CS412: Lecture #13

Mridul Aanjaneya

March 3, 2015

The Cubic Spline

As always, our goal in this interpolation task is to define a curve $s(x)$ which interpolates the n data points

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \quad \text{(where } x_1 < x_2 < \ldots < x_n\text{)}$$

In the fashion of piecewise polynomials, we will define $s(x)$ as a different cubic polynomial $s_k(x)$ at each sub-interval $I_k = [x_k, x_{k+1}]$, i.e.,

$$s(x) = \begin{cases} s_1(x), x \in I_1 \\ s_2(x), x \in I_2 \\ \vdots \\ s_{n-1}(x), x \in I_{n-1} \end{cases}$$

Each of the s_k's is a cubic polynomial:

$$s_k(x) = a_3^{(k)} x^3 + a_2^{(k)} x^2 + a_1^{(k)} x + a_0^{(k)}$$

where $a_3^{(k)}$, $a_2^{(k)}$, $a_1^{(k)}$, $a_0^{(k)}$ are unknown coefficients. Since we have $n-1$ piecewise polynomials, in total we shall have to determine $4(n-1) = 4n - 4$ unknown coefficients. The points $(x_2, x_3, \ldots, x_{n-1})$ where the formula for $s(x)$ changes from one cubic polynomial (s_k) to another (s_{k+1}) are called *knots*.

Note: In some textbooks, the extreme points x_1 and x_n are also included in the definition of what a knot is. We will stick with the definition we stated above.

The piecewise polynomial interpolation method described as *cubic spline* also requires the neighboring polynomials s_k and s_{k+1} to be joined at x_{k+1} with a certain degree of smoothness. In detail:

- The curve should be continuous: $s_k(x_{k+1}) = s_{k+1}(x_{k+1})$
- The derivative (slope) should be continuous: $s'_k(x_{k+1}) = s'_{k+1}(x_{k+1})$
- The 2nd derivative should be continuous as well: $s''_k(x_{k+1}) = s''_{k+1}(x_{k+1})$

(*Note:* If we force the next (3rd) derivative to match, this will force s_k and s_{k+1} to be exactly identical.)

When determining the unknown coefficients $\{a_i^{(j)}\}$, each of these 3 smoothness constraints (for knots $k = 2, 3, \ldots, n-1$) needs to be satisfied, for a total of $3(n-2) = 3n-6$ constraint equations. We should not forget that we additionally want to *interpolate* all n data points, i.e.,

$$s(x_i) = y_i \quad \text{for } i = 1, 2, \ldots, n$$

In total, we have $3n - 6 + n = 4n - 6$ total equations to satisfy, and $4n - 4$ unknowns! Consequently, we will need 2 more equations to ensure that the unknown coefficients will be uniquely determined. Several plausible options exist on how to do that:

1. The "not-a-knot" approach: We stipulate that at the locations of the first knot (x_2) and last knot (x_{n-1}) the *third* derivative of $s(x)$ should also be continuous, e.g.:

   $$s_1'''(x_2) = s_2'''(x_2) \quad \text{and} \quad s_{n-2}'''(x_{n-1}) = s_{n-1}'''(x_{n-1})$$

   As we discussed before, these two additional constraints will effectively cause $s_1(x)$ to be identical with $s_2(x)$, and $s_{n-2}(x)$ to coincide with $s_{n-1}(x)$. In this sense, x_2 and x_{n-1} are no longer "knots" in the sense that the formula for $s(x)$ "changes" at these points (hence the name).

2. Complete spline: If we have access to the derivative f' of the function being sampled by the y_i's (i.e., $y_i = f(x_i)$), we can formulate the two additional constraints as:

   $$s_1'(x_1) = f'(x_1) \quad \text{and} \quad s_{n-1}'(x_n) = f'(x_n)$$

   Note that qualitatively, using the complete spline approach is a better utilization of the flexibility of the spline curve in matching yet one more property of f. In contrast, the not-a-knot approach makes the spline "less flexible" by removing two degrees of freedom, in order to obtain a unique solution. However, we cannot always assume knowledge of f'.

3. The natural cubic spline: We use the following two constraints:

   $$s''(x_1) = 0 \quad \text{and} \quad s''(x_n) = 0$$

   Thus, $s(x)$ reaches the endpoints looking like a straight line (instead of a curved one).

4. Periodic spline: The following two constraints are used:

   $$s'(x_1) = s'(x_n) \quad \text{and} \quad s''(x_1) = s''(x_n)$$

   This is useful when the underlying function f is also known to be periodical over $[a, b]$.

Since $s(x)$ is piecewise cubic, its second derivative $s''(x)$ is piecewise linear on $[x_1, x_n]$. The linear Lagrange interpolation formula gives the following representation for $s''(x) = s''_k(x)$ on $[x_k, x_{k+1}]$:

$$s''_k(x) = s''(x_k)\frac{x - x_{k+1}}{x_k - x_{k+1}} + s''(x_{k+1})\frac{x - x_k}{x_{k+1} - x_k}$$

Defining $m_k = s''(x_k)$ and $h_k = x_{k+1} - x_k$ gives

$$s''_k(x) = \frac{m_k}{h_k}(x_{k+1} - x) + \frac{m_{k+1}}{h_k}(x - x_k)$$

for $x_k \le x \le x_{k+1}$ and $k = 1, 2, \ldots, n-1$. Integrating the above equation twice will introduce two constants of integration, and the result can be manipulated so that it has the form:

$$s_k(x) = \frac{m_k}{6h_k}(x_{k+1} - x)^3 + \frac{m_{k+1}}{6h_k}(x - x_k)^3 + p_k(x_{k+1} - x) + q_k(x - x_k) \qquad (1)$$

Substituting x_k and x_{k+1} into equation (1) and using the values $y_k = s_k(x_k)$ and $y_{k+1} = s_k(x_{k+1})$ yields the following equations that involve p_k and q_k respectively:

$$y_k = \frac{m_k}{6}h_k^2 + p_k h_k \quad \text{and} \quad y_{k+1} = \frac{m_{k+1}}{6}h_k^2 + q_k h_k$$

These two equations are easily solved for p_k and q_k, and when these values are substituted into equation (1), the result is the following expression for the cubic function $s_k(x)$:

$$\begin{aligned} s_k(x) &= \frac{m_k}{6h_k}(x_{k+1} - x)^3 + \frac{m_{k+1}}{6h_k}(x - x_k)^3 + \left(\frac{y_k}{h_k} - \frac{m_k h_k}{6}\right)(x_{k+1} - x) \\ &+ \left(\frac{y_{k+1}}{h_k} - \frac{m_{k+1} h_k}{6}\right)(x - x_k) \qquad (2) \end{aligned}$$

Notice that equation (2) has been reduced to a form that involves only the unknown coefficients $\{m_k\}$. To find these values, we must use the derivative of equation (2), which is

$$\begin{aligned} s'_k(x) &= -\frac{m_k}{2h_k}(x_{k+1} - x)^2 + \frac{m_{k+1}}{2h_k}(x - x_k)^2 - \left(\frac{y_k}{h_k} - \frac{m_k h_k}{6}\right) \\ &+ \frac{y_{k+1}}{h_k} - \frac{m_{k+1} h_k}{6} \qquad (3) \end{aligned}$$

Evaluating equation (3) at x_k and simplifying the result yields:

$$s'_k(x_k) = -\frac{m_k}{3}h_k - \frac{m_{k+1}}{6}h_k + d_k, \quad \text{where } d_k = \frac{y_{k+1} - y_k}{h_k} \qquad (4)$$

Similarly, we can replace k by $k-1$ in equation (3) to get the expression for $s'_{k-1}(x)$ and evaluate it at x_k to obtain

$$s'_{k-1}(x_k) = \frac{m_k}{3}h_{k-1} + \frac{m_{k-1}}{6}h_{k-1} + d_{k-1} \qquad (5)$$

Now using the continuity of derivatives and equations (4) and (5) gives an important relation involving m_{k-1}, m_k and m_{k+1}:

$$h_{k-1}m_{k-1} + 2(h_{k-1} + h_k)m_k + h_k m_{k+1} = u_k \tag{6}$$

where $u_k = 6(d_k - d_{k-1})$ for $k = 2, \ldots, n-1$. Observe that the unknowns in equation (6) are the desired values $\{m_k\}$, and the other terms are constants obtained by performing simple arithmetic with the data points $\{x_k, y_k\}$. Therefore, in reality, system (6) is an underdetermined system of $n-2$ linear equations involving n unknowns. Hence, two additional equations must be supplied. They are used to eliminate m_1 and m_n. Consider the natural cubic spline strategy where m_1 and m_n are given ($= 0$). The first equation (for $k = 2$) of system (6) is:

$$2(h_1 + h_2)m_2 + h_2 m_3 = u_2 - h_1 m_1 \tag{7}$$

and similarly, the last equation is:

$$h_{n-2}m_{n-2} + 2(h_{n-2} + h_{n-1})m_{n-1} = u_{n-1} - h_{n-1}m_n \tag{8}$$

Equations (7) and (8) with (6) used for $k = 3, 4, \ldots, n-2$ form a tridiagonal $(n-2) \times (n-2)$ linear system $HM = V$ involving the coefficients $m_2, m_3, \ldots, m_{n-1}$:

$$\begin{bmatrix} b_2 & c_2 & & & & \\ a_3 & b_3 & c_3 & & & \\ & & \ddots & & & \\ & & & a_{n-3} & b_{n-2} & c_{n-2} \\ & & & & a_{n-2} & b_{n-1} \end{bmatrix} \begin{bmatrix} m_2 \\ m_3 \\ \vdots \\ m_{n-2} \\ m_{n-1} \end{bmatrix} = \begin{bmatrix} v_2 \\ v_3 \\ \vdots \\ v_{n-2} \\ v_{n-1} \end{bmatrix}$$

After the coefficients $\{m_k\}$ are determined, the spline coefficients $a_k^{(j)}$ for $s_k(x)$ are computed using the formulas

$$a_k^{(0)} = y_k, \quad a_k^{(1)} = d_k - \frac{h_k}{6}(2m_k + m_{k+1}), \quad a_k^{(2)} = \frac{m_k}{2}, \quad a_k^{(3)} = \frac{m_{k+1} - m_k}{6h_k}$$

Error analysis

For simplicity, we will again assume that

$$h_2 = h_3 = \ldots = h_{n-1} = h \qquad (h_k = x_{k+1} - x_k)$$

For the not-a-knot method, we have

$$|f(x) - s(x)| \lessapprox \frac{5}{384}||f^{(4)}||_\infty \cdot h^4$$

The "approximate" inequality is used because the interpolation error can be slightly larger near the endpoints of the interval $[a, b]$. This is a very comparable result with the (non-smooth) piecewise cubic polynomial method:

$$|f(x) - s(x)| \leq \frac{9}{384}||f^{(4)}||_\infty \cdot h^4$$

Note though that the computation of the piecewise cubic method was *very local* and simple (every interval could be independently evaluated) while the computation of the coefficients of the cubic spline is more elaborate.