Lecture

Spline interpolation

It is not possible to fit sometimes functions representing given set of data points by interpolating polynomials. In such situations, it is better to approximate function piecewise with a polymonial of low-degree. Spline is an example of piecewise connecting polynomials.

The name spline comes from a flexible rod used in draftsmanship. Below figure shows an example of it.

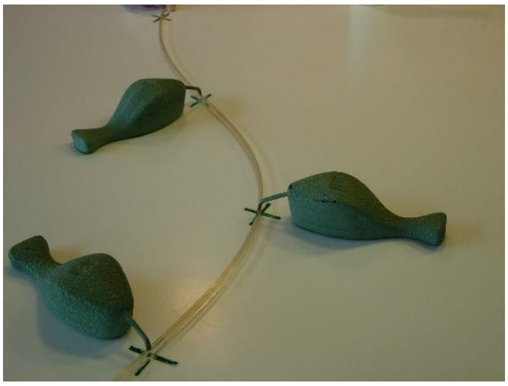


Fig.: Spline (taken from internet).

There are different types of splines.

It could be linear in which each interval is fitted by a straight line. With linear splines problem of discontinuity creeps in at the points where these splines from different intervals meet. The data points at which these splines joins are known as “knots”. Higher order splines do away with this problem. There are quadratic splines where each interval of the function may be fitted by a quadratic polynomial. In most of the cases, a cubic spline is used. Hence, we will discuss cubic splines below.

Cubic splines:

Let us consider we have (x_i, y_i), where i = 0, 1, 2, ……., n; (n+1) data points (not necessarily evenly spaced). We have to find piecewise polynomials which joins like it is shown in Fig. below.

We define here an interval between (x_i, y_i) and (x_{i+1}, y_{i+1}) in which the cubic polynomial is $f_i(x)$.

Writing equation for $f_i(x)$,

$$f_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i, \text{ i = 0,1, 2,….., n-1} \quad (1)$$

We assume here now certain conditions:

(i) $f_i(x_i) = y_i$ for i = 0, 1, 2,……, n-1 (n-interpolating data points)
(ii) $f_i(x_{i+1}) = f_{i+1}(x_{i+1})$, i = 0, 1, 2,…, n-2, continuity condition
(iii) $f_i'(x_{i+1}) = f_{i+1}'(x_{i+1})$, i = 0, 1, 2,…, n-2, continuity condition for slopes
(iv) $f_i''(x_{i+1}) = f_{i+1}''(x_{i+1})$, i = 0, 1, 2,…, n-2, continuous curvature

We shall impose some conditions later known as "end conditions".

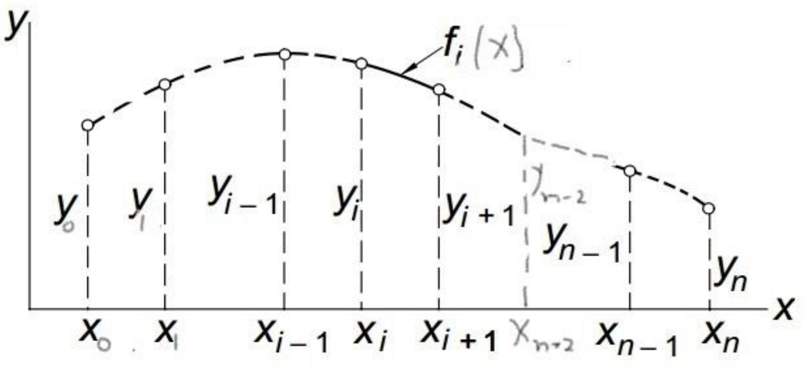


Fig. Cubic spline.

We have 4n unkowns i.e. a_i, b_i, c_i and d_i for i = 0, 1,……., n-1, so from condition (i), putting x = x_i, we get

$$d_i = y_i \tag{2}$$

If we put x = x_{i+1}, we get

$y_{i+1} = f_{i+1}(x_{i+1}) = f_i(x_{i+1}) = a_i(x - x_{i+1})^3 + b_i(x - x_{i+1})^2 + c_i(x - x_{i+1}) + y_i$,
i = 0,1, 2,….., n-1 (3)

Differentiating equation (1),

$f_i'(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i$, i = 0,1, 2,….., n-1 (4)

And differentiation equation (4),

$f_i''(x) = 6a_i(x - x_i) + 2b_i$, i = 0,1, 2,….., n-1 (5)

This is linear equation in interval (x_i, y_i) and (x_{i+1}, y_{i+1}).

Putting x = x_i and letting $S_i = f''(x_i)$ for i = 0, 1,2,….., n-1 and $S_n = f_{n-1}''(x_n)$, we get

$$S_i = 6a_i(x_i - x_i) + 2b_i$$

$$S_i = 2b_i$$

So,

$$b_i = \frac{S_i}{2} \tag{6}$$

Putting x = x_{i+1},

$$S_{i+1} = 6a_i(x_{i+1} - x_i) + 2b_i$$

Now considering $h_i = (x_{i+1} - x_i)$,

We get,

$$S_{i+1} = 6a_ih_i + 2\frac{S_i}{2}$$

Solving,

$$a_i = \frac{S_{i+1} - S_i}{6h_i} \tag{7}$$

Substituting from equations (2), (6) and (7) for d_i, b_i and a_i into equation (1) and putting x = x_{i+1}, we get

$$y_{i+1} = \frac{S_{i+1} - S_i}{6h_i}(x_{i+1} - x_i)^3 + \frac{S_i}{2}(x_{i+1} - x_i)^2 + c_i(x_{i+1} - x_i) + y_i$$

(here $f_i(x_i) = y_i$, from condition (i) so $f_i(x_{i+1}) = y_{i+1}$,)

Using $h_i = (x_{i+1} - x_i)$, we get

$$y_{i+1} = \frac{S_{i+1} - S_i}{6h_i}{h_i}^3 + \frac{S_i}{2}{h_i}^2 + c_ih_i + y_i$$

Hence,

$$c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \tag{8}$$

Imposing condition (iii) i.e. the slopes of two cubic polynomials joining at (x_i, y_i) are same. Thus, with $x = x_i$ for ith interval equation (iii) becomes,

$$y_i' = 3a_i(x_i - x_i)^2 + 2b_i(x_i - x_i) + c_i$$

$$y_i' = c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \tag{9}$$

If we choose slope in x_{i-1} to x_i, we get at right side.

$$y_i' = 3a_{i-1}(x_i - x_{i-1})^2 + 2b_{i-1}(x_i - x_{i-1}) + c_{i-1}$$

Or,

$$y_i' = 3a_{i-1}{h_{i-1}}^2 + 2b_{i-1}h_{i-1} + c_{i-1} \tag{10}$$

where,

$$h_{i-1} = (x_i - x_{i-1})$$

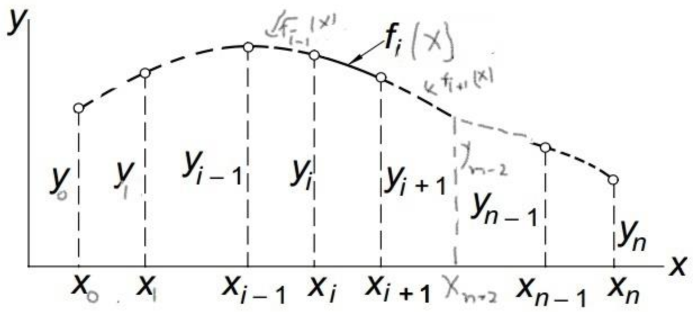
We have,

$a_{i-1} = \frac{S_i - S_{i-1}}{6h_{i-1}}$, $b_{i-1} = \frac{S_{i-1}}{2}$ and $c_{i-1} = \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}S_{i-1} + h_{i-1}S_i}{6}$

Using these relations, we obtain from equation (10),

$$y_i' = 3\frac{S_i - S_{i-1}}{6h_{i-1}}{h_{i-1}}^2 + 2\frac{S_{i-1}}{2}h_{i-1} + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}S_{i-1} + h_{i-1}S_i}{6} \tag{11}$$

Using condition of continuity for y_i', (See Figure below, the slopes in the interval x_{i-1} to x_i and x_i to x_{i+1} should match at $x = x_i$)



Hence, equating equations (9) and (11),

$$y_i' = 3\frac{S_i - S_{i-1}}{6h_{i-1}}\ {h_{i-1}}^2 + 2\frac{S_{i-1}}{2}h_{i-1} + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}S_{i-1} + h_{i-1}S_i}{6}$$
$$= \frac{y_{i+1} - y_i}{h_i} - \frac{2h_iS_i + h_iS_{i+1}}{6}$$

Simplifying, we obtain

$h_{i-1}S_{i-1} + (2h_{i-1} + 2h_i)S_i + h_iS_{i+1} = 6\left(\frac{y_{i+1}-y_i}{h_i} - \frac{y_i-y_{i-1}}{h_{i-1}}\right)$, i = 1,2, …., n-1 (12)

Or,

$$h_{i-1}S_{i-1} + (2h_{i-1} + 2h_i)S_i + h_iS_{i+1} = 6(f[x_i, x_{i+1}] - f[x_{i-1}, x_i])$$

As one can see, equation (12) contains divided differences.

In matrix form, we can write above equation as

$$\begin{bmatrix} h_0 & 2(h_0 + h_1) & h_1 & & \\ \vdots & h_1 & 2(h_1 + h_2) & h_2 & \vdots \\ & \cdots & & & h_{n-1} \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_{n-1} \\ S_n \end{bmatrix} =$$

$$6\begin{bmatrix} f[x_1, x_2] - f[x_0, x_1] \\ f[x_2, x_3] - f[x_1, x_2] \\ \vdots \\ f[x_{n-1}, x_n] - f[x_{n-1}, x_{n-1}] \end{bmatrix} \qquad (13)$$

Note here that there are (n-1) equations and (n+1) unknowns. Solution of this equation can be carried out under certain conditions discussed below. This is a tridiagonal matrix and it could be solved with suitable technique.

If the intervals, $h_{i-1} = h_i = h_{i+1} = h$ i.e. evenly spaced data points, then equation (12) becomes,

$$S_{i-1} + 4S_i + S_{i+1} = \frac{6}{h^2}(y_{i+1} - 2y_i + y_{i-1}),\ \text{i = 1, 2,….., n-1} \qquad (14)$$

In order to solve these equations, we must know S_0 *and* S_n. These are called end conditions. Mostly, four conditions are chosen.

1. S_0 *and* S_n= 0. It means in the extremities the cubics reach lineality. This is known as "natural spline".
2. The end slopes could be taken as specified values. If the information is not available, we could choose slopes from the points. Let f'(x_0) = P and f'(x_n) = Q, so
   Left end, we can write: $2h_0S_0 + h_0S_1 = 6(f[x_0, x_1] - P)$
   Right end: $2h_{n-1}S_n + h_{n-1}S_{n-1} = 6(Q - f[x_{n-1}, x_n])$

   (Use of divided differences been made)
3. $S_0 = S_1$ *and* $S_{n-1} = S_n$ i.e. at the end cubics approach parabola.
4. We consider:
   At left end: $\frac{S_1 - S_0}{h_0} = \frac{S_2 - S_1}{h_1}$ *and* $S_0 = \frac{(h_0 + h_1)\, S_1 - h_0 S_2}{h_1}$
   Right end: $\frac{S_n - S_{n-1}}{h_{n-1}} = \frac{S_{n-1} - S_{n-2}}{h_{n-2}}$ *and* $S_n = \frac{(h_{n-2} + h_{n-1})\, S_{n-1} - h_{n-1} S_{n-1}}{h_{n-2}}$ (15)
   This is known as "not a knot condition".

In a separate sheet, the coefficient matrix for each condition is discussed.