§7.2 (Natural) Cubic Splines

If you recall the definition of the linear and quadratic splines, probably you can guess the definition of the spline of degree k:

Definition 1 (Splines of Degree k). A function S is a spline of degree 1 on $[a, b]$ if

1. The domain of S is $[a, b]$.
2. $S, S', S'', \ldots, S^{(k-1)}$ are continuous on $[a, b]$.
3. There is a partition $a = t_0 < t_1 < \ldots < t_n = b$ such that on $[t_i, t_{i+1}]$, S is a polynomial of degree $\leq k$.

You would also expect that a spline of degree k has $k - 1$ degrees of freedom. That is if $n + 1$ knots are given, the spline of degree k is defined by $n(k + 1)$ parameters. The given data

t	t_0	t_1	$\ldots$	t_n
y	y_0	y_1	$\ldots$	y_n

provide $2n$ equations. The continuity of $S', S'', \ldots, S^{(k-1)}$ at the $n - 1$ internal knots gives $(k - 1)(n - 1)$ equations. This is a total of $n(k + 1) - (k - 1)$ equations. Thus we have $k - 1$ more unknowns than equations. Thus, barring some singularity, we can (and must) add $k - 1$ constraints to uniquely define the spline.

Often k is chosen as 3. This yields cubic splines. We must add 2 extra constraints to define the spline. The usual choice is to make

$$S''(t_0) = S''(t_n) = 0.$$

This yields the *natural cubic spline*.

Why Natural Cubic Splines?

It turns out that natural cubic splines are a good choice in the sense that they are the "interpolant of minimal H^2 seminorm." The corollary following this theorem states this in more easily understandable terms:

Theorem 2. Suppose f has two continuous derivatives, and S is the natural cubic spline interpolating f as knots $a = t_0 < t_1 < \ldots < t_n = b$. Then

$$\int_a^b \left[S''(x)\right]^2 \mathrm{d}x \leq \int_a^b \left[f''(x)\right]^2 \mathrm{d}x$$

Proof. We let $g(x) = f(x) - S(x)$. Then $g(x)$ is zero on the $(n+1)$ knots t_i. Derivatives are linear, meaning that

$$f''(x) = S''(x) + g''(x).$$

Then

$$\int_a^b \left[f''(x)\right]^2 \mathrm{d}x = \int_a^b \left[S''(x)\right]^2 \mathrm{d}x + \int_a^b \left[g''(x)\right]^2 \mathrm{d}x + \int_a^b 2S''(x)g''(x)\mathrm{d}x.$$

We show that the last integral is zero. Integrating by parts we get

$$\int_a^b S''(x)g''(x)\mathrm{d}x = S''g'\Big|_a^b - \int_a^b S'''g'\mathrm{d}x = -\int_a^b S'''g'\mathrm{d}x,$$

because $S''(a) = S''(b) = 0$. Then notice that S is a polynomial of degree ≤ 3 on each interval, thus $S'''(x)$ is a piecewise constant function, taking value c_i on each interval $[t_i, t_{i+1}]$. Thus

$$\int_a^b S'''g'\mathrm{d}x = \sum_{i=0}^{n-1} \int_a^b c_i g'\mathrm{d}x = \sum_{i=0}^{n-1} c_i g\Big|_{t_i}^{t_{i+1}} = 0,$$

with the last equality holding because $g(x)$ is zero at the knots. □

Corollary 3. The natural cubic spline is best twice-continuously differentiable interpolant for a twice-continuously differentiable function, under the measure given by the theorem.

Proof. Let f be twice-continuously differentiable, and let S be the natural cubic spline interpolating $f(x)$ at some given nodes $\{t_i\}_{i=0}^n$. Let $R(x)$ be some twice-continuously differentiable function which also interpolates $f(x)$ at these nodes. Then $S(x)$ interpolates $R(x)$ at these nodes. Apply the theorem to get

$$\int_a^b \left[S''(x)\right]^2 \mathrm{d}x \leq \int_a^b \left[R''(x)\right]^2 \mathrm{d}x$$

□

An Example

We construct the natural cubic spline for the following data:

t	-1	0	2
y	3	-1	3

The natural cubic spline is defined by eight parameters:

$$S(x) = \begin{cases} ax^3 + bx^2 + cx + d & x \in [-1, 0] \\ ex^3 + fx^2 + gx + h & x \in [0, 2] \end{cases}$$

We interpolate to find that $d = h = -1$ and

$$\begin{aligned} -a + b - c - 1 &= 3 \\ 8e + 4f + 2g - 1 &= 3 \end{aligned}$$

We take the derivative of S:

$$S'(x) = \begin{cases} 3ax^2 + 2bx + c & x \in [-1, 0] \\ 3ex^2 + 2fx + g & x \in [0, 2] \end{cases}$$

Continuity at the middle node gives $c = g$.

We take the second derivative of S:

$$S''(x) = \begin{cases} 6ax + 2b & x \in [-1, 0] \\ 6ex + 2f & x \in [0, 2] \end{cases}$$

Continuity at the middle node gives $b = f$. The natural cubic spline condition gives $-6a + 2b = 0$ and $12e + 2f = 0$. Solving this by "divide and conquer" gives

$$S(x) = \begin{cases} x^3 + 3x^2 - 2x - 1 & x \in [-1, 0] \\ -\frac{1}{2}x^3 + 3x^2 - 2x - 1 & x \in [0, 2] \end{cases}$$

Computing Natural Cubic Splines

Finding the constants for the previous example was fairly tedious. And this is for the case of only three nodes. We would like a method easier than setting up the $4n$ equations and unknowns.

The book does a fairly good job of describing the algorithm. I see little value in covering this in detail in class because its much more drudgework than ideas.

Here's the basic description: You want to find the natural cubic spline, $S(x)$ interpolating the data:

t	t_0	t_1	$\ldots$	t_n
y	y_0	y_1	$\ldots$	y_n

Define $z_i = S''(t_i)$. Other than $z_0 = z_n = 0$, the z_i are not known up front and need to be determined. Not surprisingly, you can write $S_i(x)$ as some cubic polynomial involving z_i, z_{i+1} and two constants yet to be determined. When you use the continuity of S, S' at the internal nodes, the two constants for each S_i are determined to be related to the data t_i, y_i in some detailed way. This leads to a symmetric tridiagonal system:

$$\begin{bmatrix} u_1 & h_1 & 0 & \cdots & 0 \\ h_1 & u_2 & h_2 & \cdots & 0 \\ 0 & h_2 & u_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{n-1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-1} \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \end{bmatrix} \tag{1}$$

The h_i's and u_i's are easy to describe–they are related to the width of the intervals:

$$h_i = t_{i+1} - t_i, \quad u_i = 2\left(h_{i-1} + h_i\right).$$

The v_i's are not so easy to explain, and involve the y_i as well.

To solve this system there are fancy methods for tridiagonal systems. It turns out that naïve Gaussian Elimination works just fine in this case, however.

§7.3, 7.4 B Splines

The chapter on cubic splines was weak. I'll talk a little bit about B splines. This is not on the final exam or the quizzes. It is interesting and useful, however, so you should pay attention.

The B splines form a *basis* for spline functions, which is where the name comes from. We presuppose the existence of an infinite number of knots:

$$\ldots < t_2 < t_1 < t_0 < t_1 < t_2 < \ldots,$$

with $\lim_{k\to-\infty} t_k = -\infty$ and $\lim_{k\to\infty} t_k = \infty$.

The B splines of degree 0 are defined as single "blocks":

$$B_i^0(x) = \begin{cases} 1 & t_i \le x < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

The zero degree B splines are continuous from the right, are nonzero only on one subinterval $[t_i, t_{i+1})$, sum to 1 everywhere.

We justify the description of B splines as basis splines: If S is a spline of degree 0 on the given knots and is continuous from the right then

$$S(x) = \sum_i S(x_i) B_i^0(x).$$

Amazing.

Then we can define the B splines of degree k recursively:

$$B_i^k(x) = \left(\frac{x - t_i}{t_{i+k} - t_i}\right) B_i^{k-1}(x) + \left(\frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}}\right) B_{i+1}^{k-1}(x).$$

These quickly become unweildy. We focus on the case $k = 1$. The B spline $B_i^1(x)$ is

- Piecewise linear.
- Continuous.
- Nonzero only on (t_i, t_{i+2}).
- 1 at t_{i+1}.

They're sometimes called *hat* functions. Imagine wearing a hat shaped like this! Whatever.

The nice thing about the hat functions is they allow us to use analogy. Harken back to polynomial interpolation and the Lagrange Functions. The hat functions play a similar role because

$$B_i^1(t_j) = \delta_{(i+1)j} = \begin{cases} 1 & (i+1) = j \\ 0 & (i+1) \ne j \end{cases}$$

Then if we want to interpolate the following data with splines of degree 1:

t	t_0	t_1	$\dots$	t_n
y	y_0	y_1	$\dots$	y_n

We can immediately set

$$S(x) = \sum_{i=0}^{n} y_i B_{i-1}^1(x).$$