A. Ralston, Ph. Rabinovich
A First Course in Numerical Analysis

3.8 SPLINE INTERPOLATION

As mentioned in Sec. 3.4, it is possible that a sequence of interpolation polynomials $\{p_n(x)\}$ over a fixed finite interval need not converge even to a smooth function. And if we investigate the behavior of these polynomials between the interpolation points we find that in many cases the polynomials oscillate quite violently while the function varies smoothly. And the higher the degree of the polynomial, the worse the situation becomes. A way to overcome this problem is by using piecewise low-order interpolating polynomials on subintervals of the given interval. With them, the oscillation between points is not significant, so that they can imitate the behavior of the function. However, the resulting function pieced together from the individual low-order polynomials may not be smooth. Since we wish to imitate the behavior of smooth functions, a requirement for these piecewise functions is that the resulting function pieced together be smooth. Such functions, with the maximal degree of smoothness, are called *splines* and we shall now describe them formally.

Let the interval $I = [a, b]$ be divided into $n - 1$ subintervals $a = a_1 < a_2 < \cdots < a_{n-1} < a_n = b$ not necessarily of equal length. A *spline* $S(x)$ of degree m is a function defined on I which:

1. Coincides with a polynomial of degree m on each subinterval $I_i = [a_{i-1}, a_i]$, $i = 2, \ldots, n$.
2. Has continuous derivatives up to order $m - 1$.

The abscissas $\{a_i\}$ are called the *nodes* or *knots* of the spline. A spline $S(x)$ is said to interpolate to the data points (a_i, y_i) if $S(a_i) = y_i$, $i = 1, \ldots, n$.

The word spline derives from the instrument often used by draftsmen in fairing a curve through data points. The simplest spline, that of degree 1, is a piecewise linear function which is not very smooth but very useful if the spacing between nodes is small. In fact, every table of functional values in which linear interpolation is used leads to an approximation of the underlying function by a linear spline. Splines of degree 2 can be defined, but since there is only one degree of freedom in their definition, there is a lack of symmetry in their determination with relation to the endpoints of the interval. Furthermore, the resulting functions are not sufficiently smooth. Thus the most prevalent spline in use is the cubic spline, which involves two parameters chosen to reflect the behavior at the endpoints of the interval. One type of end condition is that $S''(a) = S''(b) = 0$. The cubic spline which satisfies these conditions is called the *natural cubic spline*. A second condition, which generally yields better results, is

$$S'(a) = f'(a) \qquad S'(b) = f'(b)$$

This, of course, requires knowledge of the derivative at the endpoints.

Since we shall subsequently be dealing exclusively with cubic interpolatory splines, we shall drop the adjectives and simply write spline. One of several representations of splines can be derived as follows. Set

$$h_i = a_i - a_{i-1} \qquad i = 2, 3, \ldots, n \tag{3.8-1}$$

Since $S(x)$ is piecewise cubic, $S'(x)$ is piecewise quadratic and $S''(x)$ is piecewise linear and continuous. Hence, we can write

$$S''(x) = M_{i-1}\frac{a_i - x}{h_i} + M_i\frac{x - a_{i-1}}{h_i} \qquad \text{on } I_i \tag{3.8-2}$$

for certain constants M_i, where in fact

$$S''(a_i) = M_i \qquad i = 1, 2, \ldots, n \tag{3.8-3}$$

Integrating (3.8-2) twice and writing the arbitrary linear function in the form indicated, we obtain

$$S(x) = M_{i-1}\frac{(a_i - x)^3}{6h_i} + M_i\frac{(x - a_{i-1})^3}{6h_i} + c_i(a_i - x) + d_i(x - a_{i-1}) \tag{3.8-4}$$

on I_i. Since we wish the spline to interpolate at the knots, we have that $S(a_{i-1}) = y_{i-1}$ and $S(a_i) = y_i$. This determines the c_i and d_i, yielding

$$\begin{aligned} S(x) &= M_{i-1}\frac{(a_i - x)^3}{6h_i} + M_i\frac{(x - a_{i-1})^3}{6h_i} \\ &+ \left(y_{i-1} - \frac{M_{i-1}h_i^2}{6}\right)\frac{a_i - x}{h_i} \\ &+ \left(y_i - \frac{M_i h_i^2}{6}\right)\frac{x - a_{i-1}}{h_i} \end{aligned} \tag{3.8-5}$$

on I_i. Differentiating (3.8-5), we obtain

$$S'(x) = -M_{i-1}\frac{(a_i - x)^2}{2h_i} + M_i\frac{(x - a_{i-1})^2}{2h_i} + \frac{y_i - y_{i-1}}{h_i} - \frac{M_i - M_{i-1}}{6}h_i \tag{3.8-6}$$

on I_i. We note the particular values

$$\begin{aligned} S'(a_i^-) &= \frac{h_i}{6}M_{i-1} + \frac{h_i}{3}M_i + \frac{y_i - y_{i-1}}{h_i} \\ S'(a_i^+) &= -\frac{h_{i+1}}{3}M_i - \frac{h_{i+1}}{6}M_{i+1} + \frac{y_{i+1} - y_i}{h_{i+1}} \end{aligned} \tag{3.8-7}$$

Since $S'(x)$ is required to be continuous, these two values must be equal; this yields the equations

$$\frac{h_i}{6}M_{i-1} + \frac{h_i + h_{i+1}}{3}M_i + \frac{h_{i+1}}{6}M_{i+1} = \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \qquad i = 2, 3, \ldots, n-1 \tag{3.8-8}$$

These form a set of $n - 2$ linear equations in $M_1, \ldots, M_n$, so that two more conditions must be added. Once the M's are determined, the interpolation spline is completely determined through (3.8-5). We shall abbreviate Eqs. (3.8-8) by setting

$$\begin{aligned} \sigma_i &= \frac{y_i - y_{i-1}}{h_i} \qquad i = 2, 3, \ldots, n \\ \lambda_i &= \frac{h_{i+1}}{h_i + h_{i+1}} \qquad \mu_i = 1 - \lambda_i \qquad d_i = \frac{6(\sigma_{i+1} - \sigma_i)}{h_i + h_{i+1}} \\ & \qquad i = 2, 3, \ldots, n-1 \end{aligned} \tag{3.8-9}$$

We then get the set of equations

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i \qquad i = 2, 3, \ldots, n-1 \tag{3.8-10}$$

We shall write the two additional conditions in the form

$$2M_1 + \lambda_1 M_2 = d_1 \qquad \mu_n M_{n-1} + 2M_n = d_n \tag{3.8-11}$$

and indicate below several possible choices of the constants $\lambda_1, d_1, \mu_n, d_n$. The combined system now becomes

$$\begin{bmatrix} 2 & \lambda_1 & 0 & & & & \\ \mu_2 & 2 & \lambda_2 & & \bigcirc & & \\ 0 & \mu_3 & 2 & & & & \\ & & & \ddots & & & \\ & & & & 2 & \lambda_{n-2} & 0 \\ & \bigcirc & & & \mu_{n-1} & 2 & \lambda_{n-1} \\ & & & & 0 & \mu_n & 2 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ M_{n-2} \\ M_{n-1} \\ M_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-2} \\ d_{n-1} \\ d_n \end{bmatrix} \tag{3.8-12}$$

with a tridiagonal matrix.

1. If one selects $\lambda_1 = d_1 = \mu_n = d_n = 0$, then $M_1 = M_n = 0$. This yields the natural spline as defined above.
2. Using the second endpoint condition proposed above

$$S'(a) = y'_1 \qquad S'(b) = y'_n \tag{3.8-13}$$

we obtain

$$\lambda_1 = 1 = \mu_n \qquad d_1 = \frac{6}{h_2}\left(\frac{y_2 - y_1}{h_2} - y'_1\right)$$

$$d_n = \frac{6}{h_n}\left(y'_n - \frac{y_n - y_{n-1}}{h_n}\right) \tag{3.8-14}$$

3. A third possibility is to choose a_i, $i = 2, \ldots, n-1$, as the spline nodes, i.e., not require the endpoints to be nodes, and to impose the conditions $S(a_1) = y_1$ and $S(a_n) = y_n$.

In this case, the index i in (3.8-10) runs from 3 to $n-2$, and the two additional conditions are similar in form to those of (3.8-11). The values of λ_2 and d_2 can readily be evaluated by letting $x = a_1$ in (3.8-5) and setting $S(a_1) = y_1$. Similarly, by setting $x = a_n$, we can find the values of λ_{n-1} and d_{n-1} {30}. This scheme, which does not require any additional information, appears to be better than choice 1 for the average function that comes up in practice since the second derivative does not generally vanish at the endpoints of the interval.

From (3.8-9) we see that $0 < \lambda_i < 1$ and $0 < \mu_i < 1$ for $i = 2, \ldots, n-1$. If, therefore, $|\lambda_1| < 2$, $|\mu_n| < 2$, the matrix in (3.8-12) will be diagonally dominant. In this case, it can be shown (see Sec. 9.1) that unique solutions to (3.8-12) will exist for arbitrary $d_1, \ldots, d_n$. Thus, in cases 1 and 2, we are assured of the existence of the spline. Similarly, in case 3, a solution always exists {30}.

Splines have the following important properties which can be readily proved {31, 32}.

1. Given data points (a_i, y_i), $i = 1, \ldots, n$. Of all the functions $f(x)$ with continuous second derivatives which interpolate to these data the spline $S(x)$ which also satisfies $S''(a) = S''(b) = 0$ uniquely minimizes the integral

$$E(g) = \int_a^b [g''(x)]^2\, dx \tag{3.8-15}$$

Similarly, of all the functions $f(x)$ with continuous second derivatives which interpolate to these data and satisfy $f'(a) = y'_1, f'(b) = y'_n$ the spline $S(x)$ which also satisfies (3.8-13) uniquely minimizes (3.8-15).

2. If we define $h = \max_i h_i$ and let $S(x)$ be the natural spline interpolating $f(x)$ at a_i, $i = 1, \ldots, n$, where $f(x)$ has a continuous second derivative, then

$$\max_{a \le x \le b} |f(x) - S(x)| \le h[hE(f)]^{1/2} \tag{3.8-16}$$

$$\max_{a \le x \le b} |f'(x) - S'(x)| \le [hE(f)]^{1/2} \tag{3.8-17}$$

A similar theorem holds for the spline $S(x)$ satisfying $S'(a) = f'(a)$, $S'(b) = f'(b)$.

A stronger, albeit only asymptotic, result states that if $f(x)$ has a continuous fourth derivative, and if $\max_i (h/h_i) \le \beta < \infty$ as $h \to 0$ for a fixed β, then

$$\max_{a \le x \le b} |f^{(k)}(x) - S^{(k)}(x)| = O(h^{4-k}) \qquad k = 0, 1, 2 \tag{3.8-18}$$

Example 3.8 Determine the natural cubic spline $S(x)$ which interpolates to the values of y_i at the points a_i, $i = 1, \ldots, 5$, where

$a_i = .25$	.30	.39	.45	.53
$y_i = .5000$	.5477	.6245	.6708	.7280

We have that $h_2 = .05$, $h_3 = .09$, $h_4 = .06$, $h_5 = .08$, so that using (3.8-9) leads to

$$\lambda_2 = \tfrac{9}{14} \quad \mu_2 = \tfrac{5}{14} \quad \lambda_3 = \tfrac{2}{5} \quad \mu_3 = \tfrac{3}{5} \quad \lambda_4 = \tfrac{4}{7} \quad \mu_4 = \tfrac{3}{7}$$

$$\sigma_2 = .9540 \quad \sigma_3 = .8533 \quad \sigma_4 = .7717 \quad \sigma_5 = .7150$$

$$d_2 = -4.3157 \quad d_3 = -3.2640 \quad d_4 = 2.4300$$

Inserting these values into (3.8-10), we get

$$\begin{aligned} \tfrac{5}{14}M_1 + 2M_2 + \tfrac{9}{14}M_3 \qquad\qquad &= -4.3157 \\ \tfrac{3}{5}M_2 + 2M_3 + \tfrac{2}{5}M_4 \qquad &= -3.2640 \\ \tfrac{3}{7}M_3 + 2M_4 + \tfrac{4}{7}M_5 &= -2.4300 \end{aligned}$$

Since we wish to find a natural spline, we have that $M_1 = M_5 = 0$, so that we are left with a tridiagonal system of three equations. Solving this system by the algorithm given in Sec. 9.11 [cf. (9.11-8) to (9.11-10)], we find that

$$M_2 = -1.8806 \qquad M_3 = -.8226 \qquad M_4 = -1.0261$$