Biometrika (2013), **xx**, x, *pp*. 1–11

Printed in Great Britain

Supplementary Material for "Smoothing Splines with Varying Smoothing Parameter"

BY XIAO WANG

Department of Statistics, Purdue University, 250 N. University Street, West Lafayette, Indiana, 47907, USA

wangxiao@purdue.edu

PANG DU

Department of Statistics, Virginia Tech, 406-A Hutcheson Hall Blacksburg, Virginia 24061, USA

pangdu@vt.edu

AND JINGLAI SHEN

Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland, 21250, USA

shenj@umbc.edu

PROOF OF THEOREM 1

For any $f, g \in W_2^m$ and $\delta \in \mathbb{R}$,

$$\psi(f+\delta g) - \psi(f) = 2\delta\psi_1(f,g) + \delta^2\Big[\int_0^1 g^2(t)d\omega_n(t) + \lambda\int_0^1 \rho(t)\{g^{(m)}(t)\}^2 dt\Big], \qquad (1)$$

where

$$\psi_1(f,g) = \int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}g(t)d\omega_n(t) + \lambda\int_0^1 \rho(t)f^{(m)}(t)g^{(m)}(t)dt. \qquad (2)$$

LEMMA 1. *$f \in W_2^m$ minimizes $\psi(f)$ in (2), if and only if, $\psi_1(f,g) = 0$ for all $g \in W_2^m$.*

Proof. If $f \in W_2^m$ minimizes $\psi(f)$, $\psi(f+\delta g) - \psi(f) \geq 0$ for all $g \in W_2^m$ and any $\delta \in \mathbb{R}$. Then $\psi_1(f,g) = 0$ follows since δ can be either negative or positive. On the other hand, if $\psi_1(f,g) = 0$, we have $\psi(f+\delta g) - \psi(f) \geq 0$ by (1). Thus, f minimizes $\psi(f)$. □

Let $g(t) = t^k (k = 0, \ldots, m-1)$ in (2). An application of Lemma 1 shows that if f minimizes $\psi(f)$, then

$$\int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}\, t^k d\omega_n(t) = 0 (k = 0, 1, \ldots, m-1).$$

We first have

$$\check{l}_1(f,1) - \check{l}_1(h,1) = \int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}d\omega_n(t) = 0.$$

Further,

$$\check{l}_2(f,1)-\check{l}_2(h,1)=\int_0^1\int_0^s \sigma^{-2}(t)\{f(t)-h(t)\}d\omega_n(t)ds=\int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}\,t\,d\omega_n(t)=0.$$

Similarly, it is shown that $\check{l}_k(f,1)=\check{l}_k(h,1)(k=1,\dots,m)$.

LEMMA 2. *If $f\in W_2^m$ satisfies $\check{l}_k(f,1)=\check{l}_k(h,1)(k=1,\dots,m)$, then for all $g\in W_2^m$,*

$$\psi_1(f,g)=\int_0^1 \psi_2(f)\,g^{(m)}(t)dt, \tag{3}$$

where

$$\psi_2(f)=\lambda\,\rho(t)\,f^{(m)}(t)+(-1)^m\,\{\check{l}_m(f,t)-\check{l}_m(h,t)\}. \tag{4}$$

Proof. If $f\in W_2^m$ satisfies $\check{l}_k(f,1)=\check{l}_k(h,1)$ $(k=1,\dots,m)$, we have

$$\begin{aligned}\int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}g(t)d\omega_n(t)&=\int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}\;\{g(t)-g(1)\}d\omega_n(t)\\&=-\int_0^1 \sigma^{-2}(t)\{f(t)-h(t)\}\int_t^1 g'(s)dsd\omega_n(t)\\&=-\int_0^1\{\check{l}_1(f,s)-\check{l}_1(h,s)\}\;g'(s)ds\\&=\cdots\\&=(-1)^m\int_0^1\{\check{l}_m(f,s)-\check{l}_m(h,s)\}\;g^{(m)}(s)ds\end{aligned}$$

Hence, (4) follows. □

Let $B^+=\{t\in[0,1]:\psi_2(f)>0\}$ and $B^-=\{t\in[0,1]:\psi_2(f)<0\}$. Define $g_+^{(m)}(t)=-I_{B^+}(t)$ and $g_-^{(m)}(t)=I_{B^-}(t)$, where I is the indicator function. Since $\psi_1(f,g)=0$ for all $g\in W_2^m$, we have $\psi_1(f,g_+)<0$ and $\psi_1(f,g_-)<0$, unless B^+ and B^- are of measure zero. This shows that $\psi_2(f)=0$ almost everywhere.

PROOFS OF THEOREM 2 AND COROLLARY 1

It follows from (9) that $r^{-1}(t)\hat{f}(t)=V_1(t)+V_2(t)+V_3(t)+V_4(t)$, where

$$\begin{aligned}V_1(t)&=\frac{d^m}{dt^m}\int_0^1 P(t,s)l_m(f_0,s)ds,\\V_2(t)&=\frac{d^m}{dt^m}\int_0^1 P(t,s)\{\check{l}_m(h,s)-\check{l}_m(f_0,s)\}ds,\\V_3(t)&=\frac{d^m}{dt^m}\int_0^1 P(t,s)\{l_m(\hat{f}-f_0,s)-\check{l}_m(\hat{f}-f_0,s)\}ds,\end{aligned}$$

and $V_4(t)=\sum_{k=1}^{2m}a_kC_k^{(m)}(t)$. Let $\bar{f}$ be the minimizer of the functional

$$\int_0^1 r^{-1}(s)\{f(s)-f_0(s)\}^2ds+\lambda\int_0^1\rho(t)\{f^{(m)}(s)\}^2ds.$$

Similar to Theorem 1, we have

$$(-1)^m \lambda \rho(t) \bar{f}^{(m)}(t) + l_m(\bar{f}, t) = l_m(f_0, t), \tag{5}$$

and

$$l_m(\bar{f}, t) = \int_0^1 P(t, s) l_m(f_0, s) ds. \tag{6}$$

Hence, $V_1(t) = r^{-1}(t)\bar{f}(t)$. Taking the mth derivative of both sides of (5), we get

$$(-1)^m \lambda \{\rho(t) \bar{f}^{(m)}(t)\}^{(m)} + r^{-1}(t)\bar{f}(t) = r^{-1}(t) f_0(t).$$

Recall that f_0 is $2m$ times continuously differentiable and $\beta = \lambda^{-1/(2m)}$. Combining this with (6), it is easy to show that $\bar{f}^{(k)}(t) \to f_0^{(k)}(t) (k = 1, \ldots, 2m)$ as $\beta \to \infty$. Therefore,

$$V_1(t) = r^{-1}(t) f_0(t) + (-1)^{m-1} \lambda \{\rho(t) f_0^{(m)}(t)\}^{(m)} + o(\lambda).$$

PROPOSITION 1. *Assume that a function $\tilde{J}(t, s)$ satisfies*

$$(-1)^m \frac{\partial^m}{\partial s^m} \tilde{J}(t, s) = \frac{\partial^m}{\partial t^m} P(t, s), \quad t, s \in [0, 1].$$

Then,

$$\tilde{J}(t, s) + \sum_{k=0}^{m-1} (-1)^k \zeta_{k+1}(s) \tilde{J}_k(t) = \frac{r(s)}{r(t)} J(t, s),$$

where

$$\zeta_k(s) = \int_s^1 \cdots \int_{s_{k-3}}^1 \int_{s_{k-2}}^1 ds_{k-1} ds_{k-2} \cdots ds_1, \quad \tilde{J}_k(t) = \frac{\partial^k}{\partial s^k} \tilde{J}(t, s) \mid_{s=1},$$

and $J(t, s)$ is the Green's function for

$$(-1)^m \lambda r(t) \{\rho(t) u^{(m)}(t)\}^{(m)} + u(t) = 0. \tag{7}$$

Proof. Consider the integral equation $(-1)^m \lambda \rho(t) f^{(m)}(t) + l_m(f, t) = l_m(g, t)$. If we write this equation as a differential equation for f, we have

$$(-1)^m \lambda \{\rho(t) f^{(m)}(t)\}^{(m)} + r^{-1}(t) f(t) = r^{-1}(t) g(t). \tag{8}$$

Further writing (8) as a differential equation for $l_m(f, t)$, we obtain

$$(-1)^m \lambda \rho(t) \frac{d^m}{dt^m} \Big\{ r(t) \frac{d^m}{dt^m} l_m(f, t) \Big\} + l_m(f, t) = l_m(g, t). \tag{9}$$

It follows from (9) that $l_m(f, t) = \int_0^1 P(t, s) l_m(g, s) ds$. Hence,

$$\begin{aligned} r^{-1}(t) f(t) = \frac{d^m}{dt^m} l_m(f, t) &= \int_0^1 \frac{\partial^m}{\partial t^m} P(t, s) l_m(g, s) ds = (-1)^m \int_0^1 \frac{\partial^m}{\partial s^m} \tilde{J}(t, s) l_m(g, s) ds \\ &= (-1)^m \sum_{k=1}^{m} (-1)^{k+1} l_{m-k+1}(g, 1) \tilde{J}_{m-k}(t) + \int_0^1 \tilde{J}(t, s) r^{-1}(s) g(s) ds \\ &= \int_0^1 \Big\{ \sum_{k=0}^{m-1} (-1)^k \zeta_{k+1}(s) \tilde{J}_k(t) + \tilde{J}(t, s) \Big\} r^{-1}(s) g(s) ds \end{aligned}$$

From (7) and (8), we have $f(t) = \int_0^1 J(t,s)g(s)ds$. Thus,

$$\int_0^1 \Big[r^{-1}(t)J(t,s) - \big\{ \sum_{k=0}^{m-1} (-1)^k \zeta_{k+1}(s)\tilde{J}_k(t) + \tilde{J}(t,s)\big\} r^{-1}(s)\Big] g(s)ds = 0.$$

Since the above equation is true for all $g \in L_2[0,1]$, the proposition follows. □

By applying Proposition 1, we have, for any $t \in (0,1)$,

$$\begin{aligned} V_2(t) &= \int_0^1 (-1)^m \frac{\partial^m}{\partial s^m} \tilde{J}(t,s) \check{l}_m(h - f_0, s)ds \\ &= \int_0^1 \tilde{J}(t,s) d\{\check{l}_1(h - f_0, s)\} + (-1)^m \sum_{k=1}^{m-1} (-1)^{k-1} \tilde{J}_{m-k}(t) \check{l}_{m-k+1}(h - f_0, 1) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{r(t_i)}{r(t)} J(t, t_i)\sigma^{-1}(t_i)\epsilon_i + \text{ higher order terms }. \end{aligned}$$

Eggermont & LaRiccia (2006) established the uniform error bounds for regular smoothing splines. We adopt the same approach as in Eggermont & LaRiccia (2006) for adaptive smoothing splines; the details are omitted here. For $\lambda \ll (n^{-1} \log n)^{2m/(1+4m)}$, we obtain

$$\|\hat{f} - f_0\| = O\Big[\frac{\max\big(\{\log 1/\lambda\}^{1/2}, \{\log\log n\}^{1/2}\big)}{n^{1/2}\lambda^{1/(4m)}} \Big].$$

Therefore, $\|V_3\| \le O(\beta^m) D_n \|\hat{f} - f_0\|$. Finally, it follows from Lemma 4 in next section that $\|V_4\|$ is of order $O(\beta^m)\exp\{-\beta Q_\beta(t)\{Q_\beta(1) - Q_\beta(t)\}\}$, and thus a negligible term in the asymptotic expansion of $r^{-1}(t)\hat{f}(t)$. This completes the representation for $\hat{f}$.

Proof of Corollary 1. Define $U_\beta(t) = \frac{1}{n}\sum_{i=1}^n \{\sigma(t_i)/q(t_i)\} J(t,t_i)\epsilon_i$. For any $t \in (0,1)$, the variance of $U_\beta(t)$ is given by

$$\begin{aligned} E\{U_\beta^2(t)\} &= \frac{1}{n}\int_0^1 \frac{\sigma^2(s)}{q^2(s)} J^2(t,s) d\omega_n(s) \\ &= \frac{\beta^2}{n}\int_0^1 \frac{\sigma^2(s)}{q^2(s)} \varrho(s)^2 \{Q'_\beta(s)\}^2 L^2\big\{\beta|Q_\beta(t) - Q_\beta(s)|\big\} d\omega_n(s) \\ &= \frac{\beta^2}{n}\int_0^1 r(s)\{Q'_\beta(s)\}^2 L^2\big\{\beta|Q_\beta(t) - Q_\beta(s)|\big\} ds + \text{ higher order terms} \\ &= \frac{\beta}{n}\int_{-\beta\{Q_\beta(1) - Q_\beta(t)\}}^{\beta Q_\beta(t)} r\big[\Upsilon_\beta\{Q_\beta(t) - \frac{x}{\beta}\}\big] Q'\big[\Upsilon_\beta\{Q_\beta(t) - \frac{x}{\beta}\}\big] L^2(|x|)dx + \text{ higher order terms} \\ &= \frac{\beta}{n} r(t) Q'(t) \int_{-\infty}^{\infty} L^2(|x|)dx + \text{ higher order terms} \end{aligned}$$

where Υ_β is the inverse function of Q_β, the second last equality is obtained by letting $\beta\big\{Q_\beta(t) - Q_\beta(s)\big\} = x$, and the last equality is from a simple Taylor expansion. We invoke the Lindeberg–Lévy Central Limit Theorem to verify that $(n/\beta)^{1/2} U_\beta(t)$ converges in distribution to $N\big\{0, r(t)^{1-1/(2m)} \rho(t)^{-1/(2m)} L_0\big\}$. Lindeberg's condition is easily satisfied if the ϵ_i have the finite fourth moment. The corollary follows by incorporating the bias term from Theorem 2.

Green's Functions

A critical step in the representation of the adaptive smoothing spline estimator is to use the Green's function to solve a two-point boundary value problem with a large parameter. This step is of independent technical interest. In this section, we derive the Green's functions for the time-varying ordinary differential equation

$$(-1)^m \lambda w(t)\big\{r(t)F^{(m)}(t)\big\}^{(m)} + F(t) = G(t), \quad t \in [0,1], \tag{10}$$

subject to the natural boundary conditions.

We firstly focus on the homogeneous equation with a large parameter $\beta = \lambda^{-1/(2m)}$:

$$(-1)^m w(t)\{r(t)F^{(m)}(t)\}^{(m)} + \beta^{2m}F(t) = 0, \quad t \in [0,1]. \tag{11}$$

We exploit the techniques in Coddington & Levinson (1955) to establish approximations to its fundamental solutions and their derivatives up to the $(m-1)$th order. Let

$$\mu_k = \cos\Big\{\frac{(1+2k)\pi}{2m}\Big\}, \omega_k = \sin\Big\{\frac{(1+2k)\pi}{2m}\Big\} \; (k = 0, \dots, \frac{m}{2}-1)$$

when m is even, and let

$$\mu_k = \cos\Big(\frac{k\pi}{m}\Big), \omega_k = \sin\Big(\frac{k\pi}{m}\Big) \; (k = 1, \dots, \frac{m-1}{2})$$

when m is odd.

LEMMA 3. *Given an m, suppose that $w(\cdot)$ and $r(\cdot)$ are positive functions in $C^{m+1}[0,1]$. Denote $\gamma(t) = \{w(t)r(t)\}^{-1/(2m)}$. Then, for all β sufficiently large, a fundamental solution of (11) is given by*

$$h_k(t) = \exp\Big\{\beta\big(\pm\mu_k \pm \imath\,\omega_k\big)Q_\beta(t)\Big\},$$

for any k defined above (corresponding to either an even m or an odd m), where

$$Q_\beta(s) = \int_0^t \gamma(s)\big\{1 + O(\beta^{-1})\big\}ds.$$

Furthermore, for each k and $\ell = 1, \dots, m-1$,

$$h_k^{(\ell)}(t) = \Big\{\beta(\pm\mu_k \pm \imath\omega_k)\gamma(t)\Big\}^\ell \exp\Big[\beta\big(\pm\mu_k \pm \imath\,\omega_k\big)\int_0^t \gamma(s)\big\{1 + O(\beta^{-1})\big\}ds\Big].$$

Proof. We exploit the techniques in Coddington & Levinson (1955) to establish approximations to its fundamental solutions of (10) and their derivatives up to the $(m-1)$th order. The homogeneous ordinary differential equation (11) can be written as

$$(-1)^m w(t)\big\{r(t)F^{(2m)}(t) + mr'(t)F^{(2m-1)}(t) + \cdots + r^{(m)}(t)F^{(m)}(t)\big\} + \beta^{2m}F(t) = 0$$

Since $w(t)r(t) > 0$ on $[0,1]$, we have

$$F^{(2m)}(t) + \beta a_1(t,\beta)F^{(2m-1)}(t) + \beta^2 a_2(t,\beta)F^{(2m-2)}(t) + \cdots + \beta^{2m}a_{2m}(t,\beta)F(t) = 0 \tag{12}$$

where

$$a_1(t,\beta) = (-1)^{m-1}\frac{mr'(t)}{\beta r(t)}, \;\; \dots, \;\; a_m(t,\beta) = (-1)^{m-1}\frac{r^{(m)}(t)}{\beta^m r(t)},$$

and

$$a_\ell(t,\beta) \equiv 0 \; (\ell = m+1,\ldots,2m-1), \quad a_{2m}(t,\beta) = \frac{(-1)^{m-1}}{w(t)r(t)}$$

In particular, $a_i(t,\beta) = (-1)^{m-1}p_i(t)/\{\beta^i w(t)r(t)\}$ $(i = 1,\ldots,2m-1)$, where $p_i(t)$ is a linear combination of finite products of $w(t)$ and the derivatives of $r(t)$ up to the mth order. Hence each $a_i(t,\beta)$ can be written as $a_i(t,\beta) = \sum_{j=0}^{m} a_{ij}(t)\beta^{-j}$ for suitable functions $a_{ij}(t)$ which are at least in $C^1[0,1]$. Therefore, $a_i(t,\beta) \to 0$ $(i = 1,\ldots,2m-1)$ uniformly in t on $[0,1]$ as $\beta \to \infty$. Let $x_1(t) \equiv F(t)$ and $x_{i+1}(t) \equiv x_i'(t)/\beta$ $(i = 1,\ldots,2m-1)$. Then the ordinary differential equation (12) can be written as

$$x_i' = \beta x_{i+1}, \; (i = 1,\ldots,2m-1), \quad x_{2m}' = -\beta\big\{a_{2m}(t,\beta)x_1 + \cdots + a_1(t,\beta)x_{2m}\big\} \tag{13}$$

Let $x = (x_1,\ldots,x_{2m})^{\mathrm{T}} \in \mathbb{R}^{2m}$. The ordinary differential equation (13) becomes $\dot{x} = \beta A(t,\beta)x$, where

$$A(t,\beta) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_{2m}(t,\beta) & -a_{2m-1}(t,\beta) & \cdots & \cdots & -a_2(t,\beta) & -a_1(t,\beta) \end{bmatrix} = \sum_{k=0}^{2m-1} \beta^{-k} A_k(t)$$

Here

$$A_0(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_{2m,0}(t) & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

where $a_{2m,0}(t) = (-1)^{m-1}/\{w(t)r(t)\}$, and $A_k(t)$ $(k = 0,\cdots,2m-1)$ are at least of the class $C^1[0,1]$. Note that for each $t \in [0,1]$, $A_0(t)$ has $2m$ distinct eigenvalues given by $(\pm\mu_k \pm \imath\omega_k)/\{w(t)r(t)\}^{1/2m}$, where μ_k and ω_k are defined before this proposition. Since the angle between any two distinct roots of the equation $(-1)^m x^{2m} + 1 = 0$ is given by $\pm q\pi/m$, where $q \in \{1,\cdots,m\}$, it follows from the similar argument in Abramovich & Grinshtein (1999) that the hypothesis of Coddington & Levinson (1955) holds. Let $\{\lambda_k(t)\}$ represent the $2m$ distinct eigenvalues of $A_0(t)$ (or equivalently the distinct roots of $(-1)^m x^{2m} + 1 = 0$). Consequently, by using Coddington & Levinson (1955), we conclude that the fundamental solution $X(t) \in \mathbb{R}^{2m\times 2m}$ of the homogeneous ordinary differential equation $\dot{X}(t) = \beta A(t,\beta)X(t)$ can be represented as $X(t) = B_0(t)\widehat{P}(t,\beta)e^{\beta Q_0(t)+Q_1(t)} + O(\beta^{-1})$, where $Q_0'(t) = \mathrm{diag}(\lambda_1(t),\cdots,\lambda_{2m}(t))$ with $Q(0) = 0$, $Q_1'(t)$ is a diagonal matrix whose diagonal entries are bounded on $[0,1]$ with $Q_1(0) = 0$,

$$\widehat{P}(t,\beta) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ O(\beta^{-1}) & \cdots & \cdots & \cdots & O(\beta^{-1}) & 1 + O(\beta^{-1}) \end{bmatrix}$$

and

$$B_0(t) = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1(t) & \cdots & \lambda_{2m}(t) \\ \lambda_1^2(t) & \cdots & \lambda_{2m}^2(t) \\ \vdots & \cdots & \vdots \\ \lambda_1^{2m-1}(t) & \cdots & \lambda_{2m}^{2m-1}(t) \end{bmatrix}$$

Recall that $F^{(\ell)} = \beta^\ell x_{\ell+1}$ and each $\lambda_k(t)$ is continuous and bounded away from 0 on $[0, 1]$. These results, together with the discussion in (Coddington & Levinson, 1955, Section 6.5), show that the ℓth derivative of each fundamental solution is of the form $\{\beta\lambda_k(t)\}^\ell \exp\left[\beta(\int_0^t \lambda_k(s)\{1 + O(\beta^{-1})\}ds)\right]$ ($\ell = 0, \ldots, m-1$). Hence, the lemma follows.□

The functions represented by $O(\beta^{-1})$ in Lemma 3 might be different, although they have the same order.

PROPOSITION 2. *Let $P(t, s)$ be the Green's function for (10). Let $P(t, s)$ equal to $P_1(t, s)$ when $t > s$ and $P_2(t, s)$ when $t < s$. If*

$$\frac{\partial^k}{\partial t^k} P_1(t, s)\Big|_{s=t} = \frac{\partial^k}{\partial t^k} P_2(t, s)\Big|_{s=t} \ \ (k = 0, 1, \ldots, 2m-2),$$

and

$$\frac{\partial^{2m-1}}{\partial t^{2m-1}} P_1(t, s)\Big|_{s=t} - \frac{\partial^{2m-1}}{\partial t^{2m-1}} P_2(t, s)\Big|_{s=t} = \frac{1}{(-1)^m \lambda w(t) r(t)},$$

then $F_0(t) = \int_0^1 P(t, s)G(s)$ is a solution of (10).

Proof. We write $F_0(t) = \int_0^t P_1(t, s)G(s)ds + \int_t^1 P_2(t, s)G(s)ds$ for all $t \in [0, 1]$. Since $\partial^k P(t, s)/\partial t^k$ ($k = 0, 1, \ldots, 2m-1$) is continuous at $s = t$, we have

$$F_0^{(k)}(t) = \int_0^1 \frac{\partial^k}{\partial t^k} P(t, s)G(s)ds \ \ (k = 1, 2, \ldots, 2m-1).$$

Furthermore, due to the jump of $\partial^{2m-1}P(t, s)/\partial t^{2m-1}$ at $s = t$, we have

$$F_0^{(2m)}(t) = \frac{G(t)}{(-1)^m \lambda w(t) r(t)} + \int_0^1 \frac{\partial^{2m}}{\partial t^{2m}} P(t, s)G(s)ds.$$

For any fixed s, $P(t, s)$ is a solution of the homogeneous differential equation (11). This verifies that $F_0(t) = \int_0^1 P(t, s)G(s)ds$ is a solution of (10). □

Now consider a special case when $w(t) \equiv 1$ and $r(t) \equiv 1$ for $t \in [0, 1]$. This leads to the time-invariant ordinary differential equation

$$(-1)^m \lambda F^{(2m)}(t) + F(t) = G(t), \quad t \in [0, 1].$$

Under this circumstance, the Green's function has been obtained explicitly before, see Berlinet & Thomas-Agnan (2004) and Wang et al. (2010). The homogeneous differential equation $\lambda F^{(2m)}(t) + F(t) = 0$ has $2m$ solutions $\exp\{(\pm\mu_k \pm \imath\omega_k)\beta t\}$. Let $K(t, s)$ be the corresponding Green's function. It turns out that $K(t, s)$ is a function of $|t - s|$ and can be written as

$K(t,s) \equiv \beta\, L(\beta|t-s|)$. The explicit formula for L is

$$L(t) \equiv \sum_{k=0}^{\frac{m}{2}-1} e^{-\mu_k t}\big\{\, c_k\, \cos(\omega_k t) + d_k\, \sin(\omega_k t)\big\}, \tag{14}$$

for an even m, and

$$L(t)\ \equiv\ c_0\, e^{-t} + \sum_{k=1}^{(m-1)/2} e^{-\mu_k t}\big\{\, c_k\, \cos(\omega_k t) + d_k\, \sin(\omega_k t)\big\}, \tag{15}$$

for an odd m. Similar to Proposition 2, the coefficients c_k, d_k can be uniquely determined from the following conditions:

$$L^{(k)}(t)\big|_{t=0} = 0\ \ (k = 1, 3, \ldots, 2m-3)\ \ \text{ and }\ \ L^{(2m-1)}(t)\big|_{t=0} = \frac{(-1)^m}{2}. \tag{16}$$

Define

$$P(t,s) = \beta\ \varrho(s)\ Q'_\beta(s)\ L\Big\{\beta|Q_\beta(t) - Q_\beta(s)|\Big\}, \tag{17}$$

where $\varrho(s) = 1 + O(\beta^{-1})$ holds uniformly in $s \in [0,1]$. It is easy to verify that $P(t,s)$ in (17) satisfies the conditions in Proposition 2. Therefore, it is the Green's function for (10).

The homogeneous differential equation (11) has $2m$ linearly independent solutions:

$$\begin{aligned} &C_{k1}(t) = e^{-\beta\mu_k Q_\beta(t)} \cos\{\beta\omega_k Q_\beta(t)\},\ \ C_{k2}(t) = e^{-\beta\mu_k Q_\beta(t)} \sin\{\beta\omega_k Q_\beta(t)\}, \\ &C_{k3}(t) = e^{-\beta\mu_k \{Q_\beta(1)-Q_\beta(t)\}} \cos\{\beta\omega_k Q_\beta(t)\},\ \ C_{k4}(t) = e^{-\beta\mu_k \{Q_\beta(1)-Q_\beta(t)\}} \sin\{\beta\omega_k Q_\beta(t)\}, \end{aligned} \tag{18}$$

where $k = 0, \ldots, m/2-1$ when m is even or $k = 0, \ldots, (m-1)/2$ when m is odd. These solutions, together with the $2m$ boundary conditions for (10) $F^{(k)}(0) = 0$ and $F^{(k)}(1) = G^{(k)}(1)$, $k = 0, 1, \ldots, m-1$, yield:

LEMMA 4. *The solution to (*10*) subject to the boundary conditions can be written as*

$$F(t) = \int_0^1 P(t,s)G(s)ds + \sum_k \Big\{a_{k1}C_{k1}(t) + a_{k2}C_{k2}(t) + a_{k3}C_{k3}(t) + a_{k4}C_{k4}(t)\Big\}, \tag{19}$$

*where the Green's function $P(t,s)$ is given in (*17*), and the coefficients $a_{k1}, a_{k2}, a_{k3}, a_{k4}$ are unique and bounded for all β sufficiently large.*

Proof. Consider an even m; the case of an odd m is similar and is omitted. For notational convenience, let

$$\chi_{j+1} = -\mu_j + \imath\omega_j\ (j = 0, \ldots, \frac{m}{2} - 1),$$

where each $\mu_j > 0$ and $\omega_j > 0$, and $\gamma(t) = \{w(t)r(t)\}^{-1/(2m)}$. Let $\lambda_{2j-1}(t) = \chi_j\gamma(t)$, $\lambda_{2j}(t) = \overline{\chi}_j\gamma(t)(t)$ $(j = 1, \ldots, m/2)$, where $\overline{\chi}_j$ denotes the conjugate of χ_j. Let $h_k(t) = \exp\big\{\beta(\int_0^t \lambda_k(s)\{1 + O(\beta^{-1})\}ds)\big\}$ be the corresponding m fundamental solutions. Similarly, define $g_k(t) = \exp\big[\beta\{\int_t^1 \lambda_k(s)(1 + O(\beta^{-1}))ds\}\big]$. Hence, for each odd k, $h_{k+1}(t)$ is the conjugate of $h_k(t)$. Define the complex number $p_k = b_{k1} + \imath b_{k2}$, $p_k^+ = b_{k1}^+ + \imath b_{k2}^+$ $(k = 1, 3, \ldots, m -$

1). Therefore, $F(t)$ can be written as

$$F(t) = F_0(t) + \sum_{j=1}^{m/2} \{p_{2j-1}h_{2j-1}(t) + \overline{p}_{2j-1}h_{2j}(t) + p^+_{2j-1}g_{2j-1}(t) + \overline{p}^+_{2j-1}g_{2j}(t)\},$$

where b_{kj}, b^+_{kj} are to be determined. Define $b = \Big(b_{11}, b_{12}, \dots, b_{\frac{m}{2}1}, b_{\frac{m}{2}1}, b^+_{11}, b^+_{12}, \dots, b^+_{\frac{m}{2}1}, b^+_{\frac{m}{2}1}\Big)$ $\|G\|$, and $G = \big(\|G\|, G(1), G'(1), \dots, G^{(m-1)}(1)\big)$. It is easy to verify that the coefficients $a_{k1}, a_{k2}, a_{k3}, a_{k4}$ are unique and bounded for all β sufficiently large if and only if b is so. The boundary conditions lead to the linear equation $DB\widetilde{b} = v$, where

$$\widetilde{b} = (p_1, \overline{p}_1, p_3, \overline{p}_3, \dots, p_{m/2-1}, \overline{p}_{m/2-1}, p^+_1, \overline{p}^+_1, p^+_3, \overline{p}^+_3, \dots, p^+_{m/2-1}, \overline{p}^+_{m/2-1}) \in \mathbb{C}^{2m},$$

$v^{\mathrm{T}} = [v_0, v_1]$ with

$$v_0 = \Big[-F_0(0), -\frac{F_0'(0)}{\beta}, \dots, -\frac{F_0^{(m-1)}(0)}{\beta^{m-1}} \Big],$$

$$v_1 = \Big[-F_0(1) + G(1), \frac{-F_0'(1) + G'(1)}{\beta}, \dots, \frac{-F_0^{(m-1)}(1) + G^{(m-1)}(1)}{\beta^{m-1}} \Big],$$

and the matrices $D \in \mathbb{R}^{2m\times 2m}$ and $B \in \mathbb{C}^{2m\times 2m}$ are

$$D = \mathrm{diag}\Big(1, \gamma(0), \gamma^2(0), \dots, \gamma^{m-1}(0), 1, \gamma(1), \gamma^2(1), \dots, \gamma^{m-1}(1)\Big),$$

and

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where the matrix blocks $B_{ij} \in \mathbb{C}^{m\times m}$ are obtained from the derivatives of h_{2j-1} at $t = 0$ given in Lemma 3:

$$B_{11} = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ \chi_1 & \overline{\chi}_1 & \chi_2 & \overline{\chi}_2 & \cdots & \chi_{\frac{m}{2}} & \overline{\chi}_{\frac{m}{2}} \\ \chi_1^2 & \overline{\chi}_1^2 & \chi_2^2 & \overline{\chi}_2^2 & \cdots & \chi_{\frac{m}{2}}^2 & \overline{\chi}_{\frac{m}{2}}^2 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \chi_1^{m-1} & \overline{\chi}_1^{m-1} & \chi_2^{m-1} & \overline{\chi}_2^{m-1} & \cdots & \chi_{\frac{m}{2}}^{m-1} & \overline{\chi}_{\frac{m}{2}}^{m-1} \end{bmatrix},$$

and B_{22} has the similar structure, and each entry of B_{12}, B_{21} is of order $O(e^{-\beta})$. Therefore, B_{11} is a Vandermonde matrix and is invertible, and there is a uniform bound on the entries of the inverse of B_{11} for all large β. The similar argument can be applied to B_{22}. Consequently, B and D are invertible and the elements of B^{-1}, D^{-1} are uniformly bounded. Furthermore, using the argument in Proposition 2 and uniform bounds on the derivatives of $\gamma(t)$ up to the $(m-1)$th order, it can be shown that for $t_* = 0$ or 1,

$$\frac{\big|F_0^{(j)}(t_*)\big|}{\beta^j} \le 2m\kappa \int_0^1 \beta e^{-\beta\varrho\tau} d\tau \, \|G\| \le (2m\kappa/\varrho) \, \|G\| \;\; (j = 1, \dots, m-1).$$

for some suitable constants $\kappa, \varrho > 0$. As a result, the equation $DBx = \mathbf{v}$ has a unique solution $\widetilde{b}$ that satisfies the desired bound. This also holds true for b. □

PROOF OF THEOREM 3

It follows from Assumption 2 that $\widetilde{u}(t) = f_0^{(2m)}(t) \in L_2[0,1]$. Letting $\widetilde{x}_0 = \{f_0^{(m)}(0), \ldots, f_0^{(2m-2)}(0), f_0^{(2m-1)}(0)\}^{\mathrm{T}} \in \mathbb{R}^m$, we have $z(t) = f_0^{(m)}(t)$ such that $(\widetilde{u}, \widetilde{x}_0) \in \mathcal{P}$ for a large $\mu > 0$ and a small $\varepsilon > 0$. This shows that $\mathcal{P}$ is nonempty. It is also straightforward to verify that $\mathcal{P}$ is convex based on the strict convexity of the function $(\cdot)^{-1/(2m)}$ on $(0, \infty)$.

LEMMA 5. *The objective functional J is strictly convex over the set $\mathcal{P} \subseteq L_2[0,1] \times \mathbb{R}^m$.*

Proof. The lemma holds true even if the upper and lower bounds specified in Assumption 6 are removed. For each $(u, x_0) \in \mathcal{P}$, let ρ denote $z(t)/f_0^{(m)}(t)$ for $z(t)$ generated from (u, x_0), recalling that we require $\rho(t) > 0$ once $t \in \mathcal{N}$. Denote the set of such ρ's by $\mathcal{S}$. It is clear that ρ is strictly positive and $\rho^{-1/(2m)}$ is Lebesgue integrable on $[0,1]$. Noting that ρ shares the same convex combination relation with (u, x_0), the set $\mathcal{S}$ is convex and it suffices to show that $\Pi(\rho)$ is strictly convex on $\mathcal{S}$. Write Π in (14) as $\Pi(\rho) = \Pi_1(\rho) + \Pi_2(\rho)$, where

$$\Pi_1(\rho) = \int_0^1 r^2(t) \left[\{\rho(t) f_0^{(m)}(t)\}^{(m)}\right]^2 dt, \quad \Pi_2(\rho) = \int_0^1 L_0 r(t)^{1-1/(2m)}\, \rho(t)^{-1/(2m)} dt.$$

Obviously, Π_1 is convex on $\mathcal{S}$. We show next that $\Pi_2(\rho)$ is strictly convex on $\mathcal{S}$. Using the fact that the function $(\cdot)^{-1/(2m)}$ is strictly convex on $(0, \infty)$, we have, for any $\rho_1, \rho_2 \in \mathcal{S}$ with $\rho_1 \neq \rho_2$ and any $\alpha \in (0,1)$,

$$\{\alpha\rho_1(t) + (1-\alpha)\rho_2(t)\}^{-1/(2m)} < \alpha\{\rho_1(t)\}^{-1/(2m)} + (1-\alpha)\{\rho_2(t)\}^{-1/(2m)}, \ t \in [0,1].$$

Hence, $g(t) = \alpha\{\rho_1(t)\}^{-1/(2m)} + (1-\alpha)\{\rho_2(t)\}^{-1/(2m)} - \{\alpha\rho_1(t) + (1-\alpha)\rho_2(t)\}^{-1/(2m)}$ is strictly positive on $[0,1]$. Therefore, $L_0 r(t)^{1-1/(2m)} g(t) > 0$, $t \in [0,1]$. Since $\{\rho_1(t)\}^{-1/(2m)}$ and $\{\rho_2(t)\}^{-1/(2m)}$ are Lebesgue integrable, so is $g(t)$. Furthermore, it follows from Assumption 2 that $r(t)$ is strictly positive and continuous on $[0,1]$. Hence, there exists a positive constant κ such that $0 < r(t) \leq \kappa$, $t \in [0,1]$. In view of this, we deduce that $L_0 r(t)^{1-1/(2m)} g(t)$ is Lebesgue integrable. Thus it is easy to show that $\Pi_2(\alpha\rho_1 + (1-\alpha)\rho_2) < \alpha\Pi_2(\rho_1) + (1-\alpha)\Pi_2(\rho_2)$, leading to the strict convexity of Π_2. By this result, we further have, for any $\rho_1, \rho_2 \in \mathcal{S}$ with $\rho_1 \neq \rho_2$ and any $\alpha \in (0,1)$,

$$\Pi(\alpha\rho_1 + (1-\alpha)\rho_2) = \Pi_1(\alpha\rho_1 + (1-\alpha)\rho_2) + \Pi_2(\alpha\rho_1 + (1-\alpha)\rho_2),$$

where $\Pi_1(\alpha\rho_1 + (1-\alpha)\rho_2) \leq \alpha\Pi_1(\rho_1) + (1-\alpha)\Pi_1(\rho_2)$ and $\Pi_2(\alpha\rho_1 + (1-\alpha)\rho_2) < \alpha\Pi_2(\rho_1) + (1-\alpha)\Pi_2(\rho_2)$. Hence, $\Pi(\alpha\rho_1 + (1-\alpha)\rho_2) < \alpha\Pi(\rho_1) + (1-\alpha)\Pi(\rho_2)$. This shows that $\Pi(\rho)$ is strictly convex on $\mathcal{S}$, so is $J(u, x_0)$ on $\mathcal{P}$. □

Next, we use this lemma to derive the proof of Theorem 3.

Proof. The first part of the proof follows from a standard argument in functional analysis. Consider the Hilbert space $L_2[0,1] \times \mathbb{R}^m$ endowed with the standard inner product and norm. Since $\mathcal{P}$ is nonempty, pick an arbitrary $(\widetilde{u}, \widetilde{x}) \in \mathcal{P}$ and define the level set $\mathcal{L} = \{(u, x) \in \mathcal{P} : J(u, x) \leq J(\widetilde{u}, \widetilde{x})\}$. It follows from the condition $\|x_0\| \leq \mu$ and the structure of J that $\mathcal{L}$ is bounded. Further, due to the convexity of J, the set $\mathcal{L}$ is also convex. Since the space $L_2[0,1] \times \mathbb{R}^m$ is reflexive and self-dual, it follows from the Banach–Alaoglu Theorem that an arbitrary sequence $\{(u_n, x_n)\}$ in $\mathcal{L}$ has a subsequence $\{(u_n', x_n')\}$ that attains a weak*, thus weak, limit $(u^*, x^*) \in L_2[0,1] \times \mathbb{R}^m$.

In the next, we show that $(u^*, x^*) \in \mathcal{L}$. Let $z_{(u,x)}(t)$ denote $z(t)$ generated from (u, x). Since $u \in L_2[0,1]$, it belongs to $L_1[0,1]$. In view of $z_{(u,x)}(t)$ defined in (15), we see that $z_{(u,x)}(\cdot)$ is absolutely continuous on $[0,1]$ for any $(u,x) \in L_2[0,1] \times \mathbb{R}^m$ and $m \in \mathbb{N}$. Obviously, $\|x^*\| \leq \mu$. Furthermore, for any $t \in [0,1]$,

$$\left| \int_0^t u_n'(s)ds - \int_0^t u^*(s)ds \right| \leq \int_0^t |u_n'(s) - u^*(s)|ds \leq \int_0^1 |u_n'(s) - u^*(s)|ds \leq \|u_n' - u^*\|_{L_2},$$

where $\|\cdot\|_{L_2}$ denotes the L_2-norm on $L_2[0,1]$. Together with the convergence of x_n' to x^*, we deduce that $z_{(u_n',x_n')}(\cdot)$ converges to $z_{(u^*,x^*)}(\cdot)$ pointwise on $[0,1]$. Therefore, we have $z_{(u^*,x^*)}(t)/f_0^{(m)}(t) \geq \varepsilon$, $t \in [0,1]$. Moreover, a similar argument based on the property of r as in Lemma 5 shows that $r(t)^{1-1/(2m)}\{z_{(u_n',x_n')}(t)/f_0^{(m)}(t)\}^{-1/(2m)}$ is Lebesgue integrable for all n. In addition $0 \leq \{z_{(u_n',x_n')}(t)/f_0^{(m)}(t)\}^{-1/(2m)} \leq \varepsilon^{-1/(2m)}$, $t \in [0,1]$. This implies that for all n, $|r(t)^{1-1/(2m)}\{z_{(u_n',x_n')}(t)/f_0^{(m)}(t)\}^{-1/(2m)}| \leq r(t)^{1-1/(2m)}\varepsilon^{-1/(2m)}$ on $[0,1]$, where the latter function is clearly Lebesgue integrable since $r(\cdot)$ is continuous. By Lebesgue's Dominated Convergence Theorem, $r(t)^{1-1/(2m)}\{z_{(u^*,x^*)}(t)/f_0^{(m)}(t)\}^{-1/(2m)}$ is Lebesgue integrable on $[0,1]$ and as $n \to \infty$,

$$\int_0^1 r(t)^{1-1/(2m)} \left\{ \frac{z_{(u_n',x_n')}(t)}{f_0^{(m)}(t)} \right\}^{-1/(2m)} dt \longrightarrow \int_0^1 r(t)^{1-1/(2m)} \left\{ \frac{z_{(u^*,x^*)}(t)}{f_0^{(m)}(t)} \right\}^{-1/(2m)} dt. \tag{20}$$

This shows that $(u^*, x^*) \in \mathcal{P}$. Using the property of $r(t)$ deduced from Assumption 1 again, we obtain a positive constant κ' such that for any $\epsilon > 0$,

$$\begin{aligned}\int_0^1 r^2(t)\left|(u_n')^2(t) - (u^*)^2(t)\right|dt &\leq \kappa' \int_0^1 \left|(u_n')^2(t) - (u^*)^2(t)\right|dt \leq \kappa'\|u_n' - u^*\|_{L_2}\|u_n' + u^*\|_{L_2} \\ &\leq \kappa'\|u_n' - u^*\|_{L_2}(2\|u^*\|_{L_2} + \epsilon)\end{aligned}$$

for all n sufficiently large, where the Cauchy–Schwarz and triangle inequalities are used. Together with (20), we see that $J(u_n', x_n')$ converges to $J(u^*, x^*)$ as $n \to \infty$.

Consequently, we have $J(u^*, x^*) \leq J(\widetilde{u}, \widetilde{x})$ such that $(u^*, x^*) \in \mathcal{L}$. This shows that $\mathcal{L}$ is weakly compact. In view of the continuity of J, we see that a global optimal solution exists on $\mathcal{L}$ (Luenberger, 1969, Section 5.10, Theorem 2), and thus on $\mathcal{P}$. Finally, since J is strictly convex on the convex set $\mathcal{P}$, the optimal solution is unique. □

REFERENCES

ABRAMOVICH, F. & GRINSHTEIN, V. (1999). Derivation of an equivalent kernel for general spline smoothing: a systematic approach. *Bernoulli* **5**, 359-79.

BERLINET, A. & THOMAS-AGNAN, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Kluwer: Wiley.

CODDINGTON, E. A. & LEVINSON, N. (1955). *Theory of Ordinary Differential Equations*. New York: McGraw-Hill.

EGGERMONT, P. P. B. & LARICCIA, V. N.(2006). Uniform error bounds for smoothing splines. *IMS Lecture Notes-Monograph Series: High Dimensional Probability* **51**, 220-37.

LUENBERGER, D. G. (1969). *Optimization by Vector Space Methods*. John Wiley & Sons Inc.

WANG, X., SHEN, J. & RUPPERT, D.(2010). Local asymptotics of P-splines. *Electr. J. Statist.* **5**, 1-17.