Cubic Spline Interpolation

Abstract

This methodological framework examines the construction of a piecewise cubic function with continuous first and second derivatives that passes through a given set of data points. Cubic spline interpolation uses the interpolation conditions at the data points and the continuity of the first and second derivatives across adjacent intervals to determine the unknown coefficients of the cubic polynomials. The objective of this article is to improve the accuracy of temporal disaggregation by converting annual or quarterly data into higher-frequency monthly or quarterly data using cubic spline interpolation.

Keywords: curve function, derivative, interpolation, cubic spline, matrix.

Prepared by the Financial Stability Department of the Central Bank of Uzbekistan.

If you have any comments, please send them to udjumanazarov@cbu.uz.

Cubic Spline Interpolation

Cubic spline interpolation[1] is a method of generating a third degree continuous piecewise function that passes through multiple points. The function determined by this method serves to more accurately represent the data between the points.

A continuous piecewise function passing through several points according to the cubic spline interpolation methodology[2] is expressed as:

$$S(x) = \begin{cases} s_1(x) \; if \; x_1 \leq x \leq x_2 \\ s_2(x) \; if \; x_2 \leq x \leq x_3 \\ \vdots \\ s_{n-1}(x) \; if \; x_{n-1} \leq x \leq x_n \end{cases}$$

$$s_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i$$
$$i = 1, 2, \dots, n - 1$$

Here, s_i is a third degree polynomial, n is number of points, x_n are data points, $a_i, b_i, \; c_i$ and d_i are coefficients.

The first and second derivatives are obtained from the above third degree polynomial:

$$s_i'(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i$$
$$s_i''(x) = 6a_i(x - x_i) + 2b_i$$
$$i = 1, 2, \dots, n - 1$$

Also, when determining the coefficients in the function, the properties of cubic spline interpolation are used: the piecewise function $S(x)$ will interpolate at all data points, the function $S(x)$, and the first and second derivatives of this function will be continuous on the interval $[x_1, x_n]$.

Due to the fact that the function $S(x)$ will interpolate at all of the data points, the following equality holds:

[1] A third order continuous curve with continuous derivatives passing through a given set of points.

[2] McKinley, S., & Levine, M. (1998). Cubic Spline Interpolation.

$$s_i(x_i) = a_i(x_i - x_i)^3 + b_i(x_i - x_i)^2 + c_i(x_i - x_i) + d_i$$
$$s_i(x_i) = d_i$$

According to the property that a piecewise function must be continuous across its entire interval, each sub-function must join at the data points:

$$s_i(x_i) = s_{i-1}(x_i)$$
$$i = 2, \dots, n$$
$$s_{i-1}(x_i) = a_{i-1}(x_i - x_{i-1})^3 + b_{i-1}(x_i - x_{i-1})^2 + c_{i-1}(x_i - x_{i-1}) + d_{i-1}$$
$$d_i = a_{i-1}(x_i - x_{i-1})^3 + b_{i-1}(x_i - x_{i-1})^2 + c_{i-1}(x_i - x_{i-1}) + d_{i-1}$$
$$i = 2, \dots, n-1$$
$$h = x_i - x_{i-1}$$
$$d_i = a_{i-1}h^3 + b_{i-1}h^2 + c_{i-1}h + d_{i-1}$$

According to the property that the first-order derivative of the function $S(x)$ must be continuous on the interval $[x_1, x_n]$, the following equations are formed:

$$s_i'(x_i) = s_{i-1}'(x_i)$$
$$s_i'(x_i) = 3a_i(x_i - x_i)^2 + 2b_i(x_i - x_i) + c_i$$
$$s_i'(x_i) = s_{i-1}'(x_i) = c_i$$
$$c_i = 3a_{i-1}(x_i - x_{i-1})^2 + 2b_{i-1}(x_i - x_{i-1}) + c_{i-1}$$
$$c_i = 3a_{i-1}h^2 + 2b_{i-1}h + c_{i-1}$$

Based on the property of the second derivative of the function $S(x)$ must be continuous on the interval $[x_1, x_n]$, that is defined as:

$$s_i''(x_i) = s_{i-1}''(x_i)$$
$$s_i''(x_{i+1}) = s_{i+1}''(x_{i+1})$$
$$s_i''(x_i) = 6a_i(x_i - x_i) + 2b_i$$
$$s_i''(x_i) = s_{i-1}''(x_i) = 2b_i$$
$$2b_i = 6a_{i-1}(x_i - x_{i-1}) + 2b_{i-1}$$
$$2b_i = 6a_{i-1}h + 2b_{i-1}$$

$$2b_{i+1} = 6a_i h + 2b_i$$

The above expressions are simplified by substituting M_i for $s_i''(x_i)$:

$$s_i''(x_i) = 2b_i$$
$$M_i = 2b_i$$
$$b_i = \frac{M_i}{2}$$

The value d_i is equal to the value of the function $S(x)$ at the point x_i ($s_i(x_i) = d_i$). Also, the value of a_i is re-written by the above designations:

$$2b_{i+1} = 6a_i h + 2b_i$$
$$6a_i h = 2b_{i+1} - 2b_i$$
$$a_i = \frac{2b_{i+1} - 2b_i}{6h}$$
$$a_i = \frac{2\left(\frac{M_{i+1}}{2}\right) - 2\left(\frac{M_i}{2}\right)}{6h}$$
$$a_i = \frac{M_{i+1} - M_i}{6h}$$

The value of c_i is re-written by the entered designations:

$$d_{i+1} = a_i h^3 + b_i h^2 + c_i h + d_i$$
$$c_i h = d_{i+1} - a_i h^3 - b_i h^2 - d_i$$
$$c_i = \frac{d_{i+1} - a_i h^3 - b_i h^2 - d_i}{h}$$
$$c_i = (-a_i h^2 - b_i h) - \frac{d_i - d_{i+1}}{h}$$
$$c_i = \frac{d_{i+1} - d_i}{h} - \left(\frac{M_{i+1} - M_i}{6h} h^2 - \frac{M_i}{2} h\right)$$
$$c_i = \frac{y_{i+1} - y_i}{h} - \left(\frac{M_{i+1} + 2M_i}{6}\right) h$$

The following expressions are formed to calculate the coefficients of the function through the defined equations:

$$
\begin{cases}
a_i = \dfrac{M_{i+1} - M_i}{6h} \\
b_i = \dfrac{M_i}{2} \\
c_i = \dfrac{y_{i+1} - y_i}{h} - \left(\dfrac{M_{i+1} + 2M_i}{6}\right) h \\
d_i = y_i
\end{cases}
$$

Determining the second derivative of third degree polynomials ($s_i''(x_i) = M_i$) is enough to calculate the coefficients. The expressions can be handled more conveniently by putting them into matrix form as follows:

$$c_{i+1} = 3a_i h^2 + 2b_i h + c_i$$

$$3\left(\frac{M_{i+1} - M_i}{6h}\right) h^2 + 2\left(\frac{M_i}{2}\right) h + \frac{y_{i+1} - y_i}{h} - \left(\frac{M_{i+1} + 2M_i}{6}\right) h$$

$$= \frac{y_{i+2} - y_{i+1}}{h} - \left(\frac{M_{i+2} + 2M_{i+1}}{6}\right) h$$

$$h\left(\frac{3M_{i+1} - 3M_i}{6} + \frac{6M_i}{6} - \left(\frac{M_{i+1} + 2M_i}{6}\right) + \left(\frac{M_{i+2} + 2M_{i+1}}{6}\right)\right) = \frac{y_i - 2y_{i+1} + y_{i+2}}{h}$$

$$\frac{h}{6} * (M_i + 4M_{i+1} + M_{i+2}) = \frac{y_i - 2y_{i+1} + y_{i+2}}{h}$$

$$M_i + 4M_{i+1} + M_{i+2} = 6\left(\frac{y_i - 2y_{i+1} + y_{i+2}}{h^2}\right)$$

$$i = 1,2,3, \dots, n-1$$

$$
\begin{bmatrix}
1 & 4 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\
0 & 1 & 4 & 1 & \dots & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & \dots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & 4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 1 & 4 & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & 4 & 1
\end{bmatrix}
\begin{bmatrix}
M_1 \\ M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \\ M_n
\end{bmatrix}
= \frac{6}{h^2}
\begin{bmatrix}
y_1 - 2y_2 + y_3 \\
y_2 - 2y_3 + y_4 \\
y_3 - 2y_4 + y_5 \\
\vdots \\
y_{n-4} - 2y_{n-3} + y_{n-2} \\
y_{n-3} - 2y_{n-2} + y_{n-1} \\
y_{n-2} - 2y_{n-1} + y_n
\end{bmatrix}
$$

The resulting matrix equation contains n unknowns ($M_1, M_2, \dots, M_{n-1}, M_n$) and $n-2$ equalities. This does not allow to fully determining the coefficients. Therefore, the coefficients are determined by introducing natural splines, parabolic runout splines, and cubic runout splines for polynomials at the endpoints.

The condition of natural splines includes that the second derivative of the third degree polynomials ($s_i''(x_i) = M_i$) is equal to zero at the endpoints ($M_1 = M_n = 0$). As a result, the above matrix equation becomes:

$$\begin{bmatrix} 1 & 4 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \\ 0 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \end{bmatrix}$$

For reasons of convenience, the first and last columns of this matrix can be eliminated, as they correspond to the M_1 and M_n values, which are both zero.

$$\begin{bmatrix} 4 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 4 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \end{bmatrix}$$

It can be seen from the resulting matrix that there are $n-2$ unknowns ($M_2, M_3, \dots, M_{n-2}, M_{n-1}$) and $n-2$ equalities. As a result, unknown values for M_2 through M_{n-1} will be determined. Based on the determined M_i values, the coefficients of the third degree polynomials are calculated, and the continuous piecewise function passing through several points is determined.

The condition of parabolic runout spline that the second derivatives of the third degree polynomials ($s_i''(x_i) = M_i$) at the endpoints are equal to the second and one point before the end, respectively ($M_1 = M_2$ and $M_n = M_{n-1}$). As a result, the matrix equation above becomes the following:

$$\begin{bmatrix} 5 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 4 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \end{bmatrix}$$

By solving the matrix equation, the unknown values for M_2 through M_{n-1} are determined. Based on the determined M_i values, the coefficients of the third degree polynomials are calculated, and the continuous piecewise function passing through several points is determined.

Cubic runout spline condition contains certain relations $M_1 = 2M_2 - M_3$ and $M_n = 2M_{n-1} - M_{n-2}$) of the second derivative ($s_i''(x_i) = M_i$) and the above matrix equation becomes:

$$\begin{bmatrix} 6 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 4 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \end{bmatrix}$$

The resulting matrix contains $n - 2$ unknowns ($M_2, M_3, \dots, M_{n-2}, M_{n-1}$) and $n - 2$ equalities. In this way, unknown values for M_2 through M_{n-1} are determined.

The unknown coefficients of the third degree polynomials $S(x)$ are calculated based on the values of M_i determined according to the above conditions of cubic spline interpolation. As a result, a third order continuous piecewise function passing through several points is found.