

Reconstruction of Smooth Functions from Incomplete Data

Peter Massopust[†]

*GSF-National Research Center for Environment and Health,
Institute of Biomathematics and Biometry,
and
Centre of Mathematics M6,
Technische Universität München, Germany
Email: massopust@ma.tum.de*

(Received March 28, 2007)

A method for the reconstruction of a smooth function from incomplete data by means of Schoenberg's variation diminution and Hermite interpolation is presented. Error bounds on the reconstruction algorithm are derived and numerical experiments evaluated.

AMS Subject Classification: 41A15, 65D07, 65D10.

1 Introduction

In many situations, one is confronted with the problem of trying to reconstruct an unknown continuous function from a finite set of data points or measurements where some of this data is incorrect, missing, or unreliable. If a contiguous set of data points is missing, the problem of reconstruction is highly ill-posed with an infinite dimensional solution set. However, if more information about the function is available, such as smoothness, monotonicity behavior, or convexity, then it might be possible to reconstruct the function to a certain degree from a set of incomplete or incorrect data.

There exist, of course, stochastic approaches (see for instance [1, 2]), but sometimes a deterministic algorithm may be preferential. In this short paper, such a deterministic method is presented to reconstruct a smooth function f when the values of f are unknown over an interval of positive Lebesgue measure. (This also includes the case when the function is unknown over a finite number of intervals of positive Lebesgue measure.) Of course, some assumptions about the function f must be made in order for this process to work. These assumptions are not too stringent; the main issue being the existence of exactly one critical value on the interval where the information is missing, and for computational simplicity symmetry with respect to this critical value. The proposed method uses B-splines approximation techniques, in particular, the so-called Schoenberg variation diminution, as this procedure is ideally suited for preserving non-negativity, monotonic-

[†] Research was partially supported by the grant MEXT-CT-2004-013477, Acronym MAME-BIA, of the European Commission.

ity, and convexity properties of data. In addition, the variation diminishing splines are used to predict the derivative of the unknown values of f via Hermite interpolation.

As an application for this type of setting, we like to mention the measurement of induced static electric and magnetic fields generated by small cavities in certain materials. The amplitude of the static electric and magnetic fields due to these small cavities can be extremely large when the applied fields are high. Due to finite bit allocation, measurement devices sometimes fail to register the large amplitude values and clip the measurements. These clipped data values fall into the above description of incomplete data and the electric and magnetic responses satisfy, in very good approximation, the assumptions that were made above.

The structure of this paper is as follows. In Section 2, a brief review of Spline Theory is presented. Particular emphasis is placed on B-splines, the Schoenberg variation diminution, and Hermite interpolatory splines. In addition, the for the purposes of this paper relevant approximation-theoretic results are also given. The reconstruction algorithm for smooth functions is presented in Section 3 together with the derivation of error estimates for this reconstruction method. The last section, presents some numerical experiments and validation of the reconstruction algorithm.

2 Brief Review of Spline Theory

In this section, we present some definitions and results from the theory of splines, in particular, B-Splines and Hermite interpolatory splines. For more details and for proofs of the stated theorems, the interested reader is referred to [4, 5].

In the sequel, we use the following notation. The linear space of n -times continuously differentiable functions on an interval $[a, b]$, $a < b$, is denoted by $C^n[a, b]$, for $n \in \mathbb{N}$, and by $C[a, b]$ in the case $n = 0$, where it is called the space of continuous functions on $[a, b]$. The linear space of real polynomials of degree $< k$ or order k , $k \in \mathbb{N}$, will be written as Π^k .

2.1 B-Splines

A spline function or, for short, a spline is a piecewise polynomial function joined together on subintervals with certain continuity or smoothness conditions. More precisely, let $X := \{a := x_0 < x_1 < \dots < x_{n+1} := b\}$ be a sequence of real numbers and let $k \in \mathbb{N}$. A spline of degree $k - 1$ or order k on the interval $[a, b]$ is a function s such that

- (i) $s|_{[x_{i-1}, x_i]} \in \Pi^k$, $i = 1, \dots, n + 1$;
- (ii) $s \in C^{k-2}[a, b]$.

We refer to the set X also as the knot sequence for s . It can be shown [5] that the set $\mathcal{S}_{X,k}$ of all spline functions s of order k and with knot sequence X forms a real vector space of dimension $n + k$.

B-Splines are local basis functions for $\mathcal{S}_{X,k}$ in the sense that they are nonnegative only on k contiguous intervals $[x_i, x_{i+1}]$ and vanish elsewhere. Moreover, B-splines provide a computationally efficient and numerically stable framework for the evaluation and approximation by splines.

Let $t_1 \leq t_2 \leq \dots \leq t_{n+k}$ be a nondecreasing sequence of real numbers. The quantity $\mathbf{t} := (t_1, t_2, \dots, t_{n+k})$ is referred to as a knot vector. Note that some of the knots t_i may be repeated, i.e., it may happen that $t_i = t_{i+1} = \dots = t_{i+\nu_i}$, for some $\nu_i > 0$. The number ν_i is called the multiplicity of knot t_i , and if $\nu_i = 1$, then the knot t_i is called simple.

Now suppose that positive integers n and k are given. The B-spline $B_{i,k,\mathbf{t}}$ of order k with knot vector \mathbf{t} and corresponding to knot t_i is recursively defined as follows.

$$B_{i,1,\mathbf{t}}(x) := \begin{cases} 1, & \text{for } t_i \leq x < t_{i+1} \\ 0, & \text{otherwise,} \end{cases}$$

and for $1 < k \in \mathbb{N}$

$$B_{i,k,\mathbf{t}}(x) := \frac{x - t_i}{t_{i+k-1} - t_i} B_{i,k-1,\mathbf{t}}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} B_{i+1,k-1,\mathbf{t}}(x), \quad (2.1)$$

for $i = 1, \dots, n$. Should it happen that $t_{i+k-1} = t_i$ or $t_{i+k} = t_{i+1}$ for one of the coefficients in the above recursive definition of $B_{i,k,\mathbf{t}}$, then the entire coefficient is set equal to zero.

Next, we list some of the main properties of B-splines.

- (i) $B_{i,k,\mathbf{t}}(x) > 0$ for $t_i < x < t_{i+k}$;
- (ii) $B_{i,k,\mathbf{t}}$ is a polynomial of order k on each interval $[t_i, t_{i+1})$ and a piecewise polynomial on each interval $[t_i, t_{i+k})$.
- (iii) $B_{i,k,\mathbf{t}} \equiv 0$ on $(-\infty, t_1) \cup (t_{n+k}, +\infty)$.
- (iv) If ν_i is the multiplicity of knot t_i , i.e., $t_{i-1} \neq t_i = t_{i+1} = \dots = t_{i+\nu_i-1} \neq t_{i+\nu_i}$, then $B_{i,k,\mathbf{t}}$ is $(k-1-\nu_i)$ -times continuously differentiable in a neighborhood of t_i .
Loosely speaking,

degree = smoothness at knot + multiplicity of knot.

- (v) Partition of Unity Property: $\sum_{i=1}^n B_{i,k,\mathbf{t}}(x) \equiv 1$ on $[t_k, t_{n+1}]$.

Now consider the spline space $\mathcal{S}_{X,k}$ whose elements are splines of order k . The elements of $\mathcal{S}_{X,k}$ have prescribed degrees of smoothness $k-\nu_i$ at each of the knots x_i , $i = 1, \dots, n$, whereas the smoothness of B-splines is controlled by the multiplicity of knots coinciding with x_i . This observation shows that for any space $\mathcal{S}_{X,k}$ we can find a knot vector $\mathbf{t} = (t_j)_{1 \leq j \leq m+k}$ whose associated B-splines form a basis for $\mathcal{S}_{X,k}$. To this end, let $\mathbf{t} = (t_j)_{1 \leq j \leq m+k}$ be defined by

$$\begin{aligned} t_1 = \dots = t_k &=: x_0; & t_{k+\sum_{\ell=1}^j \nu_{\ell+i}} &:= x_j, & i &= 0, 1, \dots, \nu_j - 1; & j &= 1, \dots, m; \\ t_{k+\sum_{\ell=1}^n \nu_{\ell+i}} &:= x_{n+1}, & i &= 1, \dots, k; & m &= k + \sum_{j=1}^n \nu_j. \end{aligned} \quad (2.2)$$

Then the functions $B_{j,k,\mathbf{t}}$, $j = 1, \dots, m$, are linearly independent and elements of $\mathcal{S}_{X,k}$. Thus, they form a basis of $\mathcal{S}_{X,k}$.

Theorem 2.1 (Representation Theorem for Splines) *Every spline $s \in \mathcal{S}_{X,k}$ defined on*

the interval $[a, b]$ has a unique expansion in terms of B-splines of order k of the form:

$$s = \sum_{j=1}^m c_j B_{j,k,\mathbf{t}} \quad c_j \in \mathbb{R}.$$

For a fixed $k \in \mathbb{N}$ and a fixed knot vector \mathbf{t} , denote by

$$\mathfrak{S}_{k,\mathbf{t}} := \left\{ \sum_{j=1}^m c_j B_{j,k,\mathbf{t}} \mid c_j \in \mathbb{R} \right\}$$

the vector space spanned by the $B_{j,k,\mathbf{t}}$, $j = 1, \dots, m$. Then the above theorem states that

$$\mathcal{S}_{X,k} = \mathfrak{S}_{k,\mathbf{t}}.$$

In the following, we require derivatives of B-spline representations. Employing again the recursive definition of B-splines, derivatives can be easily computed.

Theorem 2.2 (Differentiation of B-Splines) *Let $s = \sum_{j=1}^m c_j B_{j,k,\mathbf{t}}$ be the B-spline representation of s . Then, for all $\nu \in \mathbb{N}$,*

$$\frac{d}{dx^\nu} \left(\sum_{j=1}^m c_j B_{j,k,\mathbf{t}}(x) \right) = \sum_{j=1}^{m+\nu} c_j^{[\nu+1]} B_{j,k-\nu,\mathbf{t}}(x),$$

with

$$c_r^{[\nu+1]} := \begin{cases} c_r, & \nu = 0, \\ \frac{c_r^{[\nu]} - c_{r-1}^{[\nu]}}{(t_{r+k-\nu} - t_r)/(k-\nu)}, & \nu > 0. \end{cases} \quad (2.3)$$

Here, the coefficients $c_r^{[\nu]}$ whose subscripts are outside the range $1, \dots, m$ are set equal to zero.

Remark 2.1 *In case there is an index r_0 such that $t_{r_0+k-\nu} - t_{r_0} = 0$, the entire coefficient $c_{r_0}^{[\nu+1]}$ is set equal to zero.*

Example 2.1 *As an example for the above recursive formula, we consider the case $\nu = 1$. Then*

$$\frac{d}{dx} \left(\sum_{j=1}^m c_j B_{j,k,\mathbf{t}}(x) \right) = (k-1) \sum_{j=1}^{m+1} \frac{c_j - c_{j-1}}{t_{j+k-1} - t_j} B_{j,k-1,\mathbf{t}}(x), \quad (2.4)$$

with $c_0 := 0 =: c_{m+1}$.

Definition 2.1 *A B-spline whose knots are all simple and uniformly spaced, i.e., $t_{j+1} - t_j = \text{constant} > 0$, for all j , is termed a cardinal B-spline.*

We may choose the constant to be 1 and the knots to be $t_i = i$, $i = 0, \dots, n$. Then the cardinal B-splines $B_n := B_{0,n,(0,1,\dots,n)}$, $n \in \mathbb{N}$, can be generated by iterated convolution

of the characteristic function $\chi_{[0,1]}$ of the unit interval:

$$B_n = B_{n-1} * B_1 = \int_0^1 B_{n-1}(\cdot - t) dt = \sum_{j=1}^{n-1} \chi_{[0,1]}, \quad 2 \leq n \in \mathbb{N}.$$

2.2 Schoenberg's Variation Diminishing Spline Approximation

Sometimes, particularly in the presence of noise, it is advantageous to approximate rather than interpolate measured data. One possible approach is to use Schoenberg's variation diminishing spline approximation. To this end, let f be a function belonging to some function space \mathcal{F} and $\mathbf{t} = (t_i)_{1 \leq i \leq n+k}$ a knot vector contained in the domain of f . The image of the linear operator $V : \mathcal{F} \rightarrow \mathfrak{S}_{k,\mathbf{t}}$ defined by³

$$Vf := \sum_{i=1}^n f(t_{ik}^*) B_{i,k,\mathbf{t}^*}, \quad (2.5)$$

where $\mathbf{t}^* := (t_{ik}^*)_{1 \leq i \leq n+k}$ with

$$t_{ik}^* := \frac{t_{i+1} + \cdots + t_{i+k-1}}{k-1}, \quad (2.6)$$

is called Schoenberg's variation diminishing spline approximation of f . Using the properties of B-splines, it is not difficult to show that the operator V has the following characteristics. (See [5] for more details.)

- (i) $VL = L$, for any line L , or equivalently, for any polynomial L of degree one.
- (ii) The spline approximation Vf to f crosses any straight line at most as many times as f does. (Variation Diminution Property.)
- (iii) V preserves nonnegativity, monotonicity, and convexity. Hence, Vf is a shape-preserving positive linear approximation of f . In particular, V is a positive monotone linear operator.
- (iv) Vf is a local approximation of f ; on any interval $[t_i, t_{i+1}]$, the value of Vf depends only on the nearby values $t_{i-k+1,k}^*, \dots, t_{i,k}^*$.
- (v) V is in general not interpolatory, i.e., $(Vf)(x_j) \neq x_j$ with x_j defined as in Eqn. (2.2).

The operator V is also called a quasi-interpolant (of order one) and the induced process (i.e. Schoenberg's variation diminution) an example of quasi-interpolation.

Now suppose that $f = \sum_{i=1}^{n+k} c_i B_{i,k,\mathbf{t}}$ is the B-spline representation of f (which interpolates the x_j). The points (t_{ik}^*, c_i) are called the control points for f and the linear spline interpolating the control points the control polygon. The graph of Vf is contained within the control polygon.

For our later purposes, we will concentrate on fourth order B-splines. To obtain a variation diminishing fourth order B-spline representation of the samples $\{f(x_i) : i = 1, \dots, n\}$ of a function f defined on an interval $[a, b]$ with $a = x_1 < x_2 < \dots < x_n = b$, it

³ The operator V depends on the B-spline order k and the knot vector \mathbf{t} . However, unless necessary, this dependence will be suppressed.

is customary to choose as a knot vector

$$\boldsymbol{\tau} := (x_1, x_1, x_1, x_1, x_3, x_4, \dots, x_{n-2}, x_n, x_n, x_n).$$

This particular choice is motivated by the fact that at the endpoints no continuity requirements are necessary; hence x_1 and x_n have knot multiplicity four. As the total number of knots has to be equal to $n + 4$, and the highest possible smoothness, namely $\nu = 3$, is desired in the interior of $[a, b]$, x_2 and x_{n-1} need to be excluded from the knot sequence. Moreover, f is only defined on $[a, b]$ and no smoothness information is available beyond the endpoints a and b .

The image of the operator V now reads

$$Vf = \sum_{i=1}^n f(x_i) B_{i,4,\boldsymbol{\tau}} =: \sum_{i=1}^n f(x_i) B_i.$$

In the expression for the B-spline function, we will drop the index $k = 4$ and the reference to the specific knot vector $\boldsymbol{\tau}$ whenever we refer to this special case and no ambiguity arises.

2.3 Hermite Interpolatory Splines

In this section, we summarize some results regarding Hermite interpolation. More material and details may be found in [8].

Definition 2.2 Let $X := \{a = x_1 < x_2 < \dots < x_{n-1} < x_n = b\}$ be a partition of $[a, b]$ and let $d \in \mathbb{N}$. A function $\varphi : [a, b] \rightarrow \mathbb{R}$ is called a *Hermite spline* if it enjoys the following properties.

- (i) $\varphi \in C^{d-1}[a, b]$;
- (ii) φ when restricted to any subinterval $I_i := [x_i, x_{i+1}] \subset [a, b]$, $i = 1, \dots, n-1$, agrees with a polynomial of degree at most $2d - 1$, i.e., $\varphi|_{I_i} \in \Pi^{2d-1}$.

The linear space of all such Hermite splines will be denoted by H_X^d .

Of particular interest to us is the case $d := 2$, namely the cubic Hermite splines. It can be shown [8] that

$$H_X^2 = \text{span} \{J_i, S_i : i = 1, \dots, n\},$$

where J_i and S_i satisfy

$$\begin{aligned} J_i(x_j) &= \delta_{ij}, & \left(\frac{d}{dx} J_i\right)(x_j) &= 0 \\ S_i(x_j) &= 0, & \left(\frac{d}{dx} J_i\right)(x_j) &= \delta_{ij}, \end{aligned}$$

for all $i, j = 1, \dots, n$. Here δ_{ij} denotes the Kronecker delta symbol. The functions J_i and S_i are explicitly given by

$$J_i(x) := \begin{cases} 3 \left(\frac{x - x_{i-1}}{x_i - x_{i-1}} \right)^2 - 2 \left(\frac{x - x_{i-1}}{x_i - x_{i-1}} \right)^3, & x \in [x_{i-1}, x_i]; \\ 3 \left(\frac{x_{i+1} - x}{x_{i+1} - x_i} \right)^2 - 2 \left(\frac{x_{i+1} - x}{x_{i+1} - x_i} \right)^3, & x \in (x_i, x_{i+1}]; \\ 0, & \text{otherwise;} \end{cases}$$

and

$$S_i(x) := \begin{cases} -\frac{(x - x_{i-1})^2}{x_i - x_{i-1}} + \frac{(x - x_{i-1})^3}{(x_i - x_{i-1})^2}, & x \in [x_{i-1}, x_i]; \\ \frac{(x_{i+1} - x)^2}{x_{i+1} - x_i} - \frac{(x_{i+1} - x)^3}{(x_{i+1} - x_i)^2}, & x \in (x_i, x_{i+1}]; \\ 0, & \text{otherwise.} \end{cases}$$

Here quantities corresponding to indices outside the range $1, \dots, n$ are set equal to zero. Hence, for any $\varphi \in H_X^d$ we have the representation

$$\varphi = \sum_{i=1}^n \varphi(x_i) J_i + \frac{d\varphi}{dx}(x_i) S_i.$$

In particular, suppose one is given a finite collection of triplets $\Delta := \{(x_i, y_i, \eta_i) : i = 1, \dots, n\}$ with x_i being the abscissae and (approximate) ordinates y_i and (approximate) derivatives η_i are given, then

$$\sum_{i=1}^n y_i J_i + \eta_i S_i$$

represents a cubic Hermite spline approximation to Δ .

2.4 Approximation-theoretic Results

In this subsection, we briefly present some of the approximation-theoretic results for B-splines and Hermite splines. The proofs of these results may be found in [3, 8] or any other work on the approximation theory of splines.

For the remainder of this section, we assume that we are given a knot vector $\mathbf{t} = (t_i)_{1 \leq i \leq n+k}$ with $t_i < t_{i+k}$ for all i , $t_1 = t_2 = t_3 = t_4 =: a$, $t_{n+1} = t_{n+2} = t_{n+3} = t_{n+4} =: b$, and a function $f : [a, b] \rightarrow \mathbb{R}$.

Before stating the approximation results, we recall the concept of modulus of continuity.

Definition 2.3 Suppose $f : [a, b] \rightarrow \mathbb{R}$ and $h > 0$. Then the modulus of continuity of f is defined as

$$\omega(f; h) := \max\{|f(x) - f(x')| : |x - x'| \leq h; x, x' \in [a, b]\}.$$

Note that for fixed f , the modulus $\omega(f; h)$ is a function of h and the continuity and smoothness properties of f depend on how fast $\omega(f; h)$ goes to zero as $h \rightarrow 0$.

Now we are ready to state some approximation-theoretic results about B-splines and Hermite splines. To this end, let

$$\text{dist}(f, \mathfrak{S}_{k,\mathbf{t}}) := \min\{\|f - \mathfrak{s}\|_\infty : \mathfrak{s} \in \mathfrak{S}_{k,\mathbf{t}}\}$$

be the distance between f and its B-spline approximation, i.e., the best uniform approximation of f by elements from $\mathfrak{S}_{k,\mathbf{t}}$. (Here $\|\cdot\|_\infty$ denotes the sup-norm.)

Theorem 2.3 *For any $j \in \{0, 1, \dots, k-1\}$ there exist constants $c(j, k)$ depending only on j and k such that for all knot vectors \mathbf{t} of the form introduced above and for all $f \in C^j[a, b]$*

$$\text{dist}(f, \mathfrak{S}_{k,\mathbf{t}}) \leq c(j, k) \|\mathbf{t}\|^j \omega(f^{(j)}; \|\mathbf{t}\|),$$

where $\|\mathbf{t}\| := \max\{|t_i - t_{i-1}| : i = 2, \dots, n+k\}$ denotes the fineness of the knot sequence \mathbf{t} .

If $j = k-1$, then as $\omega(f^{(k-1)}; h) \leq h \|f^{(k)}\|_\infty$, one has

$$\text{dist}(f, \mathfrak{S}_{k,\mathbf{t}}) \leq c(k) \|\mathbf{t}\|^k \|f^{(k)}\|_\infty.$$

The corresponding result for Schoenberg variation diminishing splines is as follows.

Theorem 2.4 *Assume that $f \in C^2[a, b]$. Then*

$$\|f - Vf\|_\infty \leq c(k) \|\mathbf{t}\|^2 \|f^{(2)}\|_\infty,$$

where the constant $c(k)$ depends only on k .

It can be shown that even if $f \in C^r$, with $r > 2$, the exponent of the term $\|\mathbf{t}\|$ remains unchanged. In other words, for very smooth functions f , Schoenberg's variation diminishing splines do not provide optimal approximation.

The approximation-theoretic estimate for Hermite splines is provided in the next result.

Theorem 2.5 *Suppose that a set of interpolation points $X := \{(x_i, y_i) : i = 1, \dots, n\}$ is given, where the ordinates y_i are the samples of a function $f \in C^{2d}[a, b]$. Let $\varphi \in H_X^d$ and denote by $\|X\| := \max\{|x_i - x_{i-1}| : i = 2, \dots, n\}$ the fineness of the partition X of $[a, b]$. Then*

$$\left\| f^{(\ell)} - \varphi^{(\ell)} \right\|_\infty \leq \frac{\|X\|^{2d-\ell}}{2^{2d-2\ell} \ell! (2d-2\ell)!} \left\| f^{(2d)} \right\|_\infty, \quad \ell = 0, 1, \dots, d.$$

Proof See, for instance, [3]. □

3 Reconstruction Algorithm for Smooth Functions

In this section, a reconstruction algorithm for smooth functions whose values over an interval with positive Lebesgue measure are unknown is presented and error estimates for the approximation derived.

3.1 The Reconstruction Algorithm

To this end, let $\mathbf{y} := \{y_1, \dots, y_n\}$, $10 \leq n \in \mathbb{N}$, be a set of measured data points supported on the set of abscissae $\{\mathbf{x} := x_1, \dots, x_n\}$ and interpreted as the samples of a C^2 function $f : [x_1, x_n] \rightarrow \mathbb{R}$. We will consider the case when some of the sampled or measured data \mathbf{y} is incomplete, i.e., that data points are missing and/or unreliable. To this end, we assume that there exist $p, q \in \{1, \dots, n\}$ with $q \geq p + 1$ such that the data has the following form:

$$\mathbf{y}_i = \begin{cases} y_i, & i = 1, \dots, p; \\ \emptyset, & i = p + 1, \dots, q - 1; \\ y_i, & i = q, \dots, n. \end{cases} \quad (3.1)$$

The exact values of f on the interval $[x_{p+1}, x_{q-1}]$ are unknown.⁶ We term a function whose samples are of the form (3.1) an incomplete function. Note that an incomplete function \tilde{f} is necessarily a piecewise function of the form

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [x_1, x_p] \cup [x_q, x_n]; \\ \text{undefined}, & x \in [x_p, x_q]. \end{cases}$$

We denote the first p values of \mathbf{y} by \mathbf{y}^b and the last $n - q + 1$ values of \mathbf{y} by \mathbf{y}^\sharp .

With the set of measurements \mathbf{y} , we associate knot vectors

$$\boldsymbol{\tau}^b := (x_1, x_1, x_1, x_1, x_3, \dots, x_{p-2}, x_p, x_p, x_p, x_p)$$

and

$$\boldsymbol{\tau}^\sharp := (x_q, x_q, x_q, x_q, x_{q+2}, \dots, x_{n-2}, x_n, x_n, x_n, x_n).$$

and form Schoenberg variation diminishing spline approximations of the form

$$f^b := \sum_{i=1}^p y_i B_{i, \boldsymbol{\tau}^b} \quad (3.2 a)$$

and

$$f^\sharp := \sum_{i=q}^n y_i B_{i, \boldsymbol{\tau}^\sharp}, \quad (3.2 b)$$

which approximate the measured data supported on $\mathbf{x}^b := \{x_1, \dots, x_p\}$ and $\mathbf{x}^\sharp := \{x_q, \dots, x_n\}$, respectively. Our objective is to try to recover the values of f between x_p and x_q . This, of course, is a highly ill-posed problem with infinitely many solutions unless more information about the function f is known.

Assumption (A): The function f is assumed to have exactly one critical value on the interval $[x_{p+1}, x_{q-1}]$.

⁶ There is of course also the possibility that the function values are unknown over more than one nonempty interval of positive Lebesgue measure. However, for the reconstruction algorithm, it suffices to consider only one interval as similar arguments could be applied to the case of more than one interval.

Under this assumption, we define \hat{f} as

$$\hat{f} := \begin{cases} f^b, & x \in [x_1, x_p]; \\ f^\sharp, & x \in [x_q, x_n]. \end{cases} \quad (3.3)$$

Using the representations given by Eqns. (3.2 a), the derivatives of \hat{f} are computed at $\boldsymbol{\xi} := \boldsymbol{\xi}^b \cup \boldsymbol{\xi}^\sharp$, where

$$\boldsymbol{\xi}^b := \{x_1, \dots, x_{p-1}, x_p - \delta_p\} \quad \text{and} \quad \boldsymbol{\xi}^\sharp := \{x_q + \delta_q, \dots, x_{q+1}, x_n\},$$

for chosen $\delta_p, \delta_q > 0$. The reason for evaluation of the derivatives at $x_p - \delta_p$, respectively, $x_q + \delta_q$ is due to the fact that in general $p \leq \sup\{x : f(x) \leq y_{\max}\}$ and $q \geq \inf\{x : f(x) \leq y_{\max}\}$.

Eqn. (2.4) implies for $k := 4$

$$\begin{aligned} \eta_j := \frac{d\hat{f}}{dx}(\boldsymbol{\xi})_j &= \begin{cases} \frac{d}{dx} \sum_{i=1}^p y_i B_{i,4,\boldsymbol{\tau}^b}((\boldsymbol{\xi})_j), & x \in [x_1, x_p] \\ \frac{d}{dx} \sum_{i=1}^p y_i B_{i,4,\boldsymbol{\tau}^\sharp}((\boldsymbol{\xi})_j), & x \in [x_q, x_n] \end{cases} \\ &= \begin{cases} \sum_{i=1}^{p+1} \frac{3(y_i - y_{i-1})}{\tau_{i+3}^b - \tau_i^b} B_{i,3,\boldsymbol{\tau}^b}((\boldsymbol{\xi})_j), & x \in [x_1, x_p] \\ \sum_{i=q}^{n+1} \frac{3(y_i - y_{i-1})}{\tau_{i+3}^\sharp - \tau_i^\sharp} B_{i,3,\boldsymbol{\tau}^\sharp}((\boldsymbol{\xi})_j), & x \in [x_q, x_n] \end{cases} \\ &= \begin{cases} \frac{3(y_{j-1} - y_{j-2})}{\tau_{j+2}^b - \tau_{j-1}^b} B_{j-1,3,\boldsymbol{\tau}^b}((\boldsymbol{\xi})_j) + \frac{3(y_j - y_{j-1})}{\tau_{j+3}^b - \tau_j^b} B_{j,3,\boldsymbol{\tau}^b}((\boldsymbol{\xi})_j), & x_j \in [x_1, x_p] \\ \frac{3(y_{j-1} - y_{j-2})}{\tau_{j+2}^\sharp - \tau_{j-1}^\sharp} B_{j-1,3,\boldsymbol{\tau}^\sharp}((\boldsymbol{\xi})_j) + \frac{3(y_j - y_{j-1})}{\tau_{j+3}^\sharp - \tau_j^\sharp} B_{j,3,\boldsymbol{\tau}^\sharp}((\boldsymbol{\xi})_j), & x_j \in [x_q, x_n]. \end{cases} \end{aligned}$$

Employing Eqn. (2.1) and the specific form of the knot vectors $\boldsymbol{\tau}^b$ and $\boldsymbol{\tau}^\sharp$, one obtains

for η_j the following values.

$$\eta_j = \begin{cases} 0, & j = 1; \\ \frac{3(y_2 - y_1)}{4(x_3 - x_1)}, & j = 2; \\ \frac{y_3 - y_2}{x_4 - x_1}, & j = 3; \\ \frac{3(y_4 - y_3)}{2(x_5 - x_1)}, & j = 4; \\ \frac{3(y_j - y_{j-1})}{2(x_{j+1} - x_{j-2})}, & j = 5, \dots, p-2; \\ \frac{3(y_{p-1} - y_{p-2})}{2(x_p - x_{p-4})}, & j = p-1; \\ \frac{3(y_{p-1} - y_{p-2})}{x_p - x_{p-3}} B_{p-1,3,\tau^b}(x_p - \delta_p) + \frac{3(y_p - y_{p-1})}{x_p - x_{p-2}} B_{p,3,\tau^b}(x_p - \delta_p), & j = p; \\ \frac{3(y_{q+1} - y_q)}{x_{q+2} - x_q} B_{q,3,\tau^\#}(x_q + \delta_p) + \frac{3(y_{q+2} - y_{q+1})}{x_{q+3} - x_q} B_{q+1,3,\tau^\#}(x_q + \delta_p), & j = p+1; \\ \frac{3(y_{q+2} - y_{q+1})}{2(x_{q+4} - x_q)}, & j = p+2; \\ \frac{3(y_{q-p+j} - y_{q-p+j-1})}{2(x_{q-p+j+2} - x_{q-p+j-1})}, & j = p+3, \dots, n+p-q-3; \end{cases}$$

and

$$\eta_j = \begin{cases} \frac{3(y_{n-3} - y_{n-4})}{2(x_n - x_{n-4})}, & j = n+p-q-2; \\ \frac{y_{n-2} - y_{n-3}}{x_n - x_{n-3}}, & j = n+p-q-1; \\ \frac{3(y_n - y_{n-1})}{4(x_n - x_{n-2})}, & j = n+p-q; \\ 0, & j = n+p-q+1. \end{cases}$$

Now a cubic Hermite spline interpolation is performed on the triple $(\mathbf{x}^b \cup \mathbf{x}^\#, \mathbf{y}^b \cup \mathbf{y}^\#, \boldsymbol{\eta}^*)$,

where $\boldsymbol{\eta}^* := (\eta_j)_{1 \leq j \leq n+p-q+1}$. For this purpose, set

$$\boldsymbol{x}^* := \boldsymbol{x}^b \cup \boldsymbol{x}^\sharp \quad \text{and} \quad \boldsymbol{y}^* := \boldsymbol{y}^b \cup \boldsymbol{y}^\sharp,$$

and denote the cubic Hermite splines defined for the knot set \boldsymbol{x}^* by J_j^* and S_j^* , $j = 1, \dots, n+p-q+1$. The function

$$\begin{aligned} f^* &: [x_1, x_n] \rightarrow \mathbb{R} \\ f^*(x) &:= \sum_{j=1}^{n+p-q+1} y_j^* J_j^*(x) + \eta_j^* S_j^*(x) \end{aligned}$$

is termed the reconstruction of f .

The error between the reconstruction f^* and the original function f at the extremum value is defined by

$$E(f, f^*) := |f(\mathfrak{r}) - f^*(\mathfrak{r}^*)|,$$

where $\mathfrak{r} := \operatorname{argmax} f$ and $\mathfrak{r}^* := \operatorname{argmax} f^*$. Here it was assumed that the extremum value is a maximum; otherwise consider $-f$ and $-f^*$.

Under the further assumption

Assumption (S): $y_p = y_q$ and $\eta_p = -\eta_q$,

\mathfrak{r}^* can be easily computed. This assumption is reasonable for functions that are symmetric or have small support and large ordinates. It is then also reasonable to let $\delta := \delta_p = \delta_q$. If f^* is the reconstruction of a positive function f , then $\mathfrak{r}^* \in [x_p, x_q]$. Using the assumption (S), we have for $f^*|_{[x_p, x_q]}$ the following expression:

$$\begin{aligned} f^*(x) &= y_p^* J_p^*(x) + \eta_p^* S_p^*(x) + y_q^* J_q^*(x) + \eta_q^* S_q^*(x) \\ &= \eta^* x^2 - \eta^*(x_p + x_q)x + x_p x_q \eta^* + y^*(x_p - x_q), \quad x \in [x_p, x_q]. \end{aligned}$$

Here, we set $y^* := y_p^* = y_q^*$ and $\eta^* := \eta_p^* = -\eta_q^*$. Now if f^* has a maximum on $[x_p, x_q]$ then there exists an $\tilde{x} \in (x_p, x_q)$ such that $f^{*\prime}(\tilde{x}) = 0$ and $f^{*\prime\prime}(\tilde{x}) < 0$. (Here, \prime denotes the derivative with respect to the independent variable.) It is easily shown that

$$f^{*\prime}|_{[x_p, x_q]}(x) = \frac{\eta^*(2x - x_p - x_q)}{x_p - x_q} = 0 \quad \iff \quad x = \frac{x_p + x_q}{2}$$

and that

$$f^{*\prime\prime}|_{[x_p, x_q]} \left(\frac{x_p + x_q}{2} \right) = \frac{2\eta^*}{x_p - x_q} < 0.$$

Hence,

$$\mathfrak{r}^* = \frac{x_p + x_q}{2}$$

and

$$\max\{f^* : x \in [x_p, x_q]\} = f^*(\mathfrak{r}^*) = y^* + \frac{\eta^*(x_q - x_p)}{4}. \quad (3.4)$$

A special case occurs, when the samples \boldsymbol{x} are uniformly spaced, i.e., when there exists a positive constant θ such that $x_i = i\theta$, $i = 1, \dots, n$. The quantity θ is then the reciprocal

of the sampling rate. For this scenario, we have

$$\mathfrak{x}^* = \frac{x_p + x_q}{2} = \begin{cases} \nu\theta, & \text{if } p + q = 2\nu; \\ (\nu + \frac{1}{2})\theta = \frac{x_\nu + x_{\nu+1}}{2}, & \text{if } p + q = 2\nu + 1, \end{cases}$$

and

$$\begin{aligned} \eta^* &= \frac{3(y_{p-1} - y_{p-2})}{p\theta - (p-3)\theta} B_{p-1,3,\tau^b}(p\theta - \delta) + \frac{3(y_p - y_{p-1})}{p\theta - (p-2)\theta} B_{p,3,\tau^b}(p\theta - \delta) \\ &= \theta^{-1} \left[(y_{p-1} - y_{p-2}) B_{p-1,3,\tau^b}(p\theta - \delta) + \frac{3}{2} (y_p - y_{p-1}) B_{p,3,\tau^b}(p\theta - \delta) \right]. \end{aligned}$$

Employing the recursive formula (2.1) and using $x_i = i\theta$, $i = 1, \dots, n$, the B-splines $B_{p-1,3,\tau^b}$ and $B_{p,3,\tau^b}$ read as follows.

$$\begin{aligned} B_{p-1,3,\tau^b} &= \frac{[x - (p-3)\theta]^2}{3\theta^2} \chi_{[(p-3)\theta, (p-2)\theta]} - \frac{[5x + (12-5p)\theta][x - p\theta]}{12\theta^2} \chi_{[(p-2)\theta, p\theta]}, \\ B_{p,3,\tau^b} &= \frac{[x - (p-2)\theta]^2}{4\theta^2} \chi_{[(p-2)\theta, p\theta]}. \end{aligned}$$

Evaluation at $x = p\theta - \delta$, where $0 < \delta < \theta$, yields

$$B_{p-1,3,\tau^b}(p\theta - \delta) = \left(1 - \frac{5\delta}{12\theta}\right) \frac{\delta}{\theta} \quad \text{and} \quad B_{p,3,\tau^b}(p\theta - \delta) = \left(1 - \frac{\delta}{\theta}\right)^2, \quad (3.5)$$

which then gives for η^*

$$\eta^* = -\frac{\delta \left(1 - \frac{5\delta}{12\theta}\right)}{\theta^2} y_{p-2} + \left[\frac{\delta \left(1 - \frac{5\delta}{12\theta}\right)}{\theta^2} - \frac{3}{2} \left(1 - \frac{\delta}{\theta}\right)^2 \right] y_{p-1} + \frac{3}{2} \left(1 - \frac{\delta}{\theta}\right)^2 y_p.$$

Combining Eqns. (3.4) and (3.5), we obtain as the reconstructed maximum value

$$\begin{aligned} f^*(\mathfrak{x}^*) &= \frac{\delta(p-q)(12-5\delta)}{48\theta^2} y_{p-2} + \frac{(p-q)(18\theta^3 - 12\delta\theta(1+3\theta))}{48\theta^2} y_{p-1} \\ &\quad + \left[1 + \frac{3(q-p)(\delta-\theta)^2}{8\theta} \right] y_p. \end{aligned} \quad (3.6)$$

We summarize the results obtained above in a theorem.

Theorem 3.1 *Let $10 \leq n \in \mathbb{N}$ and suppose that $f : [x_1, x_n] \rightarrow \mathbb{R}$ is an incomplete function with uniform sampling set $\mathbf{x} = \{x_i = i\theta : i = 1, \dots, n; \theta > 0\}$ satisfying assumptions (A) and (S). Then the reconstructed maximum value of f on the interval $[x_p, x_q]$, where information about the function values is missing, is given by formula (3.6).*

3.2 Error Estimates

In order to calculate the error $E(f, f^*)$, we use again the assumptions (A) and (S), which now imply that $\mathfrak{r} = \mathfrak{r}^*$. In addition, we assume that $f \in C^4[x_1, x_n]$. As two dif-

ferent approximations are used, namely Schoenberg's variation diminution and Hermite Interpolation, and the latter of the two involved approximating derivatives of f over the interval $[x_p, x_q]$, the following chain of inequalities needs to be considered.

$$\begin{aligned} E(f, f^*) &= |f(\mathbf{x}) - f^*(\mathbf{x})| = \left| f(\mathbf{x}) - \left(\sum_{j=1}^{n+p-q+1} y_j^* J_j^*(\mathbf{x}) + \eta_j^* S_j^*(\mathbf{x}) \right) \right| \\ &\leq \left| f(\mathbf{x}) - \left(\sum_{j=1}^{n+p-q+1} f(x_j^*) J_j^*(\mathbf{x}) + f'(x_j^*) S_j^*(\mathbf{x}) \right) \right| \\ &\quad + \left| f(\mathbf{x}) - \left(\sum_{j=1}^{n+p-q+1} [f'(x_j^*) - \eta_j^*] S_j^*(\mathbf{x}) \right) \right| \end{aligned}$$

The first term in the above inequality can be estimated as follows (cf. p. 57, Eqn. (2.1.5.12)):

$$\begin{aligned} &\left| f(\mathbf{x}) - \left(\sum_{j=1}^{n+p-q+1} f(x_j^*) J_j^*(\mathbf{x}) + f'(x_j^*) S_j^*(\mathbf{x}) \right) \right| \\ &\leq \frac{|(\mathbf{x} - x_p)(x_q - \mathbf{x})|^2}{4!} \max_{\xi \in [x_p, x_q]} |f^{(4)}(\xi)| \quad (3.7) \\ &= \frac{(q-p)^2 \theta^4}{384} \max_{\xi \in [x_p, x_q]} |f^{(4)}(\xi)|. \end{aligned}$$

To estimate the second term, the error due to the derivative approximation, recall that $\eta_j^* = (Vf)'(x_j)$, $j \in \{1, \dots, p-1, p+1, \dots, n+p-q+1\}$, and $\eta_p^* = (Vf)'(x_p - \delta) = -\eta_q^*$ (by assumption (S)). Since \mathbf{x} is only contained in the support of S_p^* and S_q^* , it suffices to obtain an estimate for $|f'(x_p) - (Vf)'(x_p - \delta)|$. To this end, consider

$$\begin{aligned} |f'(x_p) - (Vf)'(x_p - \delta)| &\leq |f'(x_p) - (Vf)'(x_p)| + |(Vf)'(x_p) - (Vf)'(x_p - \delta)| \\ &\leq |f'(x_p) - (Vf)'(x_p)| + |(Vf)'(x_p) - (Vf)'(x_p)| \\ &\quad + |(Vf)'(x_p) - (Vf)'(x_p - \delta)|. \end{aligned}$$

If the function f is approximated by a fourth order B-spline expansion, then its derivative has to be expressed in terms of third order B-splines. Hence, the first term in the last inequality above can be estimated by (cf. [7], p. 50)

$$|f'(x_p) - (Vf)'(x_p)| \leq \frac{\theta^2}{2} \max_{\xi \in I(x_p)} |f^{(3)}(\xi)|, \quad (3.8)$$

where $I(x_p)$ denotes the support of all those third order B-splines that do not vanish at x_p . Note that since x_p is an “endpoint”, there is only one B-spline with this property, namely, $B_{p,3,\tau^\flat}$ and, therefore, $I(x_p) = [x_{p-2}, x_p]$.

The second term in the last inequality is the error between the Schoenberg approximation of the derivative of f and the derivative of the Schoenberg approximation of f . It can be estimated as follows. Given $Vf = \sum f(\tau_{ik}^*)B_{i,k}$, then

$$Vf' = \sum f'(\tau_{i,k-1}^*)B_{i,k-1}$$

but

$$(Vf)' = \sum f(\tau_{ik}^*)B'_{ik}.$$

Now,

$$\begin{aligned} (Vf)' &= \sum f(\tau_{ik}^*)B'_{ik} = (k-1) \sum \frac{f(\tau_{ik}^*) - f(\tau_{i-1,k}^*)}{\tau_{i+k-1} - \tau_i} B_{i,k-1} \\ &= \sum \frac{f(\tau_{ik}^*) - f(\tau_{i-1,k}^*)}{\tau_{ik}^* - \tau_{i-1,k}^*} B_{i,k-1} =: \sum (\Delta f)(\tau_{ik}^*) B_{i,k-1}, \end{aligned}$$

since

$$\tau_{ik}^* - \tau_{i-1,k}^* = \frac{\tau_{i+1} + \dots + \tau_{i+k-1}}{k-1} - \frac{\tau_i + \dots + \tau_{i-1+k-1}}{k-1} = \frac{\tau_{i+k-1} - \tau_i}{k-1}.$$

Hence, the difference between Vf' and $(Vf)'$ is the replacement of the derivative f' by the finite difference Δf . Therefore,

$$\begin{aligned} |(Vf')(x_p) - (Vf)'(x_p)| &\leq \sum |f'(\tau_{i,3}^*) - (\Delta f)(\tau_{i,4}^*)| B_{i,3,\tau^\flat}(x_p) \\ &= |f'(\tau_{p,3}^*) - (\Delta f)(\tau_{p,4}^*)| B_{p,3,\tau^\flat}(x_p) \\ &= |f'x_p - (\Delta f)(x_p)|. \end{aligned}$$

Using the well-known formula

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x+\vartheta h), \quad 0 < \vartheta < 1, \quad (3.9)$$

and the fact that $x_i = i\theta$, we see that $h = \tau_{p,4}^* - \tau_{p-1,4}^* = \frac{2}{3}\theta$ and thus

$$|f'(x_p) - (\Delta f)x_p| = \frac{\theta}{3} \left| f'' \left(\left(p - \frac{2}{3} \right) \theta + \vartheta \frac{2\theta}{3} \right) \right| \leq \frac{\theta}{3} \max_{\xi \in [x_{p-1}, x_p]} |f''(\xi)|. \quad (3.10)$$

The third term, $|(Vf)'(x_p) - (Vf)'(x_p - \delta)|$, can also be estimated using Eqn. (3.9).¹⁰

$$\begin{aligned} |(Vf)'(x_p) - (Vf)'(x_p - \delta)| &= \delta |(Vf)''(x_p - \vartheta\delta)| \\ &\leq \delta \max_{\xi \in [x_p - \delta, x_p]} |(Vf)''(\xi)|. \end{aligned} \tag{3.11}$$

Putting estimates (3.8), (3.10), and (3.11) together and using the fact that $|S_p^*(\mathbf{r})| = |S_q^*(\mathbf{r})| = (q-p)\theta/8$, we obtain the following end result.

Theorem 3.2 *The error $E(f, f^*)$ between the original function f and its reconstruction f^* obeys the estimate*

$$\begin{aligned} E(f, f^*) &\leq \frac{(q-p)^2\theta^4}{384} \left\| f^{(4)} \right\|_{[x_p, x_q], \infty} \\ &\quad + \frac{(q-p)\theta^3}{8} \max \left\{ \left\| f^{(3)} \right\|_{[x_{p-2}, x_p], \infty}, \left\| f^{(3)} \right\|_{[x_q, x_{q+2}], \infty} \right\} \\ &\quad + \delta \left(\left\| (Vf)'' \right\|_{[x_p - \delta, x_p], \infty} + \left\| (Vf)'' \right\|_{[x_q, x_q + \delta], \infty} \right), \end{aligned}$$

where we used the abbreviation

$$\|\phi\|_{[a,b], \infty} := \max_{\xi \in [a,b]} |\phi(\xi)|.$$

We note that the error increases quadratically with the amount of the missing information, i.e., the length of the interval $[x_p, x_q]$.

4 A Numerical Experiment

In this section, we present an example of a reconstruction using the algorithm described in the previous section. For this purpose, we chose a function f that represents the induced static magnetic field generated by a small spherical cavity of radius 10 mm in a ferromagnetic material when an external magnetic field of strength 25,000 A/m is applied. About 65 uniformly spaced knots were used with $\theta = 2$ mm. The graph of f is then cut off at different ordinate values, namely at 90%, 80%, 70%, and 50% of the maximum value. (See the graphs on the left hand side in Figure 1.) For this particular example, assumptions (A) and (S) in Section 4 are valid. The function f is reconstructed with the results displayed on the left hand side of Figure 1.

References

- [1] CANDÈS, E. J., ROMBERG, J., & TAO, T. 2005 Stable Signal Recovery from Incomplete and Inaccurate Measurements. *Comm. Pure Appl. Math.* **59**, 1207–1223

¹⁰ The difference can be computed exactly. The result is

$$\frac{\delta}{\theta^2} \left[3(y_p - y_{p-1}) \left(2 + \frac{\delta}{\theta} \right) - (y_{p-1} - y_{p-2}) \left(1 - \frac{5\delta}{12\theta} \right)^2 \right].$$

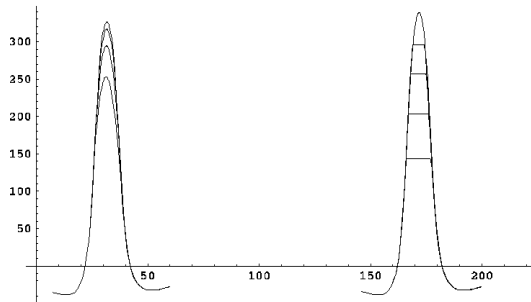


FIGURE 1. Reconstruction of a function from different sets of incomplete data..

- [2] CANDÈS, E. J., ROMBERG, J., & TAO, T. 2004 Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information. *IEEE Trans. Inform. Theory* **52**, 489–509.
- [3] CIARLET, P. G., SCHULZ, M. H., & VARGA, R. S. 1967 Numerical Methods of High-Accuracy for Nonlinear Boundary Problems I. One Dimensional Problems. *Numer. Math.* **9**, 294–430.
- [4] DAHMEN, W. 1999 B-Splines in Analysis, Algebra, and Applications. *Travaux Mathématiques*, 15–76.
- [5] DE BOOR, C. 2001 *A Practical Guide to Splines*. New York. Springer Verlag.
- [6] LYCHE, T. 1992 Knot Removal for Spline Curves and Surfaces. In *Approximation Theory VII* (Cheney, E. W, Chui, C. K. & Schumaker, L.L., eds.) New York. Academic Press, 207–226.
- [7] REIF, U. 2004 *Splineapproximation*. Vorlesungsskriptum, Technical University Darmstadt, Germany.
- [8] STOER, J. & BULIRSCH, R. 1993 *Introduction to Numerical Analysis*. New York. Springer Verlag.