1D Piecewise Smooth Function Estimation with Spline Functions

Daichi KITAHARA† Laurent CONDAT‡ Akira HIRABAYASHI†
† College of Information Science and Engineering, Ritsumeikan University
‡ GIPSA-Lab, Université Grenoble Alpes

Abstract Splines are piecewise polynomials and are widely used for interpolation and smoothing of observed data, due to their flexibility and optimality in the sense of variational problems on one-dimensional (1D) data. However, spline interpolation and smoothing are applicable only to the estimation of *continuous functions* and are not suitable for that of *piecewise smooth functions*. In this paper, we propose novel spline smoothing for 1D piecewise smooth function estimation. We define the set of splines *permitted to have discontinuous knots*. Then, we estimate a piecewise smooth function by a spline which minimizes the sum of the data fidelity, the roughness penalty, and the number of discontinuous knots.

1 INTRODUCTION

Let $f_j : (\zeta_{j-1}, \zeta_j) \to \mathbb{R}$ $(j = 1, 2, \ldots, m)$ be ρ-times continuously differentiable functions, i.e., $f_j \in C^\rho(\zeta_{j-1}, \zeta_j)$, where $-\infty \leq \zeta_0 < \zeta_1 < \cdots < \zeta_m \leq \infty$ and $\rho \in \mathbb{N} \cup \{\infty\}$. Define a *piecewise smooth function* $f : (\zeta_0, \zeta_m) \to \mathbb{R}$ by

$$f(x) := f_j(x) \qquad \text{for } x \in (\zeta_{j-1}, \zeta_j). \tag{1}$$

As the function values at ζ_j $(j = 1, 2, \ldots, m-1)$, although several cases such as $\lim_{x \to \zeta_j - 0} f_j(x)$, $\lim_{x \to \zeta_j + 0} f_{j+1}(x)$, and $\lim_{\substack{x \to \zeta_j - 0 \\ y \to \zeta_j + 0}} \frac{1}{2}(f_j(x) + f_{j+1}(y))$ can be considered, we do not take care about them in this paper. Alternatively, suppose the discontinuity of f at ζ_j $(j = 1, 2, \ldots, m-1)$, i.e., $\lim_{x \to \zeta_j - 0} f_j(x) \neq \lim_{x \to \zeta_j + 0} f_{j+1}(x)$ exactly holds. We observe noisy samples of the piecewise smooth function f by

$$z_i := f(x_i) + v_i \qquad (i = 1, 2, \ldots, n), \tag{2}$$

where $v_i \in \mathbb{R}$ is the additive white Gaussian noise, and sampling points satisfy $x_1 > \zeta_0$, $x_n < \zeta_m$, $\forall i \; x_{i+1} - x_i = h > 0$, $\forall i \, \forall j \; x_i \neq \zeta_j$ and $\forall j \, \exists i \; x_i \in (\zeta_{j-1}, \zeta_j)$. In this paper, we treat the estimation problem of the piecewise smooth function f in (1) from its finite noisy samples in (2). Such a problem appears in wide areas from science to engineering [1]–[10]. In [11], the authors assume that common basis functions of all pieces f_j $(j = 1, 2, \ldots, m)$ are known. On the other hand, in this paper, we assume that basis functions are unknown.

Spline is a function which is piecewise-defined by polynomials, and which can possess certain-times continuous differentiability including places where the polynomial pieces connect. Spline functions have been widely used for interpolation and smoothing of data in many signal and image processing areas [12], e.g., super-resolution [13], [14], computer aided design [15], [16], and regression analysis [17], [18]. The most commonly used spline functions are *cubic splines*, i.e., univariate spline functions which are expressed, on subintervals, as polynomials of degree 3 at most. This is because cubic splines are the unique solutions of the following variational problems on one-dimensional (1D) data [19]–[22].

Problem 1 (Variational Problem on 1D Interpolation) *Find* $g^* \in C^2(-\infty, \infty)$ *minimizing*

$$\int_{-\infty}^{\infty} |g''(x)|^2 \, \mathrm{d}x \tag{3}$$

subject to

$$g(x_i) = z_i \;\; \textit{for all } i = 1, 2, \ldots, n.$$

Problem 2 (Variational Problem on 1D Smoothing) *Find* $g^* \in C^2(-\infty, \infty)$ *minimizing*

$$\sum_{i=1}^{n} |g(x_i) - z_i|^2 + \lambda \int_{-\infty}^{\infty} |g''(x)|^2 \, \mathrm{d}x, \tag{4}$$

where the smoothing parameter $\lambda > 0$ *controls the trade-off between the data fidelity and the smoothness.*

Problem 1 is called *spline interpolation* and it is especially effective when noise-free data are available [13]–[16]. Problem 2 is called *spline smoothing* and it is often used for the design of continuous functions from noisy samples [17], [18]. However, spline interpolation and smoothing consider the estimation of *continuous functions*, and are not suitable for that of *piecewise smooth functions*.

In this paper, we propose novel spline smoothing for the estimation of piecewise smooth functions. For this purpose, in Section 2, as a preliminary, we newly define the set of spline functions which are *permitted to have several discontinuous knots*. The proposed piecewise smooth function estimation is given in Section 3. By assuming the sampling interval h is short enough compared with the length of each (ζ_{j-1}, ζ_j) $(j = 1, 2, \ldots, m)$, we can consider that the discontinuous knots should sparsely exist. We estimate f with a spline function minimizing the sum of the data fidelity term, the roughness penalty term, and a convex relaxation of the number of the discontinuous knots. Numerical experiments in Section 4 demonstrate the effectiveness of the proposed method by comparison with the standard spline smoothing.

2 PRELIMINARIES

2.1 Notation

Let $\mathbb{R}$ and $\mathbb{N}$ denote the set of all real numbers and nonnegative integers, respectively. For any open interval (a, b) and $\rho \in \mathbb{N} \cup \{\infty\}$, $C^\rho(a, b)$ stands for the set of all ρ-times continuously differentiable real-valued functions on (a, b). For any $d \in \mathbb{N}$, $\mathbb{P}_d$ stands for the set of all univariate real polynomials of degree d at most, i.e., $\mathbb{P}_d := \{p : \mathbb{R} \to \mathbb{R} : x \mapsto \sum_{k=0}^{d} c_k x^k \mid c_k \in \mathbb{R}\}$. We write a vector and a matrix with a boldface small letter and a boldface capital letter, respectively. For any vector $\boldsymbol{x} := (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$, the ℓ_2 norm of $\boldsymbol{x}$ is defined by $\|\boldsymbol{x}\|_2 := \sqrt{\sum_{i=1}^{n} |x_i|^2}$.

2.2 Spline Function Having Discontinuous Points

Let $\sqcup_n := \{I_i := (\xi_{i-1}, \xi_i)\}_{i=1}^n$ be a set of subintervals I_i on an open interval $I := (\xi_0, \xi_n)$ s.t. $\xi_i - \xi_{i-1} = h_i > 0$ $(i = 1, 2, \ldots, n)$. For $\sqcup_n$ and $\rho, d \in \mathbb{N}$ s.t. $0 \le \rho < d$, define

$$\mathcal{S}_d^\rho(\sqcup_n) := \left\{ s : (\xi_0, \xi_n) \to \mathbb{R} \;\middle|\; \begin{array}{l} s = p_i \in \mathbb{P}_d \text{ on } I_i, \\ s = \frac{1}{2}(p_i + p_{i+1}) \text{ at } \xi_i, \\ \text{and } s \in C^0(\xi_{i-1}, \xi_{i+1}) \\ \quad \Rightarrow s \in C^\rho(\xi_{i-1}, \xi_{i+1}) \end{array} \right\} \tag{5}$$

as the set of all univariate spline functions, *permitted to have discontinuous knots* ξ_i, of degree d and smoothness ρ on $\sqcup_n$. In what follows, we express a spline function $s \in \mathcal{S}_d^\rho(\sqcup_n)$ in the following interval normalization form:

$$s(x) := p_i(x) := \sum_{k=0}^{d} c_k^{\langle i \rangle} \left(\frac{x - \xi_{i-1}}{h_i} \right)^k \quad \text{for } x \in (\xi_{i-1}, \xi_i), \tag{6}$$

where $c_k^{\langle i \rangle} \in \mathbb{R}$ $(k = 0, 1, \ldots, d)$ are coefficients of $p_i \in \mathbb{P}_d$.

2.2.1 Quadratic Form of the Roughness Penalty Term

By restricting the domain of interest to I and the function space to $\mathcal{S}_d^\rho(\sqcup_n)$, the roughness penalty term used in (3) and (4) is expressed as

$$\int_I |s''(x)|^2 \, \mathrm{d}x = \sum_{i=1}^{n} \int_{I_i} |s''(x)|^2 \, \mathrm{d}x. \tag{7}$$

By using the expression in (6), the roughness penalty on I_i is expressed as the following quadratic form:

$$\begin{aligned} &\int_{I_i} |s''(x)|^2 \, \mathrm{d}x \\ &= \sum_{k=2}^{d} \sum_{l=2}^{d} \frac{k(k-1)l(l-1) c_k^{\langle i \rangle} c_l^{\langle i \rangle}}{h_i^4} \int_{I_i} \left(\frac{x - \xi_{i-1}}{h_i} \right)^{k+l-4} \mathrm{d}x \\ &= \sum_{k=2}^{d} \sum_{l=2}^{d} \frac{k(k-1)l(l-1)}{h_i^3 (k+l-3)} c_k^{\langle i \rangle} c_l^{\langle i \rangle} \\ &= \sum_{k=0}^{d-2} \sum_{l=0}^{d-2} \frac{(d-k)(d-k-1)(d-l)(d-l-1)}{h_i^3 (2d-k-l-3)} c_{d-k}^{\langle i \rangle} c_{d-l}^{\langle i \rangle} \\ &= \boldsymbol{c}_i^{\mathrm{T}} \boldsymbol{Q}_i \boldsymbol{c}_i, \end{aligned} \tag{8}$$

where $\boldsymbol{c}_i := (c_d^{\langle i \rangle}, c_{d-1}^{\langle i \rangle}, \ldots, c_0^{\langle i \rangle})^{\mathrm{T}} \in \mathbb{R}^{d+1}$ and a symmetric positive semidefinite matrix $\boldsymbol{Q}_i \in \mathbb{R}^{(d+1)\times(d+1)}$ is defined as

$$[\boldsymbol{Q}_i]_{k+1,l+1} := \frac{(d-k)(d-k-1)(d-l)(d-l-1)}{h_i^3 (2d-k-l-3)} \quad (k = 0, 1, \ldots, d-2 \text{ and } l = 0, 1, \ldots, d-2)$$

and $[\boldsymbol{Q}_i]_{k+1,l+1} := 0$ $(k = d-1, d$ or $l = d-1, d)$. From (7) and (8), the roughness penalty on I can be expressed as

$$\int_I |s''(x)|^2 \, \mathrm{d}x = \boldsymbol{c}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{c}, \tag{9}$$

where $\boldsymbol{c} := (\boldsymbol{c}_1^{\mathrm{T}}, \boldsymbol{c}_2^{\mathrm{T}}, \ldots, \boldsymbol{c}_n^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{n(d+1)}$ is the coefficient vector of $s \in \mathcal{S}_d^\rho(\sqcup_n)$ and $\boldsymbol{Q} \in \mathbb{R}^{n(d+1)\times n(d+1)}$ is a symmetric positive semidefinite matrix based on $\boldsymbol{Q}_i$ $(i = 1, 2, \ldots, n)$.

2.2.2 Linear Equation for the ρ-Times Differentiability

For a spline function $s \in \mathcal{S}_d^\rho(\sqcup_n)$ in (5), to ensure the ρ-times continuous differentiability on (ξ_{i-1}, ξ_{i+1}), i.e., $s \in C^\rho(\xi_{i-1}, \xi_{i+1})$, the coefficients of the adjacent polynomials p_i and p_{i+1} in (6) have to satisfy the following equations:

$$\begin{aligned} & s \in C^\rho(\xi_{i-1}, \xi_{i+1}) \\ & \Leftrightarrow p_i^{(l)}(\xi_i) = p_{i+1}^{(l)}(\xi_i) \quad (l = 0, 1, \ldots, \rho) \\ & \Leftrightarrow \frac{1}{h_i^l} \sum_{k=l}^{d} \frac{k!}{(k-l)!} c_k^{\langle i \rangle} = \frac{l!}{h_{i+1}^l} c_l^{\langle i+1 \rangle} \quad (l = 0, 1, \ldots, \rho) \\ & \Leftrightarrow \frac{1}{h_i^l} \sum_{k=l}^{d} \frac{k!}{(k-l)!} c_k^{\langle i \rangle} - \frac{l!}{h_{i+1}^l} c_l^{\langle i+1 \rangle} = 0 \quad (l = 0, 1, \ldots, \rho). \end{aligned} \tag{10}$$

From (10), there is a matrix $\boldsymbol{H}_i \in \mathbb{R}^{(\rho+1)\times 2(d+1)}$ satisfying

$$s \in C^\rho(\xi_{i-1}, \xi_{i+1}) \quad \Leftrightarrow \quad \boldsymbol{H}_i \begin{bmatrix} \boldsymbol{c}_i \\ \boldsymbol{c}_{i+1} \end{bmatrix} = \boldsymbol{0}. \tag{11}$$

In this paper, to remove the ambiguity of $\boldsymbol{H}_i$ on (i) constant multiplication and (ii) the order of the row vectors, assume that each matrix $\boldsymbol{H}_i$ $(i = 1, 2, \ldots, n-1)$ satisfies

$$\boldsymbol{H}_i \begin{bmatrix} \boldsymbol{c}_i \\ \boldsymbol{c}_{i+1} \end{bmatrix} = \begin{bmatrix} p_i(\xi_i) - p_{i+1}(\xi_i) \\ p_i'(\xi_i) - p_{i+1}'(\xi_i) \\ \vdots \\ p_i^{(\rho)}(\xi_i) - p_{i+1}^{(\rho)}(\xi_i) \end{bmatrix}. \tag{12}$$

3 Piecewise Smooth Function Estimation Based on Spline Smoothing

In this section, we estimate the piecewise smooth function f in (1) with the use of a spline function $s \in \mathcal{S}_d^\rho(\sqcup_n)$. We define $\xi_0 := x_1 - h/2$ and $\xi_i := x_i + h/2$ $(i = 1, 2, \ldots, n)$. Then from (6), the function values of s at sampling points $x_i \in (\xi_{i-1}, \xi_i)$ $(i = 1, 2, \ldots, n)$ are given by

$$s(x_i) = \sum_{k=0}^{d} c_k^{\langle i \rangle} \frac{1}{2^k} = \begin{bmatrix} \frac{1}{2^d} & \frac{1}{2^{d-1}} & \cdots & 1 \end{bmatrix} \boldsymbol{c}_i =: \boldsymbol{a}^{\mathrm{T}} \boldsymbol{c}_i. \tag{13}$$

Therefore, the data fidelity term used in (4) is expressed as

$$\sum_{i=1}^{n} |s(x_i) - z_i|^2 = \|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{z}\|_2^2, \tag{14}$$

where $\boldsymbol{z} := (z_1, z_2, \ldots, z_n)^{\mathrm{T}} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{n\times n(d+1)}$ is a matrix whose row vectors are based on $\boldsymbol{a}^{\mathrm{T}}$ in (13).

Suppose that the sampling interval h is short enough compared with the length of each (ζ_{j-1}, ζ_j) $(j = 1, 2, \ldots, m)$, i.e., there are enough samples to accurately reconstruct each f_j if the discontinuous points ζ_j $(j = 1, 2, \ldots, m-1)$ can be detected. Then, the discontinuous knots ξ_i of $s \in \mathcal{S}_d^\rho(\sqcup_n)$, approximating f, should sparsely exist on (ξ_0, ξ_n). Moreover, even if $\lim_{\substack{x\to\zeta_j-0 \\ y\to\zeta_j+0}} |f_j^{(l)}(x) - f_{j+1}^{(l)}(y)| \not\approx 0$ $(l = 0, 1, 2)$ and $x_i < \zeta_j < x_{i+1}$ hold for some i and j, the roughness penalty $\int_{\xi_{i-1}}^{\xi_{i+1}} |s''(x)|^2 \, \mathrm{d}x$ can become small for $s \notin C^0(\xi_{i-1}, \xi_{i+1})$.

On the basis of the above discussion, we consider the following non-convex optimization problem:

$$\underset{s\in\mathcal{S}_d^\rho(\sqcup_n)}{\text{minimize}} \sum_{i=1}^{n} |s(x_i)-z_i|^2 + \lambda \int_I |s''(x)|^2 \,\mathrm{d}x + w\,\|s\|_0, \tag{15}$$

where $\lambda > 0$, $w > 0$, and $\|s\|_0 \in \mathbb{N}$ denotes the number of discontinuous knots ξ_i s.t. $p_i(\xi_i) \neq p_{i+1}(\xi_i)$. From (9), (11), and (14), the problem in (15) is expressed as an optimization problem on the coefficient vector $\boldsymbol{c} \in \mathbb{R}^{n(d+1)}$:

$$\underset{\boldsymbol{c}\in\mathbb{R}^{n(d+1)}}{\text{minimize}} \|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{z}\|_2^2 + \lambda \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c} + w\sum_{i=1}^{n-1} \Gamma\left(\boldsymbol{H}_i \begin{bmatrix}\boldsymbol{c}_i\\ \boldsymbol{c}_{i+1}\end{bmatrix}\right), \tag{16}$$

where $\Gamma : \mathbb{R}^{\rho+1} \to \{0,1\}$ is defined by $\Gamma(\boldsymbol{x}) = 0$ if $\boldsymbol{x} = \boldsymbol{0}$, and $\Gamma(\boldsymbol{x}) = 1$ otherwise. In order to approximately solve the problem in (16), we use a convex relaxation.

The third term of the cost function in (16) can be considered as a group ℓ_0 (pseudo) norm without overlapping:

$$\|\boldsymbol{H}\boldsymbol{c}\|_0^G := \sum_{i=1}^{n-1} \Gamma\left(\boldsymbol{H}_i \begin{bmatrix}\boldsymbol{c}_i\\ \boldsymbol{c}_{i+1}\end{bmatrix}\right) \tag{17}$$

with the use of some matrix $\boldsymbol{H} \in \mathbb{R}^{(n-1)(\rho+1)\times n(d+1)}$ based on $\boldsymbol{H}_i$ $(i = 1,2,\ldots,n-1)$. Therefore, we replace the group ℓ_0 norm in (17) with a weighted group ℓ_1 norm:

$$\begin{aligned}\|\boldsymbol{H}\boldsymbol{c}\|_{1,\boldsymbol{w}}^G &:= \sum_{i=1}^{n-1} \left\|\boldsymbol{H}_i \begin{bmatrix}\boldsymbol{c}_i\\ \boldsymbol{c}_{i+1}\end{bmatrix}\right\|_{2,\boldsymbol{w}_i}\\ &:= \sum_{i=1}^{n-1} \sqrt{\sum_{l=0}^{\rho} w_l^{\langle i\rangle} \left|p_i^{(l)}(\xi_i) - p_{i+1}^{(l)}(\xi_i)\right|^2},\end{aligned} \tag{18}$$

where $w_l^{\langle i\rangle} > 0$, $\boldsymbol{w}_i := (w_0^{\langle i\rangle}, w_1^{\langle i\rangle}, \ldots, w_\rho^{\langle i\rangle})^{\mathrm{T}} \in \mathbb{R}^{\rho+1}$, and $\boldsymbol{w} := (\boldsymbol{w}_1^{\mathrm{T}}, \boldsymbol{w}_2^{\mathrm{T}}, \ldots, \boldsymbol{w}_{n-1}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{(n-1)(\rho+1)}$. In (18), we use the definitions of the matrices $\boldsymbol{H}_i$ in (12). As a result, we propose to solve the following convex optimization problem:

$$\underset{\boldsymbol{c}\in\mathbb{R}^{n(d+1)}}{\text{minimize}} \|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{z}\|_2^2 + \lambda \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c} + \|\boldsymbol{H}\boldsymbol{c}\|_{1,\boldsymbol{w}}^G, \tag{19}$$

for the estimation of the piecewise smooth function f. The optimal solution of the problem in (19) is computed by the alternating direction method of multipliers (ADMM) [23]:

$$\left\lfloor\begin{aligned}\boldsymbol{c}_{t+1} &= \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \lambda\boldsymbol{Q} + \frac{1}{\gamma}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{W}^2\boldsymbol{H}\right)^{-1}\\ &\qquad \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z} + \frac{1}{\gamma}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{\mu}_t - \boldsymbol{\nu}_t)\right)\\ \boldsymbol{\mu}_{t+1} &= \mathrm{prox}_{\frac{\gamma\lambda}{2}\|\cdot\|_1^G}(\boldsymbol{W}\boldsymbol{H}\boldsymbol{c}_{t+1} + \boldsymbol{\nu}_t)\\ \boldsymbol{\nu}_{t+1} &= \boldsymbol{\nu}_t + \boldsymbol{W}\boldsymbol{H}\boldsymbol{c}_{t+1} - \boldsymbol{\mu}_{t+1}\end{aligned}\right.$$

with $\gamma > 0$ and any initialization $(\boldsymbol{\mu}_0, \boldsymbol{\nu}_0) \in \mathbb{R}^{(n-1)(\rho+1)} \times \mathbb{R}^{(n-1)(\rho+1)}$, where $\boldsymbol{W} \in \mathbb{R}^{(n-1)(\rho+1)\times(n-1)(\rho+1)}$ is a diagonal matrix whose components are the square roots of $w_l^{\langle i\rangle}$.

Finally, in order to obtain a spline function $s \in \mathcal{S}_d^\rho(\sqcup_n)$ as an estimate of f, we re-solve the problem in (15). From the optimal coefficient vector $\boldsymbol{c}^*$ of the problem in (19), detect the continuous knots ξ_i of $s \in \mathcal{S}_d^\rho(\sqcup_n)$ by checking whether $|p_i^*(\xi_i) - p_{i+1}^*(\xi_i)|$ is lower than a threshold value $\tau > 0$. In (5), if ξ_i is the continuous knot, then $s \in C^\rho(\xi_{i-1}, \xi_{i+1})$ must hold. Therefore, to obtain the solution of (15), we solve

$$\begin{aligned}&\underset{\boldsymbol{c}\in\mathbb{R}^{n(d+1)}}{\text{minimize}} \|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{z}\|_2^2 + \lambda \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c}\\ &\text{subject to } \boldsymbol{H}_i \begin{bmatrix}\boldsymbol{c}_i\\ \boldsymbol{c}_{i+1}\end{bmatrix} = \boldsymbol{0} \text{ for all continuous knots } \xi_i.\end{aligned} \tag{20}$$

The problem in (20) is solved by quadratic programming.

4 Numerical Experiments

Define $\zeta_0 := 0$, $\zeta_1 := 20$, $\zeta_2 := 50$, $\zeta_3 := 70$, $\zeta_4 := 95$, $\zeta_5 := 100$, and $x_i := i - 0.5$ $(i = 1,2,\ldots,n := 100)$. As a result, knots of a spline function $s \in \mathcal{S}_3^2(\sqcup_n)$ are defined as $\xi_i := i$ $(i = 0,1,\ldots,100)$. For two piecewise smooth functions depicted by yellow lines in Figs. 1 and 2, we try to reconstruct them from their noisy samples in (2), where the standard division of the additive white Gaussian noise v_i is $\sigma = 5$. We compare the estimation results by the standard spline smoothing (see Problem 2) and the proposed methods (problem in (19) followed by problem in (20)). The smoothing parameter is set to $\lambda = 65$. The weights in the weighted group ℓ_1 norm are set to $w_0^{\langle i\rangle} = \frac{360000}{(z_{i+1}-z_i)^2+1}$, $w_1^{\langle i\rangle} = 3600$, and $w_2^{\langle i\rangle} = 0.36$. In Figs. 1 and 2, black circles denote the observed noisy samples. Blue, red, and green lines depict the estimation results by spline smoothing, the proposed method in (19), and the proposed method in (20), respectively.

From Figs. 1 and 2, we can see that the estimation results by spline smoothing lose the edges of the original piecewise smooth functions because the standard spline smoothing cannot express discontinuous points. On the other hand, the estimation results by the proposed methods are very good. In particular, we can see that the proposed method in (19) achieves both the detection of the discontinuous knots and the smoothing around the continuous knots.

5 Conclusion

In this paper, we have proposed spline smoothing for the estimation of piecewise smooth functions. For this purpose, we defined the set of spline functions permitted to have discontinuous knots. We estimated a piecewise smooth function with a spline function which minimizes the sum of the data fidelity, the roughness penalty, and the convex relaxation of the number of the discontinuous knots. The minimizer can be effectively computed by ADMM, and the numerical experiments showed the effectiveness of the proposed method.

Appendix Alternating Direction Method of Multipliers

The alternating direction method of multipliers (ADMM) solves the following convex optimization problem [23]:

$$\text{Find } \boldsymbol{x}^* \in \underset{\boldsymbol{x}\in\mathbb{R}^n}{\operatorname{argmin}}\ f(\boldsymbol{x}) + g(\boldsymbol{L}\boldsymbol{x}),$$

where $\boldsymbol{L} \in \mathbb{R}^{m\times n}$ and two functions $f : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ and $g : \mathbb{R}^m \to \mathbb{R}\cup\{\infty\}$ are *proper, lower semicontinuous*

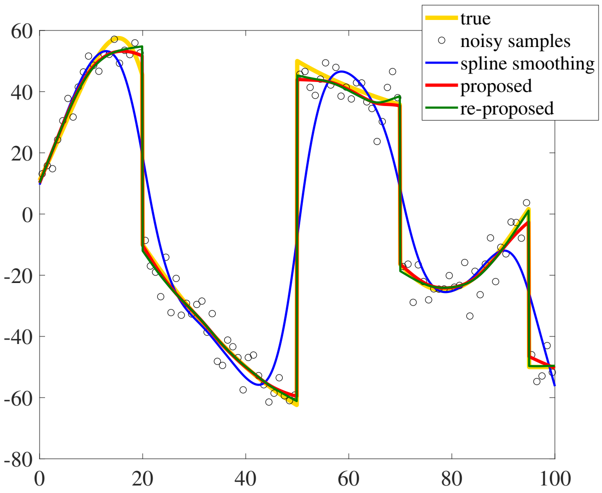

Figure 1: Experimental Results I

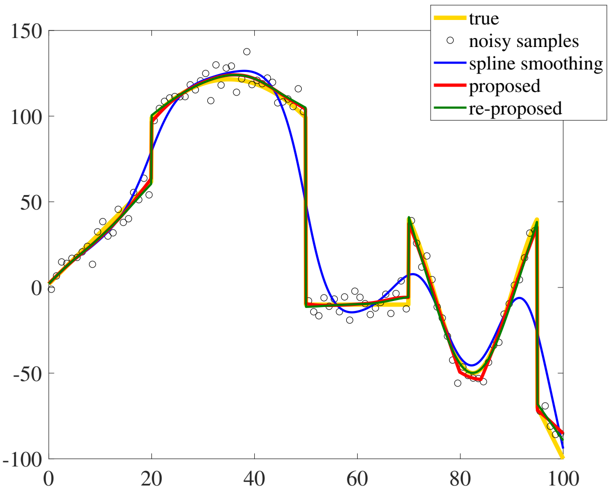

Figure 2: Experimental Results II

and *convex.*[1] The ADMM iteratively computes

$$\left\lfloor \begin{aligned} \boldsymbol{x}_{t+1} &= \operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) + \frac{1}{2\gamma}\|\boldsymbol{\mu}_t - \boldsymbol{L}\boldsymbol{x} - \boldsymbol{\nu}_t\|_2^2 \\ \boldsymbol{\mu}_{t+1} &= \operatorname{prox}_{\gamma g}(\boldsymbol{L}\boldsymbol{x}_{t+1} + \boldsymbol{\nu}_t) \\ \boldsymbol{\nu}_{t+1} &= \boldsymbol{\nu}_t + \boldsymbol{L}\boldsymbol{x}_{t+1} - \boldsymbol{\mu}_{t+1} \end{aligned} \right. \tag{21}$$

with $\gamma > 0$ and any initialization $(\boldsymbol{\mu}_0, \boldsymbol{\nu}_0) \in \mathbb{R}^m \times \mathbb{R}^m$, where $\operatorname{prox}_{\gamma g} : \mathbb{R}^m \to \mathbb{R}^m$ denotes the proximity operator[2] of γg. Then $(\boldsymbol{x}_t)_{t=1}^{\infty}$ converges to the optimal solution $\boldsymbol{x}^*$.

ACKNOWLEDGMENT

This work was supported in part by JSPS Grants-in-Aid for Research Activity Start-Up (Grant Number 17H07243).

REFERENCES

[1] X. Liu, N. D. Sidiropoulos, and A. Swami, "Blind high-resolution localization and tracking of multiple frequency hopped signals," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 889–901, Apr. 2002.

[2] A. Valyrakis, E. E. Tsakonas, N. D. Sidiropoulos, and A. Swami, "Stochastic modeling and particle filtering algorithms for tracking a frequency-hopped signal," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3108–3118, Aug. 2009.

[3] D. Angelosante, G. B. Giannakis, and N. D. Sidiropoulos, "Estimating multiple frequency-hopping signal parameters via sparse linear regression," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5044–5056, Oct. 2010.

[4] D. Angelosante, G. B. Giannakis, and N. D. Sidiropoulos, "Sparse parametric models for robust nonstationary signal analysis: Leveraging the power of sparse regression," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 64–73, Nov. 2013.

[5] R. Giryes, M. Elad, and A. M. Bruckstein, "Sparsity based methods for overparameterized variational problems," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 2133–2159, 2015.

[6] S. J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "ℓ_1 trend filtering," *SIAM Review*, vol. 51, no. 2, pp. 339–360, 2009.

[7] G. Steidl, S. Didas, and J. Neumann, "Splines in higher order TV regularization," *International Journal of Computer Vision*, vol. 70, no. 3, pp. 241–255, 2006.

[8] R. J. Tibshirani, "Adaptive piecewise polynomial estimation via trend filtering," *The Annals of Statistics*, vol. 42, no. 1, pp. 285–323, Feb. 2014.

[9] G. Ongie and M. Jacob, "Recovery of discontinuous signals using group sparse higher degree total variation," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1414–1418, Sep. 2015.

[10] I. W. Selesnick, "Generalized total variation: Tying the knots," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 2009–2013, Nov. 2015.

[11] H. Kuroda, M. Yamagishi, and I. Yamada, "Recovery of piecewise finite-dimensional continuous signals by exploiting sparsity," in *Proceedings of International Conference on Sampling Theory and Applications (SampTA)*, 2017, pp. 609–613.

[12] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[13] H. S. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 6, pp. 508–517, Dec. 1978.

[14] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[15] E. Cohen, T. Lyche, and R. Riesenfeld, "Discrete B-splines and subdivision techniques in computer-aided geometric design and computer graphics," *Computer Graphics and Image Processing*, vol. 14, no. 2, pp. 87–111, Oct. 1980.

[16] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Code*, 4th ed. New York, NY, USA: Academic Press, 1996.

[17] S. Wold, "Spline functions in data analysis," *Technometrics*, vol. 16, no. 1, pp. 87–111, Feb. 1974.

[18] L. C. Marsh and D. R. Cormier, *Spline Regression Models*, ser. Quantitative Applications in the Social Sciences. Thousand Oaks, CA, USA: SAGE, 2001, vol. 137.

[19] C. de Boor, "Best approximation properties of spline functions of odd degree," *Journal of Mathematics and Mechanics*, vol. 12, no. 5, pp. 747–749, 1963.

[20] B. W. Silverman, "Some aspects of the spline smoothing approach to non-parametric regression curve fitting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 47, no. 1, pp. 1–52, 1985.

[21] G. Wahba, *Spline Models for Observational Data*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA, USA: SIAM, 1990, vol. 59.

[22] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, 2nd ed. New York, NY, USA: Springer, 2005.

[23] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[1] A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called proper, lower semicontinous, and convex if $\operatorname{dom}(f) := \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) < \infty\} \neq \varnothing$, $\operatorname{lev}_{\leq\alpha}(f) := \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$, and $f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in (0, 1)$, respectively.

[2] The proximity operator of a proper, lower semicontinous, convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is given by

$$\operatorname{prox}_f(\boldsymbol{x}) := \operatorname*{argmin}_{\boldsymbol{y}\in\mathbb{R}^n} f(\boldsymbol{y}) + \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2 .$$