1   A representation for piecewise polynomial functions

Let $\xi = \{\xi_1 < \xi_2 < \cdots < \xi_{l+1}\}$ be a set of strictly increasing series of points, and let k be a positive integer. Further, let $P_1, \ldots, P_l$ denote a sequence of l polynomials of order k. Then the correponding piecewise polynomial (pp) function of order k is defined as follows:

$$f(x) = P_i(x) \text{ if } \xi_i < x < \xi_{i+1}$$

for $i = 1, \ldots, l$. $\{\xi\}$ are known as the breakpoints of f. At the interior breakpoints, $\xi_2, \ldots, \xi_l$, the function value is defined by specifying f to be right continuous; that is,

$$f(\xi_i) = f(\xi_i^+), \quad i = 2, \ldots, l$$

However, in a sense, without this specification, the function has two values at any interior breakpoint: the value it gets from the polynomial piece to the left of the breakpoint, $f(\xi_i^-) = P_{i-1}(\xi_i)$, in addition to the value it gets from the polynomial piece to the right of the breakpoint, $f(\xi_i^+) = P_i(\xi_i)$. To properly define the function, one can specify f to be right-continuous:

$$f(\xi_i) \equiv f(\xi_i^+) \tag{1}$$

Denote the set of pp functions of order k with breakpoints $\xi = \{\xi_1, \ldots, \xi_{l+1}\}$ by

$$\mathcal{P}_{k,\xi}.$$

$\mathcal{P}_{k,\xi}$ is a linear space having dimension kl, as it consists of l polynomials, each having k polynomial coefficients. The jth derivative of a pp f,

$$D^j f$$

is a pp function of order $k - j$ having the same breakpoint sequence and constructed from the same j^{th} derivatives of the polynomial pieces from which f was constructed. This "definition" dodges much of the complicated discussion of the derivatives of a pp function at its breakpoints and thus must be treated with considerable care in context of the fundamental theorem of calculus.

Proposition 1.1. *A pp function, f satisfies*

$$f(x) - f(a) = \int_a^x (Df)(t)\,dt \quad \text{for all} \quad x$$

if and only if f is a continuous function.

Consider a piecewise constant function f: by the previous definition, its first derivative is identically zero, and is therefore equal to the usual derivative of f if and only if f is constant. The following definition gives the information necessary for a convenient and efficient representation of this class of functions:

Definition 1.1. The *piecewise polynomial (pp) representation* of a function $f \in \mathcal{P}_{k,\xi}$ consists of

I. integers k and l, specifying the order and number of polynomial pieces, respectively,

II. a strictly increasing set of breakpoints, $\xi_1, \xi_2, \ldots, \xi_{l+1}$,

III. and the set of values of the right derivatives at each of the breakpoints,

$$c_{ij} = D^j f\left(\xi_i^+\right), \quad j = 1, \ldots, k; \quad i = 1, \ldots, l$$

2 The truncated power basis and the spaces $\mathcal{P}_{k,\xi,\nu}$

This prerequisite information is merely for the ability to responsibly refer to the set of piecewise polynomial functions and have a shorthand way of doing so. These means enable us to introduce two sets of basis functions: first, the truncated power basis, followed by B-spline basis functions. We will see that both are closely related, with the former having some properties which leave them unattractive for function approximation and thus present the construction of B-splines and how to use them to construct a representation of $\mathcal{P}_k$. In practice, one typically is given some information about an unknown function, g, and the task is to construct a function $f \in \mathcal{P}_{k,\xi}$ which satisfies conditions that g also satisfies, and in addition, has a certain number of continuous derivatives. These conditions define a subspace of $\mathcal{P}_{k,\xi}$, $\mathcal{P}_{k,\xi,\nu}$ for which we will need a corresponding basis.

2.1 Example: histogram smoothing by parabolic splines

For illustrative purposes, consider the task of smoothing a histogram using parabolic splines. Suppose we are given points

$$\tau_1 < \tau_2 < \cdots < \tau_{n+1}$$

and non-negative numbers $h_1, h_2, \ldots, h_n$, with h_i denoting the height of the histogram over the interval (τ_i, τ_{i+1}). The histogram is an approximate representation of some underlying density function, g. Letting $\Delta \tau_i = \tau_{i+1} - \tau_i$, one may interpret $h_i \Delta \tau_i$ as (approximately) equal to the integral of g over $[\tau_i, \tau_{i+1}]$. One may impose the following interpolation conditions on our smooth function, f:

$$\int_{\tau_i}^{\tau_{i+1}} f(x)\, dx = h_i \Delta \tau_i$$

for $i = 1, \ldots, n$. Let f be a piecewise polynomial of order 3 having continuous first derivative:

$$f \in \mathcal{P}_{3,\xi} \cap \mathcal{C}^{(1)}$$

Choose the breakpoint sequence ξ to coincide with $\tau = \{\tau_1, \ldots, \tau_{n+1}\}$. If g is smooth and vanishes outside its support, $[\tau_1, \tau_{n+1}]$, then

$$g^{(j)}(\tau_1) = g^{(j)}(\tau_{n+1}) = 0,$$

for $j = 0, 1, \ldots, d$, where d characterizes the extent of the smoothness of g, we may also wish to require f to obey two additional interpolation constraints:

$$f(\tau_1) = f(\tau_{n+1}) = 0,$$

giving a total of $n + 2$ interpolation conditions. These, along with the $2(n-1)$ continuity conditions yield a total $3n$ constraints on the $3n$ polynomial coefficients,

$$c_{ji} \equiv D^{j-1} f\left(\xi_i^+\right).$$

These conditions lead to the system of equations:

$$
\begin{array}{llllll}
c_{11} & & & & & = 0 \\[2mm]
c_{11} + & c_{21}\dfrac{\Delta\tau_1}{2!} + & c_{31}\dfrac{(\Delta\tau_1)^2}{3!} & & & = h_1 \\[2mm]
c_{11} + & c_{21}\Delta\tau_1 + & c_{31}\dfrac{2(\Delta\tau_1)^2}{3!} - & c_{12} & & = 0 \\[2mm]
\vdots & c_{21} + & c_{31}\Delta\tau_1 \quad - & c_{22} & & = 0 \\[2mm]
& & c_{12} + & c_{22}\dfrac{\Delta\tau_2}{2} + & c_{32}\dfrac{(\Delta\tau_2)^2}{3!} & = h_2 \\[2mm]
& & c_{12} + & c_{22}\Delta\tau_2 + & c_{32}\dfrac{(\Delta\tau_2)^2}{2} \quad \ldots & = 0 \\[2mm]
& & & c_{22} + & c_{32}\Delta\tau_2 \quad \ldots & = 0 \\[2mm]
& & & & \ddots & \quad (2)
\end{array}
$$

2.2 The subspace, $\mathcal{P}_{k,\xi,\nu}$

One may quickly see that this system is two-thirds homogeneous; that is, for every integral interpolation constraint, we have two continuity constraints that lead to zeros on the right hand side of the equality. To solve 2, the homogeneous equations are solved, leaving a reduced set of equations. To do this, one may construct a set of linearly independent functions $\phi_1, \phi_2, \ldots$ of the same size as the number of interpolation constraints which satisfy the homogeneous equations in 2. The smoother, f, is then constructed within this subspace of $\mathcal{P}_{3,\xi}$ and has form

$$f = \sum_j \alpha_j \phi_j.$$

The $\{\phi_j\}$ span a particular subspace of $\mathcal{P}_{k,\xi}$, which is comprised of functions which satisfy the homogeneous equations in 2. In general, we may characterize these homogeneous equations

3

in terms of a given function having a particular number of continuous derivatives, which may be expressed as follows:

$$\jmath_{ij} f = 0, \quad i = 2, \ldots, l$$
$$j = 1, \ldots, \nu_i \tag{3}$$

where

$$\jmath_{ij} f = \lim_{x \to \xi^+} D^{j-1} f(x) - \lim_{x \to \xi^-} D^{j-1} f(x) \tag{4}$$

and $\nu = (\nu_1, \ldots, \nu_l)$ is a vector of non-negative integers. Each ν_i specifies the number of continuity conditions impose on the function at the ith breakpoint, ξ, and $\jmath_{ij}$ is simply the size of the jump in the $(j-1)^{st}$ derivative at ξ.

2.3 The truncated power functions

As homogeneous conditions specified in 3 are done so in terms of a linear operator applied to the functions in the space, those functions $\{g \in \mathcal{P}_{k,\xi}\}$ constitute a linear subspace, denoted

$$\mathcal{P}_{k,\xi,\nu}$$

Now that the subspace is defined, we need a basis for it. One such candidate set of basis functions is the truncated power basis. Define

$$(x - t)_+ = \max(0, x - t)$$

We may then define the truncated power function as follows:

$$(x - t)_+^p = \left((x - t)_+\right)^p$$
$$= \begin{cases} (x - t)^p, & x \geq t \\ 0 & x < t \end{cases} \tag{5}$$

The function $g(x) = ()_+^p$ is a piecewise polynomial with a single breakpoint at ξ, and is continuous at this breakpoint for $p > 0$. For $p = 0$, it has a jump of size 1 at ξ. Since

$$D(\cdot - \xi)_+^p = p(\cdot - \xi)_+^{p-1},$$

it is clear that g has $p - 1$ continuous derivatives. Define the set of linear operators $\{\lambda_{ij}\}$ and corresponding functions $\{\phi_{ij}\}$ as follows:

$$\lambda_{ij} f = \begin{cases} D^j f(\xi) & i = 1 \\ D^j f(\xi^+) - D^j f(\xi^-) & i = 2, \ldots, l \end{cases}$$
$$\phi_{ij}(x) = \begin{cases} \frac{(x - \xi_i)^p}{j!} & i = 1 \\ \frac{(x - \xi_i)_+^p}{j!} & i = 2, \ldots, l \end{cases} \tag{6}$$

for $j = 0, \ldots, k-1$. Per this definition, we have that $\phi_{ij} \in \mathcal{P}_{k,\xi}$ for each $i = 1, \ldots, l$. Further, we have that

$$\lambda_{ij}\phi_{pq} = \begin{cases} 1 & i = p, \ j = q \\ 0 & \text{otherwise} \end{cases},$$

making $\{\phi_{ij}\}$ a set of kl linearly independent functions, and since $\mathcal{P}_{k,\xi}$ has dimension kl, they constitute a basis for the space. We may represent any $g \in \mathcal{P}_{k,\xi}$ in the form

$$g = \sum_{i,j} (\lambda_{ij}g)\, \phi_{ij}$$

Rewriting this expansion in the terms presented in 6, we have that we may express any function in the space as

$$g(x) = \sum_{j=0}^{k-1} \left[D^j g(\xi_1) \frac{(x - \xi_1)^j}{j!} + \sum_{i=2}^{l} [D^j g(\xi_i^+) - D^j g(\xi_i^-)] \frac{(x - \xi_i)_+^j}{j!} \right] \tag{7}$$

From this representation of the function, one can see that the coefficients of the basis functions are explicitly defined in terms of jumps of various derivatives of g at the breakpoints. Thus, enforcing the homogeneous constraints

$$\mathcal{J}_{ij}f = 0, \quad i = 2, \ldots, l; \quad j = 1, \ldots, \nu_i$$

is accomplished simply by restricting our attention to functions of the form 7 for which these coefficients are zero. This implies that the reduced set of basis functions $\{\phi_{ij}\}$, $i = 1, \ldots, l$, $j = \nu_i, \ldots, k-1$ is a basis for the subspace, $\mathcal{P}_{k,\xi,\nu}$. (We let $\nu_1 = 0$.) Then, every function $g^* \in \mathcal{P}_{k,\xi}$ in satisfying the homogeneous equations may be written in exactly one way of the form

$$g^* = \sum_{i=1}^{l} \sum_{j=\nu_i}^{k-1} \alpha_{ij}\phi_{ij} \tag{8}$$

2.3.1 The pitfalls of the truncated power basis

While the truncated power basis initially appears quite attractive for smoothing problems involving piecewise polynomials, they exhibit properties that can lead to poor function representations. In order to discuss these characteristics, we must first introduce the notion of the *condition* of a function representation.

Definition 2.1. Consider a piecewise polynomial representation of a function, p:

$$p = \sum_{i=1}^{n} a_i P_i \tag{9}$$

where $a = (a_1, \ldots, a_n)$ is a coefficient vector and $\{P_i\}$ is a set of polynomial functions defined on a closed interval $[a, b]$. We define the *size* of the polynomial p to be

$$\|p\| = \max_{a \leq x \leq b} |p(x)|,$$

and similarly, we define the size of the coefficient vector a to be

$$\|a\| = \max_{1 \leq i \leq n} |a_i|,$$

Then, we can bound the size of the function p:

$$m \|a\| \leq \left\| \sum_{i=1}^{n} a_i P_i \right\| \leq M \|a\| \tag{10}$$

where

$$m = \min_a \frac{\|\sum_{i=1}^{n} a_i P_i\|}{\|a\|}, \quad \text{and}$$
$$M = \min_a \frac{\|\sum_{i=1}^{n} a_i P_i\|}{\|a\|}$$

The *condition* of the representation of p as written in 9 is given by

$$condition(P_i) = \frac{M}{m}$$

The condition of a function representation quantifies the extent to which a slight change in the coefficient vector will impact the function itself. To see this, rather than the coefficient vector a, consider instead a perturbation of a:

$$a + \delta a$$

The corresponding perturbed polynomial is given by

$$p + \delta p = \sum_{i=1}^{n} (a_i + \delta a_i) P_i$$

By 10, we then have that

$$\frac{m \|\delta a\|}{M \|a\|} \leq \frac{\|\delta p\|}{\|p\|} \leq \frac{M \|\delta a\|}{m \|a\|},$$

implying that a relative change of $\frac{\|\delta a\|}{\|a\|}$ to the coefficient vector may result in a relative change to the function p as large as $\frac{M}{m} = condition(P_i)$ times the relative change in a (and at least as large as $condition(P_i)^{-1}$). Note that the width of the interval

$$\left[condition(P_i)^{-1} \frac{\|\delta a\|}{\|a\|}, condition(P_i) \frac{\|\delta a\|}{\|a\|} \right]$$

is increasing in $condition\,(P_i)$; so large values of the condition of a representation imply that small relative changes in their corresponding coefficients may result in much smaller or much larger relative changes in the function being represented.

When representing a function $f \in \mathcal{P}_{k,\xi}$ as in 7, two issues of concern arise: first, if l is large, the value of the function at a point x can potentially rely on far more than just k of the basis coefficients. Additionally, if the breakpoints ξ are very irregularly spaced, then the truncated power basis can present poor condition, which, in turn, can result in poorly conditioned linear systems (like the specific example given by 2.) Consequently, small perturbations of the basis function coefficients result in disproportional changes in the function.
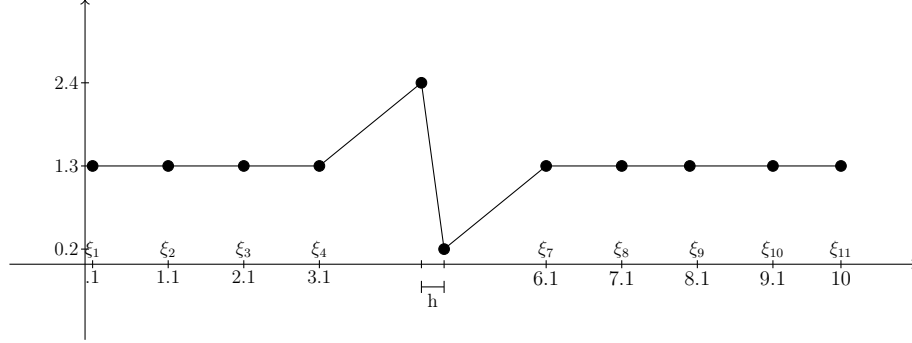

Figure 1: The linear system for the coefficients of the truncated power basis is ill-conditioned for the above piecewise linear function and choice of breakpoints, $\xi_1, \ldots, \xi_{11}$.

Example 2.1. Suppose we wish to construct a function $f \in \mathcal{P}_{k=2,\xi} \cap \mathcal{C}^{(0)}$ satisfying $f\,(\xi_i) = 1.3$ for $i = 1, \ldots, 4, 7, \ldots, 11$, $f\,(\xi_5) = 2.4$, and $f\,(\xi_6) = 0.2$ with breakpoints specified as pictured in Fig 1.

The function f is piecewise linear, so that $f \in \mathcal{P}_{2,\xi,\nu}$ with $\nu_i = 1$, $i = 1, \ldots, 11$. Then, for f is of the form

$$f\,(x) = \alpha + \beta\,(x - \xi_1) + \sum_{i=2}^{l} \alpha_i\,(x - \xi_i)_+$$

one can show that, given the continuity and interpolation constraints, $\alpha = 1.3$, $\beta = f'\,(\xi_1) = 0$, $\alpha_i = 0$ for $i \notin \{4, 5, 6, 7\}$, and

$$\begin{bmatrix} \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \end{bmatrix} = \begin{bmatrix} \frac{1.1}{\Delta\xi_4} \\ -\frac{2.2}{\Delta\xi_5} - \frac{1.1}{\Delta\xi_4} \\ \frac{1.1}{\Delta\xi_6} + \frac{2.2}{\Delta\xi_5} \\ -\frac{1.1}{\Delta\xi_6} \end{bmatrix} = \begin{bmatrix} \frac{1.1}{\Delta\xi_4} \\ -\frac{2.2}{h} - \frac{1.1}{\Delta\xi_4} \\ \frac{1.1}{\Delta\xi_6} + \frac{2.2}{h} \\ -\frac{1.1}{\Delta\xi_6} \end{bmatrix}$$

When $h = \Delta\xi_5 \to 0$, we have

$$(x - \xi_5)_+ \approx (x - \xi_6)_+$$
$$\alpha_5 \approx -\alpha_6 >> 1,$$

leading to a loss of significance in the evaluation of the function. For example, if we were to choose $\xi_5 = 4.5$, $\xi_6 = 4.8$, making $h = 0.3$, with five significant decimal digit arithmetic, then we get

$$\alpha = 1.30000, \ \beta = 0.00000, \ \alpha_4 = 0.78571, \ \alpha_5 = -8.11905, \ \alpha_6 = 8.17949, \ \alpha_7 = -0.84615$$

and the other $\alpha_i = 0$. Evaluating the corresponding function at $x = 9.6$ yields 1.299987 rather than the correct value of 1.3. This error becomes larger as h approaches 0.

A quick remedy for the problem in 2.1 is to replace the truncated power basis with the set of hat functions:

$$H_i(x) = \begin{cases} (x - \xi_i)/\Delta\xi_{i-1}, & \xi_{i-1} < x \leq \xi_i \\ (\xi_{i+1} - x)/\Delta\xi_i, & \xi_i < x \leq \xi_{i+1} \\ 0 & otherwise \end{cases}$$

To utilize the hat functions, we augment $\xi_1, \ldots, \xi_{12}$ with two additional breakpoints: $\xi_0 \leq \xi_1$ and $\xi_{12} \geq \xi_{11}$. Then,

$$f(x) = 1.3H_1(x) + \cdots + 1.3H_4(x) + 2.4H_5(x)$$
$$+ 0.2H_6(x) + 1.3H_7(x) + \cdots + 1.3H_{11}(x)$$

Even just using just two decimal digit arithmetic, we have $f(9.5) = 1.3$. Even as $h \to 0$, f is well represented using the hat functions as a basis. More generally speaking, *B-splines*, which are a generalization of these hat functions, overcome the issues that accompany the truncated power basis previously illustrated. The alternative basis is constructed by assembling linear combinations of the truncated power functions, forming a set of basis functions having "small" support: the functions vanish outside a small interval of their domain. In the following section, we will define the kth-order B-splines as scaled *divided differences* of truncated power functions. We will also show that every subspace $\mathcal{P}_{k,\xi,\nu}$ has a basis consisting of these functions, bringing forth a B-spline representation of any pp function.

3 The B-spline representation of piecewise polynomial functions

We will present an introduction to B-splines and their properties in the section to follow. As these basis functions are defined in terms of divided differences of the truncated power functions discussed in the previous section, we must first review the definition and properties of the divided difference operator. While there are many ways of defining divided differences, the following (somewhat nonconstructive definition) is intuitive and adequate for our purposes here.

3.1 The divided difference

Definition 3.1. The kth *divided difference* of a function g at the points $\tau_i, \ldots, \tau_{i+k}$, denoted

$$[\tau_i, \ldots, \tau_{i+k}] \, g,$$

is the leading coefficient of the polynomial of order $k+1$ which *agrees with g* at $\tau_i, \ldots, \tau_{i+k}$.

This leads us to the following definition:

Definition 3.2. Let $\{\tau_i\}_{i=1}^n$ denote a sequence of points. We say that a function f *agrees with* a function g at τ if, for every point τ which occurs m times in $\{\tau_i\}$, f and g agree up to m derivatives at τ; i.e.

$$f^{(i-1)} (\tau) = g^{(i-1)} (\tau), \quad i = 1, \ldots, m$$

3.1.1 Properties of the divided difference

Definition 3.1 yields the following properties of the kth divided difference:

I. Let $\{p_i\}$, $i = 1, 2, \ldots$ denote a sequence of polynomials with p_i having order i. If p_i agrees with a function g at $\tau_1, \ldots, \tau_i$ for $i = k, k+1$, then

$$p_{k+1} (x) = p_k (x) + (x - \tau_1) \cdot \cdots \cdot (x - \tau_k) [\tau_1, \ldots, \tau_{k+1}] \, g$$

Proof. Note that $p_{k+1} (x) - p_k (x)$ is a polynomial of order $k + 1$ and vanishes at $\tau_1, \ldots, \tau_k$, and, by definition, has $[\tau_1, \ldots, \tau_{k+1}] \, g$ as its leading coefficient. Consequently, $p_{k+1} - p_k$ is of the form

$$p_{k+1} (x) - p_k (x) = c \prod_{j=1}^k (x - \tau_j)$$

where

$$c = [\tau_1, \ldots, \tau_{k+1}] \, g$$

$\square$

This tells the reader that interpolating polynomials may be constructed using divided differences by adding the interpolation points one by one, giving the *Newton form* of the nth order polynomial which agrees with g at $\tau_1, \ldots, \tau_n$:

$$p_n (x) = \sum_{i=1}^n (x - \tau_1) \times \cdots \times (x - \tau_{i-1}) [\tau_1, \ldots, \tau_{i-1}] \, g$$

II. $[\tau_i, \ldots, \tau_{i+k}] \, g$ is symmetric in its arguments $\tau_i, \ldots, \tau_{i+k}$ (since the interpolating polynomial depends only on the points of interpolation and not the order in which they are specified.)

III. $[\tau_i, \ldots, \tau_{i+k}] g$ is a linear operator; if we let

$$f = \alpha g + \beta h$$

for some functions g, h and scalars α, β, then

$$[\tau_i, \ldots, \tau_{i+k}] f = \alpha [\tau_i, \ldots, \tau_{i+k}] g + \beta [\tau_i, \ldots, \tau_{i+k}] h$$

IV. If $f = gh$, then

$$[\tau_i, \ldots, \tau_{i+k}] f = \sum_{j=i}^{i+k} [\tau_i, \ldots, \tau_{i+k}] g [\tau_i, \ldots, \tau_{i+k}] h$$

V. If g is a polynomial of degree $\leq k$, then $[\tau_i, \ldots, \tau_{i+k}] g$ is constant as a function of $\tau_i, \ldots, \tau_{i+k}$. In particular,

$$[\tau_i, \ldots, \tau_{i+k}] g = 0 \quad \forall g \in \mathcal{P}_k$$

The polynomial which agrees with g at $\tau_i, \ldots, \tau_{i+k}$ or any other $k+1$ points must be g itself due to the uniqueness of interpolating polynomials. For more details, we refer the reader to A Practical Guide to Splines, DeBoor.

VI. $[\tau_i, \ldots, \tau_{i+k}] g$ is a continuous function of its $k+1$ arguments in case $g \in \mathcal{C}^{(k)}$ (that is, if g has k continuous derivatives.)

VII. If $g \in \mathcal{C}^{(k)}$, then there exists a point ξ in the smallest interval containing $\tau_i, \ldots, \tau_{i+k}$ such that

$$[\tau_i, \ldots, \tau_{i+k}] g = \frac{g^{(k)}(\xi)}{k!}$$

VIII.

$$[\tau_i, \ldots, \tau_{i+k}] g = \begin{cases} \frac{g^{(k)}(\tau_i)}{k!} & \tau_i = \cdots = \tau_{i+k}, \quad g \in \mathcal{C}^{(k)} \\ \frac{[\tau_i, \ldots, \tau_{r-1}, \tau_{r+1}, \ldots, \tau_{i+k}] g - [\tau_i, \ldots, \tau_{s-1}, \tau_{s+1}, \ldots, \tau_{i+k}] g}{\tau_s - \tau_r} & \begin{array}{l} \tau_r, \tau_s \text{ are any two distinct} \\ \text{points in } \{\tau_i, \ldots, \tau_{i+k}\} \end{array} \end{cases}$$

IX. Let $\tau = \{\tau_i\}$, $i = 1, \ldots, n$ be a nondecreasing sequence, and let g be a "smooth enough" function. Define the restriction of g to the sequence τ to be the vector $g|_\tau = (g_1, \ldots, g_n)^T$, where

$$g_i = g^{(r)}(\tau_i), \quad r = \max\{j : \tau_{i-j} = \tau_i\}$$

For any $1 \leq r \leq s \leq n$, there exist constants $\{d_i\}$, $i = 1, \ldots, n$ which depend only on r, s, and τ and are free of g such that

$$[] g = \sum_i = 1^n d_i g_i \tag{11}$$

Proof. The proof is done by induction on $s - r$. The statement is obvious for the case where $s - r = 0$; all $d_i = 0$ except the constant corresponding to $g(\tau_r)$, which is equal to 1. For any $r < s$, if $\tau_r = \cdots = \tau_s$, then

$$[\tau_r, \ldots, \tau_s] g = \frac{g^{(s-r)}(\tau_r)}{(s-r)!}$$

so that 11 holds with a single $d_j = \frac{1}{(s-r)!}$ and all other $d_i = 0$. The index $j \leq s$ of the nonzero constant depends only on the number of τ_j with $j < r$ and $\tau_j = \tau_r$.

Assume that 11 holds for $s - r < k$. Consider $s - r = k$ with $\tau_r < \tau_s$. By the induction hypothesis, there are coefficients $\{d_i'\}$ and $\{d_i''\}$ □

In the following section, we will present B-splines as divided differences of the truncated power basis and present some of their properties which come as a result of the properties of the divided difference presented in this section.

4 A B-spline representation for pp functions

Definition 4.1. Let $t = \{t_i\}$ denote a non-decreasing sequence. The ith B-spline of order k which corresponds to the knot sequence t is defined by

$$B_{i,k,t}(x) = (t_{i+k} - t_i)[t_i, \ldots, t_{i+k}](\cdot - x)_+^{k-1} \tag{12}$$

The placeholder notation, $(\cdot - x)_+^{k-1}$, is used to indicate that the kth divided difference of the function $g(t) = (t - x)_+^{k-1}$ is obtained by fixing x and applying the divided difference to $g(t)$ as a function of t alone. Henceforth, we will write B_i rather than $B_{i,k,t}$ when the spline order and knot sequence can be inferred from surrounding context.

4.1 Properties of B-splines

I. $B_i(x)$ has isolated support:

$$B_i(x) = 0, \quad x \notin [t_i, t_{i+k}]$$

To see this, note that if $x \notin [t_i, t_{i+k}]$, then $g(t) = (t - x)_+^{k-1}$ is a polynomial of degree $< k$ on $[t_i, t_{i+k}]$, thus by 3.1.1 V,

$$[t_i, \ldots, t_{i+k}] g = 0.$$

As a result, for a set of B-splines of order k corresponding to the knot sequence t, only k of them are nonzero on $[t_j, t_{j+k}]$: $B_{j-k+1}, B_{j-k+2}, \ldots, B_j$.

II. The ith B-spline of order is defined as the kth divided difference of $(\cdot - x)_+^{k-1}$ times a normalization factor: $(t_{i+k} - t_i)$. This normalization, using 3.1.1 VIII, allows us to write

$$B_i\left(x\right) = [t_{i+1}, \ldots, t_{i+k}]\left(\cdot - x\right)_+^{k-1} - [t_i, \ldots, t_{i+k-1}]\left(\cdot - x\right)_+^{k-1} \tag{13}$$

For $x \in (t_j, t_{j+1})$, by 3.1.1 I,

$$\sum_i B_i\left(x\right) = \sum_{i=j+1-k}^{j} B_i\left(x\right)$$

$$= \sum_{i=j+1-k}^{j} [t_{i+1}, \ldots, t_{i+k}]\left(\cdot - x\right)_+^{k-1} - \sum_{i=j+1-k}^{j} [t_i, \ldots, t_{i+k-1}]\left(\cdot - x\right)_+^{k-1}$$

$$= [t_{j+1}, \ldots, t_{j+k}]\left(\cdot - x\right)_+^{k-1} - [t_{j+1-k}, \ldots, t_j]\left(\cdot - x\right)_+^{k-1}$$

$$= 1 - 0 \tag{14}$$

The last equality in 14 is a consequence of the following: for $x \in (t_j, t_{j+1})$, $g\left(t\right) = (t - x)_+^{k-1}$ is a $k - 1$ degree polynomial with unit leading coefficient on $[t_{j+1}, t_{j+k}]$, so by 3.1.1 V,

$$[t_{j+1}, \ldots, t_{j+k}] g = 1.$$

On $[t_{j+1-k}, t_j]$, g is identically 0, hence $[t_{j+1-k}, \ldots, t_j] g = 0$.

III. Each $B_i\left(x\right)$ is positive on its support. Applying Leibnitz's formula (3.1.1 IV) to the product

$$[t_i, \ldots, t_{i+k}]\left(t - x\right)_+^{k-1} = [t_i, \ldots, t_{i+k}]\left(t - x\right)\left(t - x\right)_+^{k-2},$$

we have

$$[t_i, \ldots, t_{i+k}]\left(t - x\right)_+^{k-1} = [t_i, \ldots, t_{i+k}]\left(t - x\right)\left(t - x\right)_+^{k-2}$$

$$= \sum_{r=i}^{i+k} [t_i, \ldots, t_{i+r}]\left(t - x\right)[t_r, \ldots, t_{i+k}]\left(t - x\right)_+^{k-2}$$

$$= \left[[t_i]\left(t - x\right) \right]\left[[t_i, \ldots, t_{i+k}]\left(t - x\right)_+^{k-2} \right]$$

$$+ \left[[t_i, t_{i+1}]\left(t - x\right) \right]\left[[t_{i+1}, \ldots, t_{i+k}]\left(t - x\right)_+^{k-2} \right]$$

$$= \left(t_i - x\right)[t_i, \ldots, t_{i+k}]\left(t - x\right)_+^{k-2}$$

$$+ 1 \cdot [t_{i+1}, \ldots, t_{i+k}]\left(t - x\right)_+^{k-2} \tag{15}$$

since $[t_i, \ldots, t_j]\left(\cdot - x\right) = 0$ for $j > i + 1$. By 3.1.1 VIII,

$$\left(t_i - x\right)[t_i, \ldots, t_{i+k}] g = \frac{t_i - x}{t_{i+k} - t_i}\left[[t_{i+1}, \ldots, t_{i+k}] g - [t_i, \ldots, t_{i+k-1}] g \right],$$

and we may express 15 as

$$[t_i, \ldots, t_{i+k}] \left(\cdot - x \right)_+^{k-1} = \frac{x - t_i}{t_{i+k} - t_i} [t_i, \ldots, t_{i+k-1}] \left(\cdot - x \right)_+^{k-2}$$
$$+ \frac{t_{i+k} - x}{t_{i+k} - t_i} [t_{i+1}, \ldots, t_{i+k}] \left(\cdot - x \right)_+^{k-2}$$

which we can write in terms of the normalized B-spline:

$$\frac{B_{i,k}(x)}{t_{i+k} - t_i} = \frac{x - t_i}{t_{i+k} - t_i} \frac{B_{i,k-1}(x)}{t_{i+k-1} - t_i} + \frac{t_{i+k} - x}{t_{i+k} - t_i} \frac{B_{i+1,k-1}(x)}{t_{i+k} - t_{i+1}} \qquad (16)$$

This shows that we can write the ith B-spline of order k as a convex combination of the ith and $(i+1)^{st}$ B-splines of order $k-1$ since

$$\frac{x - t_i}{t_{i+k} - t_i} + \frac{t_{i+k} - x}{t_{i+k} - t_i} = 1,$$

and each of these weights are positive for $t_i < x < t_{i+1}$. If

$$B_{j,k-1}(x) > 0, \quad t_j < x < t_{j+k-1} \text{ for all } j,$$

then by 16, we have that

$$B_{i,k}(x) > 0, \qquad t_i < x < t_{i+k}$$

since $B_{j,k-1} = 0$ for $x \notin [t_j, t_{j+k}]$ by 4.1 I and by induction over k, starting with the fact that

$$B_j, 1(x) = \begin{cases} 1 & t_j \le x < t_{j+1} \\ 0 & otherwise \end{cases}$$

Properties I, II, and III demonstrate that a sequence of B-splines form a *partition of unity*: a set of non-negative functions which sum, pointwise, to one.

Example 4.1. Figure 2 show fives parabolic B-splines corresponding to the set of knots $\{0, 1, 1, 3, 4, 6, 6, 6\}$. It is clear that each spline has compact support and is non-negative on this support. The function values are provided at select domain values so that property II is evident, though it is worth noting that

$$\sum_{i=1}^{6} B_i(x) = 1$$

only on $\left[\frac{1}{6}, 1 \right]$. In particular, they do not sum to 1 on $\left(0, \frac{1}{6} \right)$.

Each B_i is piecewise parabolic, with the breakpoints being locations of discontinuity of the function or one of its derivatives. B_5 is discontinuous at $x = 6$, as the knot at 6 is repeated three times in the knot sequence defining B_5: $\{t_5, t_6, t_7, t_8\}$.

B_1, B_2, and B_4 have discontinuous first derivatives since 1 is repeated twice in the knots defining B_1 and B_2, and 6 is repeated twice in the knots defining B_4. The relationship between knot replication and smoothness will be discussed in more detail in the following section.
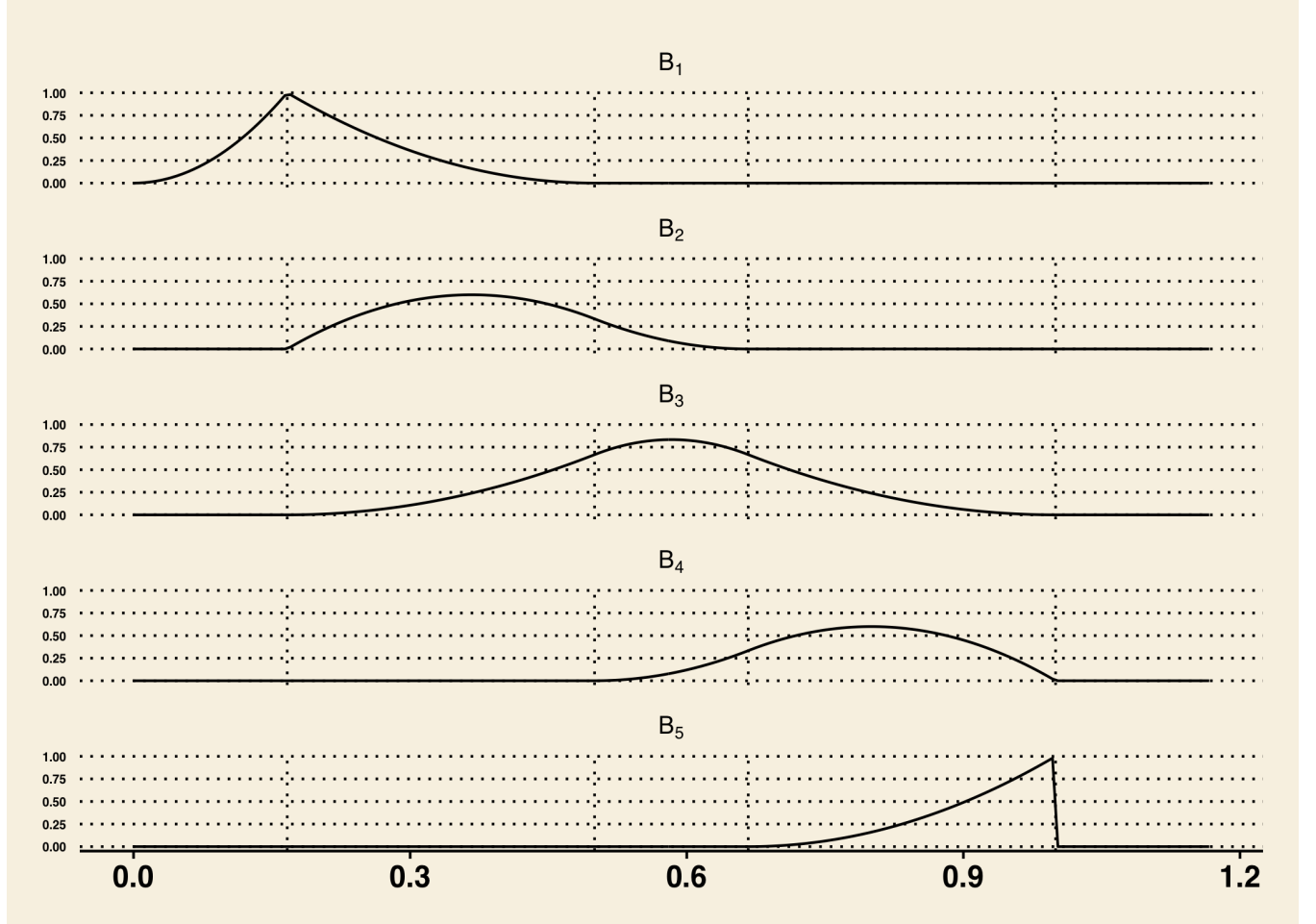
Figure 2: Parabolic B-splines corresponding to knot sequence $\{0, 1, 1, 3, 4, 6, 6, 6\}$, illustrating the connection between knot multiplicity and smoothness.

4.1.1 The Curry-Schoenberg Theorem

There is an extensive body of literature

Definition 4.2. A *spline function of order k with knot sequence t* is any linear combination of B-splines of order k for the knot sequence t. We denote this set of functions by

$$\mathcal{S}_{k,t} = \left\{ \sum_i \alpha_i B_{i,k,t} : \quad \alpha_i \in \Re \right\}$$

Theorem 4.1. *Curry-Schoenberg Theorem For a strictly increasing sequence $\xi = \{\xi_i\}$, $i = 1, \ldots, l+1$ and a sequence of non-negative integers $\nu = \{\nu_j\}$, $j = 2, \ldots, l$ with $\nu_j \leq k$*

14

for all j, define

$$n = k + \sum_{i=2}^{l} (k - \nu_i) = kl - \sum_{i=2}^{l} \nu_i$$

$$= \dim\left(\mathcal{P}_{k,\xi,\nu}\right)$$

and let $t = \{t_i\}$, $i = 1, \ldots, n + k$ be a non-decreasing sequence with

I.

$$t_1 \leq t_2 \leq \cdots \leq t_k \leq \xi_1$$
$$\xi_{l+1} \leq t_{n+1} \leq \cdots \leq t_{n+k}$$

II. *ξ_i occurs exactly $k - \nu_i$ times in t for $i = 2, \ldots, l$.*

Then, the sequence of B-splines of order k for the knot sequence t, $B_1, \ldots, B_n$, is a basis for $\mathcal{P}_{k,\xi,\nu}$ on the domain $[t_k, t_{n+1}]$.

Theorem 4.1 gives an explicit prescription for the construction of a B-spline basis for any particular pp space, $\mathcal{P}_{k,\xi,\nu}$, via the specification of t. These functions were named (B for 'basis') for this theorem. The choice of t controls the smoothness of the corresponding basis functions via knot multiplicities; the number of knots at a given breakpoint dictates the amount of smoothness at the breakpoint. Fewer knots placed at a breakpoint leads to more continuity conditions such that

$$[\# \text{ continuity conditions at } \xi] + [\# \text{ knots at } \xi] = k$$

A knot with multiplicity k yields a basis with no continuity conditions at that point, and a point where there are no knots forces k continuity conditions there.

4.1.2 The Curry-Schoenberg Proof

Proof. We begin by showing that $B_i \in \mathcal{P}_{k,\xi,\nu}$ for all i as functions defined on $[t_k, t_{n+1}]$. Each $B_i = B_{i,k,t}$ is defined as the kth divided difference of $(t - x)_+^{k-1}$ at $t_i, \ldots, t_{i+k}$ times a scalar, so by 3.1.1 IX, we can find coefficients $d_i, \ldots, d_{i+k}$ which depend only on $t_i, \ldots, t_{i+k}$ so that for any smooth function g,

$$[t_i, \ldots, t_{i+k}] g = \sum_{r=i}^{i+k} d_r g^{(j_r)}(t_r) \tag{17}$$

with

$$j_r = \max\left\{ s : \quad r - s \geq i \text{ and } t_{r-s} = t_r; \quad r = i, \ldots, i + k \right\}.$$

15

Using this, we may write

$$B_i\left(x\right) = \left(t_{i+k} - t_i\right) \sum_{r=i}^{i+k} d_r \left(t_r - x\right)_+^{k-j_r-1} \frac{\left(k-1\right)!}{\left(k - j_r - 1\right)!}. \tag{18}$$

From 18, it follows immediately that B_i is a pp function of order k with breakpoints $t_i, \ldots, t_{i+k}$ (and consequently some of the $\{\xi_i\}$.)

Now, we must establish the number of continuous derivatives of each B_i at each of its breakpoints, ξ_i, $i = 2, \ldots, l$. For any B_i, there cannot be a jump in its sth derivative across ξ_j unless

$$\xi_j = t_r \text{ and } k - 1 - j_r = s$$

for some $r \in \{i, \ldots, i+k\}$. Since

$$j_r = \#\ t_r = t_m\ i \le m < r,$$

j_r must be less than $k - \nu_j$, the total number of $\{t_m\}$ coinciding with ξ_j and hence equal to t_r, due to the construction of t. However, we must have $s \ge \nu_j$, and thus

$$D^m B_i\left(\xi_j^+\right) - D^m B_i\left(\xi_j^-\right) = 0, \qquad m = 0, \ldots, \nu_j - 1.$$

So $B_i \in \mathcal{P}_{k,\xi,\nu}$ for all $i = 1, \ldots, n$. We now only need to show that the B_i are linearly independent to complete the proof.

Lemma 4.2. *de Boor, Fix (1973) Define the linear functional λ_i by*

$$\lambda_i f = \sum_{r=0}^{k-1} \left(-1\right)^{k-r-1} \psi^{(k-r-1)}\left(\tau_i\right) D^r f\left(\tau_i\right)$$

where

$$\psi\left(t\right) = \frac{\left(t_{i+1} - t\right) \times \cdots \times \left(t_{i+1} - t\right)}{\left(k-1\right)!}$$

and where $\tau_i \in \left(t_i, t_{i+k}\right)$. Then

$$\lambda_i B_j = \delta_{ij} \quad \text{for all}\quad i, j.$$

Proof. It follows from its definition that $\lambda_i B_j$ is a pp function as a function of τ_i with breakpoints at the $\{t_i\}$. If we assume that $\tau_i \notin \{t_j\}$ for all i, then it is sufficient to show that $\lambda_i B_j = \delta_{ij}$.

By 17 and 18,

$$\lambda_i B_j = \left(t_{j+k} - t_j\right) \sum_{r=j}^{j+k} d_r \lambda_i \left[D_s^{j_r}\left(s - \cdot\right)_+^{k-1}\right]\big|_{s=t_r}$$

where $D_s^{j_r}$ denotes the operator for j_r-fold differentiation with respect to s. Now consider

$$\lambda_i \left(s - \cdot\right)_+^{k-1};$$

16

For $s < \tau_i$, $f(x) = (s - x)_+^{k-1}$ vanishes near τ_i, so that

$$\lambda_i f = 0.$$

For $s > \tau_i$, f agrees with $(s - x)^{k-1}$ in a neighborhood of τ_i, while

$$\lambda_i (s - x)^{k-1} = \sum_{r=0}^{k-1} (-1)^{k-r-1} \psi^{(k-r-1)} (\tau_i) (k - 1) \times \cdots \times (k - r) (-1)^r (s - \tau_i)^{k-r-1}$$

$$= \sum_{r=0}^{k-1} \frac{(k - 1)!}{(k - r - 1)!} \psi^{(k-r-1)} (\tau_i) (s - \tau_i)^{k-r-1} (-1)^{k-1}$$

$$= (-1)^{k-1} (k - 1)! \sum_{r=0}^{k-1} \frac{(s - \tau_i)^{k-r-1}}{(k - r - 1)!} \psi^{(k-r-1)} (\tau_i) \tag{19}$$

We recognize the sum in 19 as the truncated Taylor series in s of order k for ψ. Since ψ is itself a pp of order k, the sum must agree with ψ at s due to the uniqueness of interpolating polynomials. This implies that

$$\lambda_i (s - \cdot)^{k-1} = (-1)^{k-1} (k - 1)! \psi (s).$$

Accounting for the three cases, we have that

$$\lambda_i (s - \cdot)_+^{k-1} (-1)^{k-1} (k - 1)! \psi (s) (s - \tau_i)_+^0$$

since

$$D_s^r \lambda_i (s - \cdot)_+^{k-1} = \lambda_i \left[D_s^r (s - \cdot)_+^{k-1} \right].$$

Now we have that

$$\lambda_i B_j = (t_{j+k} - t_j) (-1)^{k-1} (k - 1)! [t_j, \ldots, t_{j+k}] \phi_i$$

where $\phi_i (s) = \psi (s) (s - \tau_i)_+^0$. Taking $[t_j, \ldots, t_{j+k}] \phi_i$ to be the leading coefficient of the order $k + 1$ polynomial which agrees with ϕ_i at $t_j, \ldots, t_{j+k}$, if we assume that

$$t_i < \tau_i < t_{i+k}$$

then

    I. ϕ_i agrees with ψ at $t_{i+1}, t_{i+2}, \ldots$, and as a polynomial of order $k + 1$, ψ has leading coefficient 0, so

$$[t_j, \ldots, t_{j+k}] \phi_i = 0, \qquad j = i + 1, i + 2, \ldots$$

    II. ϕ_i agrees with 0 at $t_{i+k-1}, t_{i+k-2}, \ldots$, so that

$$[t_j, \ldots, t_{j+k}] \phi_i = 0$$

III. ϕ_i agrees with the $(k+1)^{st}$ order polynomial

$$p\left(x\right) = \frac{\psi\left(x\right)\left(x - t_i\right)}{\left(t_{i+k} - t_i\right)}$$

at $t_i, \ldots, t_{i+k}$.

Together, I - III show that

$$\lambda_i B_j = \delta_{ij} \text{ for all } j.$$

$\square$

It is worth noting that we may apply Lemma 4.2 under the assumption that we can find τ_i in the open interval (t_i, t_{i+k}) - that is, if $t_i < t_{i+k}$ for all i. This case is of little interest, however; in the case that $t_i = t_{i+k}$, then $[t_i, t_{i+k}]$ is just a point. From 4.1 I, $B_i\left(x\right) = 0$ anywhere outside $[t_i, t_{i+k}]$, so if $t_i = t_{i+k}$, then it follows that B_i is simply the zero function. $\square$

The recipe provided by Theorem 4.1 for constructing the knot vector t such that the resulting $\{B_i\}$ form a basis for $\mathcal{P}_{k,\xi,\nu}$ leaves open the choice for the first k knots and the last k knots. It is conventional, due likely to convenience, to specify

$$t_1 = \cdots = t_k = \xi_1, \quad t_{n+1} = \cdots = t_{n+k} = \xi_{l+1},$$

and by setting $\nu_1 = \nu_{n+1} = 0$, one can include these knots according to the same method of choice of the other knots. This specification imposes no continuity conditions at the endpoints of the interval of interest; this is, however, consistent with the manner in which the B-splines are defined, spanning $\mathcal{P}_{k,\xi,\nu}$ only on $[t_k, t_{n+1}]$.

Definition 4.3. The *B-representation of* $f \in \mathcal{P}_{k,\xi,\nu}$ consists of

I. integers k and n specifying the order of f as a pp function and the number of linear parameters,

$$n = kl - \sum_i \nu_i = \dim\left(\mathcal{P}_{k,\xi,\nu}\right),$$

respectively.

II. The knot vector $t = \{t_i\}$, $i = 1, \ldots, n+k$ with elements arranged in increasing order, constructed according to Theorem 4.1, via ξ and ν.

III. The B-spline coefficients $\alpha = \{\alpha_i\}$, $i = 1, \ldots, n$ for the knot sequence, t.

Given I, II, and III in 4.3, the function value at $x \in [t_k, t_{n+1}]$ is given by

$$f\left(x\right) = \sum_{i=1}^{n} \alpha_i B_i\left(x\right),$$

18

and in particular, by I, for $x \in [t_j, t_{j+1}]$,

$$f(x) = \sum_{i=j}^{j+k-1} \alpha_i B_i(x).$$

Example 4.2 (Conversion between function representations)**.** Conversion from the B-representation to the pp-representation is straightforward the availability of stable evaluation and differentiation of the function. The $l+1$ distinct points among $t_k, \ldots, t_{n+1}$ are taken to be $\xi_1, \ldots, \xi_{n+1}$; the polynomial coefficients $\{c_{ji}\}$, $i = 1, \ldots, l$, $j = 1, \ldots, k$ are obtained by calculating $D^{j-1} f\left(\xi_i^+\right)$.

The conversion in the opposite direction, however, is much more difficult because there is no implicit information in the pp-representation about the smoothness of the function at its breakpoints, nor is it possible to numerically derive the minimal knot multiplicity necessary to represent the function as a spline using finite precision arithmetic. However, if it is known that f lies within $\mathcal{P}_{k,\xi,\nu}$ for a certain ν, then one may use Theorem 4.1 to construct t, and then α using Lemma 4.2. If such a representation for f is accessible, then the $\{\tau_i\}$ may be chosen to coincide with one of the $\{\xi_i\}$ so that the necessary derivatives may be extracted directly from the pp-representation.

Consider $f(x) = (x-3)(x-6)(x-9)$. Clearly f is a pp function of order $k = 4$. Suppose we wish to construct the B-representation of f on $[0, 10]$ as a cubic spline using interior knots $1, 2, \ldots, 9$. Then, we take

$$t_1 = t_2 = t_3 = t_4 = 0$$
$$t_{4+i} = i, \qquad\qquad\qquad\qquad i = 1, \ldots, 9$$
$$t_{13+1} = t_{13+2} = t_{13+3} = t_{13+4} = 10$$

Using Lemma 4.2, we may let

$$\tau_i = t_{i+2}$$

for all i, which gives

$$\alpha_i = f(t_{i+2}) + \frac{1}{3}\left(\Delta t_{i+2} - \Delta t_{i+1}\right) f'(t_{i+2}) - \frac{1}{3}\Delta t_{i+2}\Delta t_{i+1}\frac{f''(t_{i+2})}{2}$$

4.2   More B-spline Properties

The properties that follow illustrate how a spline $f \in (S)_{k,t}$ is related to its corresponding sequence of B-spline coefficients.

IV. **A recurrence relation for $B_{i,k}$:** The evaluation of a B-spline via its definition of a divided difference amounts to evaluating the function in terms of the truncated power basis. This approach is then plagued by the same computational issues that one hoped to sidestep with the introduction of the B-spline basis in the first place.

Via the derivation of 4.1 III, we showed that the B-splines satisfy the recurrence relation

$$B_{i,1}(x) = \begin{cases} 1, & t_i \leq x < t_{i+1} \\ 0, & otherwise \end{cases} \tag{20}$$

$$B_{i,k}(x) = \frac{x - t_i}{t_{i+k-1} - t_i} B_{i,k-1}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} B_{i+1,k-1}(x) \tag{21}$$

so that computing $B_{ik}(x)$ simply requires repeatedly taking positive linear combinations of positive quantities, starting with 20. Using this, one may express a series of B-splines as a series of lower order, but with coefficients that are polynomials in the spline argument. In general, we can write

$$\sum_i \alpha_i B_{i,k}(x) = \sum_i \alpha_i^{[j]}(x) B_{i,k-j}(x) \tag{22a}$$

$$\alpha_i^{[j]}(x) \equiv \begin{cases} \alpha_i, & j = 0 \\ \frac{(x-t_i)\alpha_i^{[j-1]}(x) + (t_{i+k-j} - x)\alpha_{i-1}^{[j-1]}(x)}{t_{i+k-j} - t_i}, & j \geq 1 \end{cases} \tag{22b}$$

and, in fact, $\alpha_i^{[k-1]}$ is a polynomial of degree $< k$ which agrees with $f = \sum_i \alpha_i B_{i,k}$ on $[t_i, t_{i+1}]$.

One may encounter 20 and 21 presented as the definition of the set of B-splines; properties 4.1 I and III follow immediately from 20 and 21, while II, is shown by induction. However, the smoothness properties of the $\{B_i\}$ presented in Theorem 4.1 are not as obvious from 20 and 21 as is their derivation from the properties of the divided difference, 3.1.1 I and 3.1.1 VI. We might even say that it is unexpected that the combination of $B_{i,k}$ and $B_{i,k-1}$ as given in 21 produces a function which has one more continuous derivative than either of the functions themselves.

V. **Differentiation of** $f \in \mathcal{S}_{k,t}$ A B-spline series may be differentiated by differencing the coefficients. (This fact will prove very useful in sections to follow.) Since the derivative of the truncated power function $g(x) = (t - x)_+^{k-1}$ is given by

$$D_x g(x) = D_x (t - x)_+^{k-1} = -(k-1)(t-x)_+^{k-2}$$

and by 13 and 17, we may write the derivative of the ith order-k B-spline as follows:

$$DB_{i,k}(x) = \left[[t_{i+1}, \ldots, t_{i+k}] - [t_i, \ldots, t_{i+k-1}] \right] D(\cdot - x)_+^{k-1}$$

$$= -(k-1)\left[[t_{i+1}, \ldots, t_{i+k}] - [t_i, \ldots, t_{i+k-1}] \right] (\cdot - x)_+^{k-2}$$

$$= -(k-1)\left[-\frac{B_{i+1,k-1}(x)}{(t_{i+k} - t_{i+1})} + \frac{B_{i,k-1}(x)}{(t_{i+k-1} - t_i)} \right]$$

20

This allows us to write

$$D_x\left[\sum_i \alpha_i B_i(x)\right] = \sum_i \alpha_i DB_{i,k}(x)$$

$$= \sum_i (k-1)\frac{\alpha_i - \alpha_{i-1}}{t_{i+k-1} - t_i}B_{i,k-1}(x) \qquad (23)$$

Note that the limits on the previous summation in 23 are left unspecified; the formula is written for biinfinite sums, and their application to finite sums is accessible after they are written formally as biinfinite sums by augmenting the appropriate zero terms.

However, if we are interested in a particular interval over the domain, say $[t_r, t_s]$, then for $x \in [t_r, t_s]$,

$$D\left[\sum_i \alpha_i B_{i,k}(x)\right] = \sum_{r-k+2}^{s-1} (k-1)\frac{\alpha_i - \alpha_{i-1}}{t_{i+k-1} - t_i}B_{i,k-1}(x)$$

since $B_{i,k-1}(x) = 0$ for all $i \notin \{r-k+2, \ldots, s-1\}$ when $t_r \le x \le t_s$. Applying 23 j times gives the jth derivative of $f \in \mathcal{S}_{k,t}$, which has form

$$D^j\left[\sum_i \alpha_i B_{i,k}(x)\right] = \sum_i \alpha_i^{(j+1)} B_{i,k-j} \qquad (24a)$$

$$\alpha_i^{(j+1)} \equiv \begin{cases} \alpha_i, & j = 0 \\ \dfrac{\alpha_i^{(j)} - \alpha_{i-1}^{(j)}}{\left(t_{i+k-j} - t_i\right)/(k-j)}, & j \ge 1 \end{cases} \qquad (24b)$$

Proof. We proceed by induction on j. We have already shown the case for $j = 1$ in the derivation of 23. Assume that the statement holds for some $j^* > 1$, so that we have

$$D^{j^*}\left[\sum_i \alpha_i B_{i,k}(x)\right] = \sum_i \frac{\alpha_i^{(j^*)} - \alpha_{i-1}^{(j^*)}}{\left(t_{i+k-j^*} - t_i\right)/(k-j^*)}B_{i,k-j^*}(x).$$

Then the $(j^* + 1)^{st}$ derivative is given by

$$D^{j^*+1}\left[\sum_i \alpha_i B_{i,k}\right] = \sum_i \frac{\alpha_i^{(j^*)} - \alpha_{i-1}^{(j^*)}}{(t_{i+k-j^*} - t_i)/(k - j^*)} DB_{i,k-j^*}$$

$$= \sum_i \alpha_i^{(j^*)} DB_{i,k-j^*}$$

$$= \sum_i \alpha_i^{(j^*)} (k - (j^* + 1)) \left[\frac{B_{i,k-(j^*+1)}}{t_{i+k-(j^*+1)} - t_i} - \frac{B_{i+1,k-(j^*+1)}}{t_{i+k-(j^*+1)+1} - t_{i+1}}\right]$$

$$= \sum_i \frac{\alpha_i^{(j^*)} - \alpha_{i-1}^{(j^*)}}{\left(t_{i+k-(j^*+1)} - t_i\right)/(k - (j^* + 1))} B_{i,k-(j^*+1)}$$

$$= \sum_i \alpha_i^{(j^*+1)} B_{i,k-(j^*+1)}$$

$\square$

The choice to write $k - j$ as a divisor in the denominator lends to the interpretation of 24a as a difference quotient, with the quantity

$$\frac{t_{i+k-j} - t_i}{k - j}$$

representing a mean mesh length of sorts on the interval $[t_i, t_{i+k-j}]$. We note that the case where t contains replicated knots leads to division by zero. This is, however, a trivial situation, since for $t_i = t_{i+k-j}$, we have $B_i = 0$, and we take $\frac{0}{0} = 0$.

VI. **The condition of the B-spline basis** To characterize the condition of the B-spline basis, we must bound the spline in terms of its B-spline coefficients and, additionally, bound the B-spline coefficients in terms of the spline. The first bound is easy to find; we need only find constants $m > 0$ and M such that

$$m \max_i |\alpha_i| \le m \max_t |\sum_i \alpha_i B_{i,k}(t)| \le M \max_i |\alpha_i|$$

holds for any coefficient vector $\alpha = \{\alpha_i\}$. Since $\{B_{i,k}\}$ form a partition of unity, we have

$$|\sum_i \alpha_i B_{i,k}(t)| \le \sum_i |\alpha_i| B_{i,k}(t) \le \sum_i \max_i |\alpha_i| B_{i,k}(t) = \max_i |\alpha_i|.$$

So for $M = 1$, the second inequality holds. To show that the first inequality holds, let $f = \sum_i \alpha_i B_{i,k}$. From Theorem 4.1, we know that we may write

$$\alpha_i = \lambda_{ik} f = \sum_{j=1}^{k} \frac{-D^{j-1}\psi_{ik}(\tau)}{(k-1)!} D^{k-j} f(\tau) \qquad (25)$$

22

for some $\tau \in [t_i, t_{i+k})$. To bound this sum in terms of $\max_t |f(t)|$, suppose that

$$\tau \in [t_l, t_{l+1}) \subset [t_i, t_{i+k}).$$

Then for some constant c_k depending on k only, and for any pp function p of order k,

$$|D^j p(\tau)| \leq c_k (\Delta t_l)^{-1} \max_{t_l \leq t \leq t_{l+1}} |p(t)|. \tag{26}$$

For the case in which $\Delta t_l = 1$, the existence of c_k follows from the fact that the space of pp functions of order k is finite dimensional; from this, the statement follows for arbitrary Δt_l via scaling. Since f agrees with some polynomial of degree $< k$ on $[t_l, t_{l+1})$,

$$|D^j f(\tau)| \leq c_k (\Delta t_l)^{-1} \max_{t_i \leq t \leq t_{i+k}} |f(t)| \tag{27}$$

However, $\psi_{ik} = (t_{i+1} - \cdot)(t_{i+2} - \cdot) \ldots (t_{i+k-1} - \cdot)$ is also a polynomial of degree $< k$, and

$$\max_{t_l \leq t \leq t_{l+1}} \psi(t) \leq c_k^* |\Delta t_{l*}|^{k-1} \tag{28}$$

for some c_k^* which depends on k only and $[t_{l*}, t_{l*+1})$ a largest interval of that form in $[t_{l*}, t_{l*+1})$. Taking $l^* = l$, from 26 with $p = \psi_{ik}$ and from 27 we obtain the bound

$$|D^{\nu-1} \psi(\tau) D^{k-\nu} s(\tau)| \leq (c_k)^2 c_k^* \max_{t_i \leq t \leq t_{i+k}} |f(t)|.$$

If we sum these bounds over ν and divide by $(k-1)!$, we have

$$|\alpha_i| = |\lambda_{ik} f| \leq c \max_{t_i \leq t \leq t_{i+k}} |f(t)|,$$

with c depending only on k.

Properties 4.2 V and VI demonstrate how closely a spline function is modeled by its B-spline coefficients. In particular, the formula for differentiation of a spline function given by V tells us that derivatives of a spline function may be obtained by simply differencing its coefficients. This fact will motivate the difference penalties discussed in section P-spline penalties to follow.

5 Single-regressor varying coefficient models via B-spline basis expansions

Hastie and Tibshirani were the first to introduce the varying coefficient model; consider data of the form

$$(x_i, y_i, t_i), \qquad i = 1, \ldots, m$$

where y_i is the response, x_i is the single (univariate) regressor variable, and t_i is an indexing variable. The single regressor varying-coefficient (VC) model extends the classical linear model by allowing the slope coefficient to vary smoothly in the dimension of the indexing variable, t. The VC model assumes that the mean response is of the form

$$E\left[Y\left(t\right)\right] = f\left(t\right)x\left(t\right)$$

where $f\left(t\right)$ is the smooth slope function of interest. This modeling approach permits interpolation x and y at values of the indexing variable t at which there is either missing data of only a single observation and slope estimation is not feasible. In the section that follows, we will discuss the approach to smoothing the coefficient vector (and *not* the regression function $x\left(t\right)$) first, for mechanical demonstration of parameterization and estimation of the coefficient function via B-spline basis expansion, at a predetermined set of values of t (knots), then following the approach of Eilers and Marx by assuming that the number and position of the knots are unknown and using penalized B-splines, or P-splines

6 Tensor product B-splines for multidimensional surface approximation