Numerical Analysis

Grinshpan

Complete Cubic Spline

Let $x_1, \dots, x_n$ be strictly increasing, $a = x_1$, $b = x_n$, and let $y(x)$ be a differentiable function on $[a, b]$ with values $y(x_k) = y_k$. We will construct a function $s(x)$ with the following properties:

1. $s(x_k) = y(x_k)$, $k = 1, \dots, n$,
2. $s'(a) = y'(a)$, $s'(b) = y'(b)$, (clamped boundary conditions)
3. $s(x)$ is twice continuously differentiable on $[a, b]$,
4. $s(x)$ on $[x_k, x_{k+1}]$ is a cubic, $k = 1, \dots, n-1$.

There is a unique function $s(x)$ with the above properties. It is called the complete cubic spline interpolant of $y(x)$.

Construction of $s(x)$

The construction goes along the same lines as for the natural cubic spline. For each $k = 1, \dots, n-1$, define

$$s(x)\Big|_{[x_k, x_{k+1}]} = \frac{M_k}{6\delta_k}\,(x_{k+1} - x)^3 + \left(\frac{y_k}{\delta_k} - \frac{M_k\delta_k}{6}\right)(x_{k+1} - x) + \frac{M_{k+1}}{6\delta_k}\,(x - x_k)^3 + \left(\frac{y_{k+1}}{\delta_k} - \frac{M_{k+1}\delta_k}{6}\right)(x - x_k),$$

where $\delta_k = x_{k+1} - x_k$. Then $s(x)$ satisfies properties 1 and 4. The parameters $M_1, M_2, \dots, M_{n-1}, M_n$ will be chosen to satisfy properties 2 and 3.

Property 3 translates into $n - 2$ equations

$$\frac{\delta_{k-1}}{6}\,M_{k-1} + \frac{\delta_k + \delta_{k-1}}{3}\,M_k + \frac{\delta_k}{6}\,M_{k+1} = \frac{y_{k+1} - y_k}{x_{k+1} - x_k} - \frac{y_k - y_{k-1}}{x_k - x_{k-1}}, \quad k = 2, \dots, n-1.$$

Property 2 adds two more:

$$-\frac{\delta_1}{3}\,M_1 - \frac{\delta_1}{6}\,M_2 + \frac{y_2 - y_1}{x_2 - x_1} = y'(a)$$

$$\frac{\delta_{n-1}}{6}\,M_{n-1} + \frac{\delta_{n-1}}{3}\,M_n + \frac{y_n - y_{n-1}}{x_n - x_{n-1}} = y'(b).$$

This gives a system of n linear equations in n unknowns M_k:

$$\begin{cases} 2\delta_1 M_1 + \delta_1 M_2 = 6\left(\dfrac{y_2 - y_1}{x_2 - x_1} - y'(a)\right) \\ \delta_{k-1} M_{k-1} + 2(\delta_k + \delta_{k-1}) M_k + \delta_k M_{k+1} = 6\left(\dfrac{y_{k+1} - y_k}{x_{k+1} - x_k} - \dfrac{y_k - y_{k-1}}{x_k - x_{k-1}}\right), \quad k = 2, \ldots, n-1, \\ \delta_{n-1} M_{n-1} + 2\delta_{n-1} M_n = 6\left(y'(b) - \dfrac{y_n - y_{n-1}}{x_n - x_{n-1}}\right). \end{cases}$$

In matrix form,

$$\begin{pmatrix} 2\delta_1 & \delta_1 & & & & \\ \delta_1 & 2(\delta_1+\delta_2) & \delta_2 & & & \\ & \delta_2 & 2(\delta_2+\delta_3) & \delta_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \delta_{n-2} & 2(\delta_{n-2}+\delta_{n-1}) & \delta_{n-1} \\ & & & & \delta_{n-1} & 2\delta_{n-1} \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ M_{n-1} \\ M_n \end{pmatrix} = 6 \begin{pmatrix} \frac{y_2-y_1}{x_2-x_1} - y'(a) \\ \frac{y_3-y_2}{x_3-x_2} - \frac{y_2-y_1}{x_2-x_1} \\ \frac{y_4-y_3}{x_4-x_3} - \frac{y_3-y_2}{x_3-x_2} \\ \vdots \\ \frac{y_n-y_{n-1}}{x_n-x_{n-1}} - \frac{y_{n-1}-y_{n-2}}{x_{n-1}-x_{n-2}} \\ y'(b) - \frac{y_n-y_{n-1}}{x_n-x_{n-1}} \end{pmatrix}.$$

$M_1, \ldots, M_n$ are uniquely determined by these equations: the coefficient matrix is diagonally dominant and hence invertible. Calculations are easy because of the tridiagonal structure.

If all subintervals $[x_k, x_{k+1}]$ are of equal length δ, we have further simplification:

$$\begin{pmatrix} 2 & 1 & & & & \\ 1 & 4 & 1 & & & \\ & 1 & 4 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & 1 \\ & & & & 1 & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ M_{n-1} \\ M_n \end{pmatrix} = \frac{6}{\delta^2} \begin{pmatrix} y_2 - y_1 - \delta y'(a) \\ y_3 - 2y_2 + y_1 \\ y_4 - 2y_3 + y_2 \\ \vdots \\ y_n - 2y_{n-1} + y_{n-2} \\ \delta y'(b) - y_n + y_{n-1} \end{pmatrix}.$$

Extremum property

THEOREM. Among all functions $f(x)$ twice continuously differentiable on $[a, b]$ and satisfying the interpolation conditions $f(x_k) = y_k$, $k = 1, \ldots, n$, $f'(a) = y'(a)$, $f'(b) = y'(b)$, the complete cubic spline $s(x)$ uniquely minimizes the integral

$$\int_a^b f''(x)^2 dx.$$

Proof. Let $f(x)$ satisfy the conditions of the theorem and let $\Delta(x) = s(x) - f(x)$. Then $\Delta(x)$ has a root at each x_k and $\Delta'(a) = \Delta'(b) = 0$. Observe that

$$\begin{aligned}
\int_a^b s''(x)\Delta''(x)dx &= s''(x)\Delta'(x)\Big|_a^b - \int_a^b s'''(x)\Delta'(x)dx \\
&= -\int_a^b s'''(x)\Delta'(x)dx && \text{(because } \Delta'(a) = \Delta'(b) = 0\text{)} \\
&= -\sum_{k=1}^{n-1} c_k \int_{x_k}^{x_{k+1}} \Delta'(x)dx && \text{(because } s'''(x)\big|_{(x_k,x_{k+1})} \text{ is constant)} \\
&= -\sum_{k=1}^{n-1} c_k \big[\Delta(x_{k+1}) - \Delta(x_k)\big] \\
&= 0 && \text{(because } \Delta(x_k) = 0\text{).}
\end{aligned}$$

Therefore

$$\begin{aligned}
\int_a^b f''(x)^2 dx &= \int_a^b (s''(x) - \Delta''(x))^2 dx \\
&= \int_a^b s''(x)^2 dx + \int_a^b \Delta''(x)^2 dx \\
&\geq \int_a^b s''(x)^2 dx.
\end{aligned}$$

Equality holds if and only if $\Delta(x)$ is linear. But then $\Delta(a) = \Delta'(a) = 0$ implies that $\Delta(x)$ is zero. □

Error estimate

It has been shown (Hall 1968, Hall & Meyer 1976) that, if $y(x)$ is four times continuously differentiable on $[a, b]$ and $\delta_{\max} = \max_k \delta_k$, then

$$|y(x) - s(x)| \leq \tfrac{5}{384} \max_{[a,b]} |y^{(4)}(x)|\ \delta_{\max}^4, \quad a \leq x \leq b.$$

This estimate is optimal in the sense that the constant $\frac{5}{384}$ is the smallest possible.