1(s)\mathrm{d}s\right)$ and $h(t) \triangleq (\Psi(t)/\sigma(t))^2$ be a weight function. Consider the following differential operator

$$\tilde{L}f(u)=\frac{d}{\mathrm{d}u}\left(\frac{\Pi_q f(u)}{\Psi(u)}\right)=\frac{D\Pi_q f(u)-\beta_1(u)\Pi_q f(u)}{\Psi(u)}.$$

Consider the model space

$$\tilde{W}_2^{q+1}[0,T] \triangleq \left\{f{:}f,\Pi_1 f,\dots,\Pi_q f \text{ are absolutely continuous and} \int_0^T (\tilde{L}f(t))^2 h(t)\mathrm{d}t<\infty\right\}$$

with inner product

$$(f,g)=\sum_{i=0}^{q}\Pi_i f(0)\Pi_i g(0)+\int_0^T \tilde{L}f(u)\tilde{L}g(u)h(u)\mathrm{d}u. \quad (41)$$

Theorem 4.2—*Assume that $a_0(t) = \Psi(t)$ and $b(t) = 1$. Then, the space $\tilde{W}_2^{q+1}[0,T]$ is an RKHS with inner product (41). Let $\tilde{\mathscr{H}}_0=\left\{f \in \tilde{W}_2^{q+1}[0,T]{:}\tilde{L}f=0\right\}$ be the kernel of $\tilde{L}$. Then,* $\{\omega_0, \omega_1,\dots, \omega_q\}$ *forms an orthonormal basis of $\tilde{\mathscr{H}}_0$ and $\tilde{W}_2^{q+1}[0,T]$ can be decomposed into $\tilde{W}_2^{q+1}[0,T]=\tilde{\mathscr{H}}_0 \oplus \tilde{\mathscr{H}}_1$, where*

$$\tilde{\mathscr{H}}_0=span\{\omega_0,\omega_1,\dots,\omega_q\},$$

$$\tilde{\mathscr{H}}_1=\left\{f \in \tilde{W}_2^{q+1}[0,T]{:}f(0)=\Pi_1 f(0)=\dots=\Pi_q f(0)=0\right\}.$$

The reproducing kernels of $\tilde{\mathscr{H}}_0$ and $\tilde{\mathscr{H}}_1$ are respectively given by

$$\tilde{R}_0(s,t)=\sum_{i=0}^{q}\omega_i(s)\omega_i(t),$$

$$\tilde{R}_1(s,t)=\int_0^{s\wedge t}\tilde{F}(s,u)\tilde{F}(t,u)[h(u)]^{-1}\mathrm{d}u.$$

The proof of Theorem 4.2 is provided in the Supplementary Materials.

Now consider the nonparametric regression model (15) where $f \in \tilde{W}_2^{q+1}[0,T]$. A general Chebyshev spline is the solution to the penalized least squares (16) with $W_2^{q+1}[0,T]$ and L

being replaced by $\tilde{W}_2^{q+1}[0,T]$ and $\tilde{L}$, respectively. Again, up to a constant function, the smoothing spline estimate equals the posterior mean of the SDM (1), (8) and (36).

Specifically, assume that $U_0, \ldots, U_q \overset{iid}{\sim} \mathrm{N}(0,a)$. Denote $\hat{U}_a = \mathrm{E}(U(t)|Y(t_i), i = 1, \ldots, n)$ and $\tilde{f}_\lambda(t)$ as the penalized least squares solution with observations $(Y(t_1) - \tilde{\mu}(t_1), \ldots, Y(t_n) - \tilde{\mu}(t_n))'$, where

$$\tilde{\mu}(t) \triangleq \int_0^t \tilde{F}(t,s)\beta_0(s)/\Psi(s)\mathrm{d}s.$$

Then for any fixed $t \in [0, T]$, when $\lambda = \sigma^2/n$, it can be shown that $\lim_{a\to\infty} \hat{U}_a(t) = \tilde{\mu}(t) + \tilde{f}_\lambda(t)$.

4.4. Chebyshev Splines to Stochastic Dynamic Models

In the previous sections we started with SDE driven SDMs and built corresponding Chebyshev splines. We now describe a strategy that builds an SDE driven SDM that corresponds to the general Chebyshev spline defined in Section 4.3.

With the general Chebyshev spline, $a_0(t), \ldots, a_q(t)$ and a general weight function $h(t) > 0$ are given. Assume that $a_i(t) \in W_2^{i+1}[0,t]$ for $i = 0, \ldots, q$ are strictly positive functions with $a_i(0) = 1$. From SDMs to Chebyshev splines, we have set $a_0(t)=\Psi(t)=\exp(\int_0^t \beta_1(s)\mathrm{d}s)$ and $h(t) = (a_0(t)/\sigma(t))^2$ in the previous sections, where $\beta_1(t)$ and $\sigma(t)$ are given in the linear SDE (8). Reversely, now we set $\beta_1(t)=a_0'(t)/a_0(t)$ and $\sigma(t)=a_0(t)/\sqrt{h(t)}$. Consider the following stochastic dynamic system

$$\Pi_q U(t)=a_0(t)U_q+V(t), \quad (42)$$

$$\mathrm{d}V(t)=\beta_1(t)V(t)\mathrm{d}t+\sigma(t)\mathrm{d}B(t), V(0)=0, \quad (43)$$

with initial conditions $U_i = \Pi_i U(0)$ for $i = 0, \ldots, q$. Then, it is not difficult to show that the general Chebyshev spline equals the posterior mean of the SDM (1), (42), and (43).

5. APPLICATIONS

Proposition 3.1 implies that with appropriate choices of parameters in the SDM and the smoothing parameter, the posterior mean coincides with the corresponding Chebyshev spline estimate. In particular, as discussed in Section 3.3.1, with $U_0, U_1 \overset{iid}{\sim} \mathrm{N}(0,a)$, $a \to \infty$ and $n\lambda=\sigma^2/\sigma_V^2$, the posterior mean of the integrated Ornstein–Uhlenbeck process (23) coincides with the exponential spline estimate, in which σ^2 is the variance of $\varepsilon(t, \omega)$ the measurement error in Equation (1) and σ_V^2 is the volatility term of integrated Ornstein–Uhlenbeck process (23), reflecting the fluctuation of the process. We now use two real data applications to illustrate this theoretical result.

Glomerular filtration rate (GFR) measures the flow rate of filtered fluid through the kidney. Progression of kidney disease is often assessed by change in GFR. Therefore it is important to estimate the trajectory of GFR based on observations (Li et al., 2014). The left panel of Figure 1 shows estimated GFR (eGFR) from a patient with chronic kidney disease. We first consider an exponential spline model discussed in Section 3.3.1 since the profile is close to exponential decay. To estimate the speed of reversion parameter θ, as in Wang (2011), we first fit a nonlinear regression model which is motivated by Equation (25) as $y_i = \beta_1 + \beta_2 \exp(-\beta_3 t_i) + \varepsilon_i$, $i = 1, \ldots, 62$, and then set $\theta = \hat{\beta}_3$ where $\hat{\beta}_3$ is the non-linear least square estimate of β_3.We then fit the exponential spline using the ssr function in the R ASSIST package (Wang, 2011) with $\theta = \hat{\beta}_3$ and smoothing parameter selected by the generalized maximum likelihood method. The exponential spline fit is shown in the left panel of Figure 1 as the solid blue line.

The Ornstein–Uhlenbeck process provides a natural model for a process stabilizing around some equilibrium point. Such phenomena is often observed in the biological or biomedical application. For example, we consider modelling the eGFR profile by the SDM with integrated Ornstein–Uhlenbeck process (23), for which the parameters have natural interpretations in terms of the convergence of process. This is a special case of the stochastic velocity model with Ornstein–Uhlenbeck process discussed in Zhu, Song, & Taylor (2011). Thus the Markov chain Monte Carlo (MCMC) algorithm developed in Zhu, Song, & Taylor (2011) can be used to compute the posterior mean of the integrated Ornstein–Uhlenbeck process. Since patients will eventually lose renal function, we set the equilibrium value $\mu = 0$.We estimated the parameter θ as the posterior mean of MCMC samples of θ, and set U_0, $U_1 \overset{iid}{\sim} \mathrm{N}(0, 10^4)$. The MCMC was run for 45,000 iterations, in which the first 35,000 runs are discarded as the burn in and every 10th draw is saved. The posterior mean is shown in the left panel of Figure 1 as the red dashed line. As expected from the theoretical result, the exponential spline estimate is almost identical to the posterior mean of the integrated Ornstein–Uhlenbeck process.

We further present another real data application for the prostate specific antigen (PSA) profile commonly used to monitor recurrence of prostate cancer after treatment. Different from the GFR example, the rate of change of the PSA profile converges or the slope of the PSA profile is stabilized at a non-zero value (Figure 1, Zhu, Song, & Taylor, 2011). It’s reasonable to fit an SDM with integrated Ornstein–Uhlenbeck process and non-zero μ. For such a case, we illustrate that exponential spline can also be applied with a simple transformation. To make the estimates by two methods comparable, we fix $\hat{\theta} = 1.15$ and $\hat{\mu} = 0.385$ (posterior means in Table 3, Zhu, Song, & Taylor, 2011), while other parameters were estimated either by the MCMC algorithm for SDM or generalized maximum likelihood method for the exponential spline. To fit the exponential spline, the transformed variable $z(t) = y(t) - \hat{\mu}\{t + (e^{-\hat{\theta} t} - 1)/\hat{\theta}\}$ is used as observations. The right panel of Figure 1 shows the observations and estimated profiles. Again, the exponential spline estimate is almost identical to the posterior mean of the integrated Ornstein–Uhlenbeck process.

6. DISCUSSION

Under the general framework of isometric mapping between the Hilbert space spanned by a second order stochastic process and the RKHS generated by its covariance kernel, we establish a specific connection between an SDM driven by a linear SDE and the Chebyshev spline. This connection provides a statistical structure and mechanism for estimating sample paths in a stochastic dynamic model as well as a justification for the somewhat ad hoc penalty in a penalized least squares. Our results extend the well-known connection between the integrated Brownian and the polynomial spline to the connection between an SDM and the Chebyshev spline, which is mutually beneficial for these two different areas. For example, fitting spline models with large data can be computationally expensive. Instead, we may fit the corresponding SDMs with efficient algorithms based on the Markov property (Kohn & Ansley, 1987; Zhu, Song, & Taylor, 2011).

In this paper we have assumed $\sigma(t)$, $\beta_0(t)$, and $\beta_1(t)$ in (8) are known, which in practice need to be estimated. Under the assumption of $a_0(t) = \Psi(t)$ and $b(t) = 1$, the SDM (1), (2) and (8) are determined by $\sigma(t)$, $\beta_0(t)$, and $\beta_1(t)$. Similarly, the Chebyshev spline model may contain unknown parameters _ (Heckman & Ramsay, 2000). For example, the parameter θ in the penalty $\int_0^T [(D^2+\theta D)f]^2 \mathrm{d}t$ of the exponential spline is usually unknown and corresponds to an unknown speed of reversion in the integrated Ornstein–Uhlenbeck process. Estimation methods have been proposed for parameters in Chebyshev spline models (Heckman & Ramsay, 2000; Wang & Ke, 2009 and references therein) and SDMs (Zhu, Song, & Taylor, 2011, and references therein). Connection and comparison between these parameter estimation methods in these two different fields will be studied in the future.

One important feature of the SDM considered in this paper is that the process $V(\cdot)$ can be non-stationary which makes it more flexible. For special cases discussed in this paper, we have assumed that $\sigma(t)$ is a constant. With a general $\sigma(t)$, Equation (21) corresponds to an adaptive penalty for spatial inhomogeneous functions (see Pintore, Speckman, & Holmes, 2006; Liu & Guo, 2010, and references therein). As a future research topic, we will explore the corresponding non-stationary processes to fit spline models with varying smoothing parameters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the editors, an associate editor and two referees for constructive comments that substantially improved an earlier draft and revision. This study was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Maryland, USA (Ruzong Fan).

BIBLIOGRAPHY

Ansley CF, Kohn R. The equivalence of two stochastic approaches to spline smoothing. Journal of Applied Probability. 1986; 23:391–405.

Berlinet, A.; Thomas-Agnan, C. Reproducing Kernel Hilbert Spaces in Probability and Statistics. Boston: Kluwer; 2004.

Bishwal, JPN. Parameter Estimation in Stochastic Differential Equations. New York: Springer; 2008.

Cramér, H.; Leadbetter, MR. Stationary and Related Stochastic Processes; Sample Function Properties and Their Applications. New York: Wiley; 1967.

Furrer EM, Nychka DW. A framework to understand the asymptotic properties of Kriging and splines. Journal of the Korean Statistical Society. 2007; 36:57–76.

Griebel M, Hegland M. A finite element method for density estimation with Gaussian process priors. SIAM Journal on Numerical Analysis. 2010; 47:4759–4792.

Gu, C. Smoothing Spline ANOVA Models. 2nd ed.. New York: Springer; 2013.

Heckman N, Ramsay JO. Penalized regression with model-based penalties. Canadian Journal of Statistics. 2000; 28:241–258.

Ikeda, N.; Watanabe, S. Stochastic Differential Equations and Diffusion Processes. 2nd ed.. Amsterdam, New York: North-Holland Pub. Co.; 1989.

Karatzas, K.; Shreve, SE. Brownian Motion and Stochastic Calculus. Springer-Verlag; 1988. p. 354-355.

Karlin S, Ziegler Z. Chebyshevian spline functions. Siam Journal on Numerical Analysis. 1966; 3:514–543.

Kimeldorf GS, Wahba G. A correspondence between Bayesian estimation on stochastic process and smoothing by splines. The Annals of Mathematical Statistics. 1970a; 41:495–502.

Kimeldorf GS, Wahba G. Spline functions and stochastic processes. Sankhya: The Indian Journal of Statistics, Series A. 1970b; 32:173–180.

Kimeldorf GS, Wahba G. Some results on Tchebycheffian spline functions. Journal of Mathematical Analysis and Applications. 1971; 33:82–95.

Kohn R, Ansley CF. A new algorithm for spline smoothing based on smoothing a stochastic process. Siam Journal on Scientific and Statistical Computing. 1987; 8:33–48.

Li L, Chang A, Rostand SG, Hebert L, Appel LJ, Astor BC, Lipkowitz MS, Wright JT, Kendrick C, Wang X, Greene TH. A within-patient analysis for time-varying risk factors of CKD progression. Journal of the American Society of Nephrology. 2014; 25

Lindgren F, Rue H, Lindstrom J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society B. 2011; 73:423–498.

Liu Z, Guo W. Data driven adaptive spline smoothing. Statistica Sinica. 2010; 20:1143–1163.

Papaspiliopoulis O, Pokern Y, Roberts GO, Stuart AM. Nonparametric estimation of diffusions: A differential equations approach. Biometrika. 2012; 99:511–531.

Pintore A, Speckman P, Holmes CC. Spatially adaptive smoothing splines. Biometrika. 2006; 93:113–125.

Rue H, Martino S. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society B. 2009; 71:319–392.

Stuart AM. Inverse problems: A Bayesian perspective. Acta Numerica. 2010; 19:451–559.

Wahba G. Improper priors, spline smoothing and the problem of guarding against model errors in regression. Journal of the Royal Statistical Society B. 1978; 40:364–372.

Wahba, G. Spline Models for Observational Data. Philadelphia, PA: Society for Industrial Mathematics; 1990.

Wang, Y. Smoothing Splines, Methods and Applications. Boca Raton, FL: CRC Press, A Chapman & Hall Book; 2011.

Wang Y, Ke C. Smoothing Spline Semi-Parametric Nonlinear Regression Models. Journal of Computational and Graphical Statistics. 2009; 18:165–183.

Wecker WE, Ansley CF. The signal extraction approach to non-linear regression and spline smoothing. Journal of the American Statistical Association. 1983; 78:81–89.

Zhu B, Song PXK, Taylor JMG. Stochastic functional data analysis: a diffusion model-based approach. Biometrics. 2011; 67:1295–1304. [PubMed: 21418053]

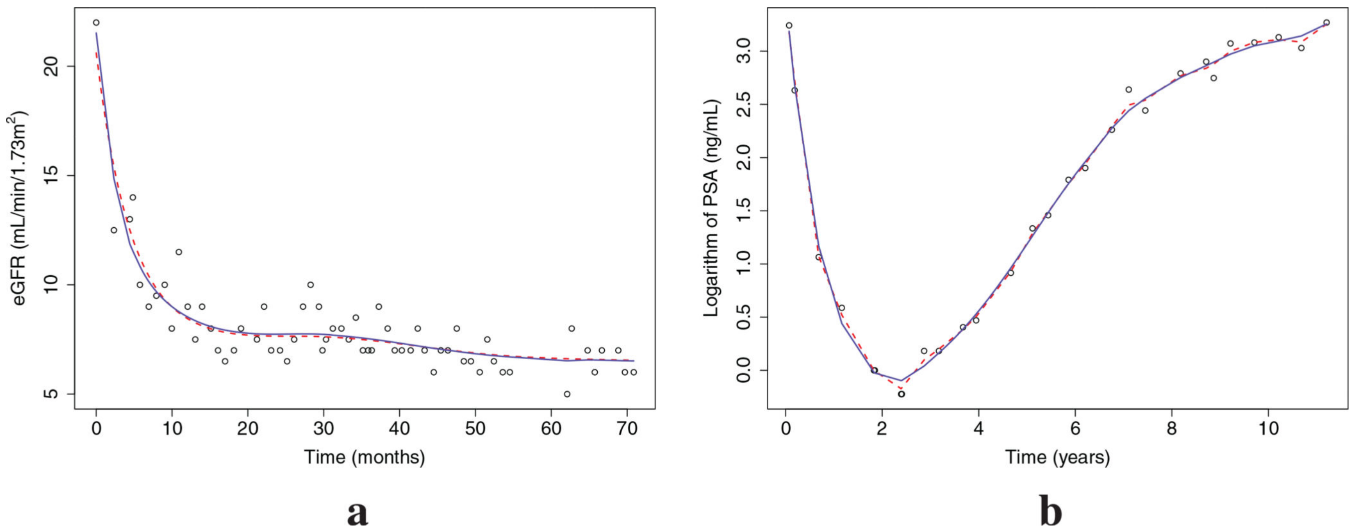

Figure 1.

GFR (a) and PSA (b) examples. Circles are observations, posterior means are dashed red lines, and exponential spline fits are solid blue lines.