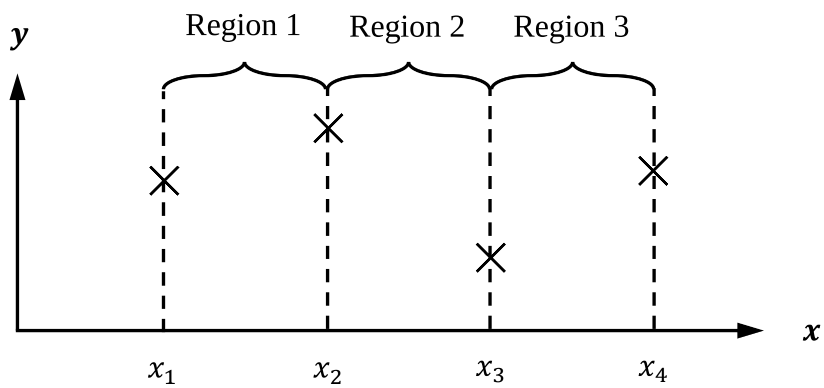
This note provides further guidance about Task A3.

Let's say that I gave you four points instead of six, giving you $n = 3$ regions:



Our end goal is to find cubic fit equations for each of the three regions, defined by:

$$g_1(x) = a_1 + b_1(x - x_1) + c_1(x - x_1)^2 + d_1(x - x_1)^3 \tag{1}$$

for region 1,

$$g_2(x) = a_2 + b_2(x - x_2) + c_2(x - x_2)^2 + d_2(x - x_2)^3 \tag{2}$$

for region 2, and

$$g_3(x) = a_3 + b_3(x - x_3) + c_3(x - x_3)^2 + d_3(x - x_3)^3 \tag{3}$$

for region 3.

This means that we have $4 \times 3 = 12$ unknowns. We need to come up with at least 12 equations relating the unknowns to each other and the values we know, which are the x and y values for our four points. These equations are called constraints, restrictions on the cubic fit equations. How do we want to constrain our fit? Well first off, we want our fit equations to pass through our points:

$$g_1(x_1) = y_1 \tag{4}$$

$$g_1(x_2) = g_2(x_2) = y_2 \tag{5}$$

$$g_2(x_3) = g_3(x_3) = y_3 \tag{6}$$

$$g_3(x_4) = y_4 \tag{7}$$

This also ensures that the fit is continuous across all regions. We don't want there to be any discontinuities (sudden jumps or steps) between one fit and the next.

Substituting into Equations (1), (2), and (3), we find that:

$$g_1(x_1) = y_1 = a_1 \tag{I}$$

$$g_2(x_2) = y_2 = a_2 \tag{II}$$

$$g_3(x_3) = y_3 = a_3 \tag{III}$$

Where I am using Roman numerals to label our constraint equations. Right off the bat, we know all three a coefficients. To make the math a little bit easier to follow, we can define the step size as:

$$h_i = x_{i+1} - x_i$$

where $i = 1 \ldots n$. Substituting the step size and the expressions for the a coefficients into Equations (5), (6), and (7) we find that:

$$g_1(x_2) = g_2(x_2) \implies a_1 + b_1(x_2 - x_1) + c_1(x_2 - x_1)^2 + d_1(x_2 - x_1)^3 = a_2$$

$$\implies b_1 h_1 + c_1 h_1^2 + d_1 h_1^3 = y_2 - y_1 \tag{IV}$$

and that:

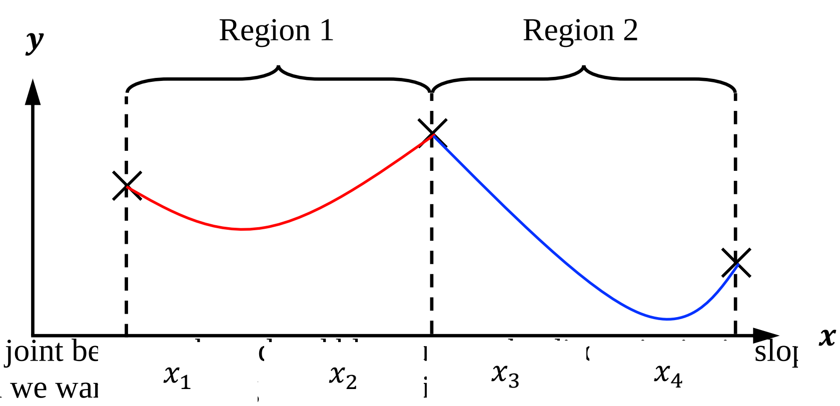
$$g_2(x_3) = g_3(x_3) \implies a_2 + b_2(x_3 - x_2) + c_2(x_3 - x_2)^2 + d_2(x_3 - x_2)^3 = a_3$$

$$\implies b_2 h_2 + c_2 h_2^2 + d_2 h_2^3 = y_3 - y_2 \tag{V}$$

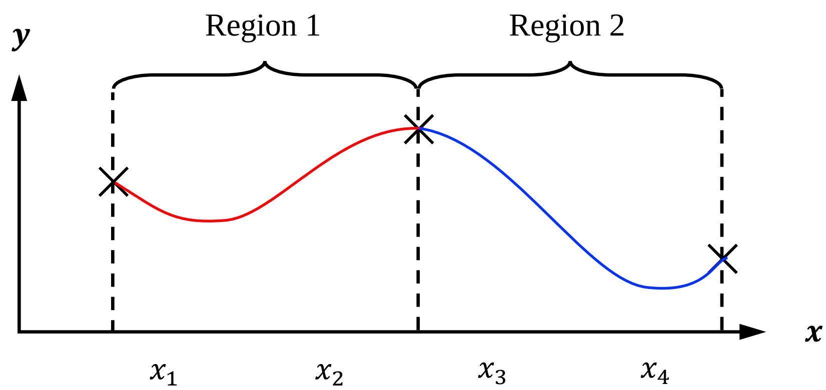
$$g_3(x_4) = y_4 = a_3 + b_3(x_4 - x_3) + c_3(x_4 - x_3)^2 + d_3(x_4 - x_3)^3$$

$$\implies b_3 h_3 + c_3 h_3^2 + d_3 h_3^3 = y_4 - y_3 \tag{VI}$$

Okay, we have 6/12 constraints figured out. How else do we want to constrain the cubic fits? At each of the inner points, (x_2, y_2) and (x_3, y_3), where two fit curves meet, we want to make sure those joints between curves are "smooth". We want to avoid something like this:



Focusing on the joint be… slop… ould be obvious. Instead we wa…



To achieve this, we constrain the cubic fits to have the same first and second derivatives at the intersection points. In other words, the tangent and tangent of the tangent of the cubic fits have to be the same at the inner data points:

First derivatives:

$$\frac{dg_1(x_2)}{dx} = \frac{dg_2(x_2)}{dx}$$

$$\frac{dg_2(x_3)}{dx} = \frac{dg_3(x_3)}{dx}$$

Second derivatives:

$$\frac{d^2g_1(x_2)}{dx^2} = \frac{d^2g_2(x_2)}{dx^2}$$

$$\frac{d^2g_2(x_3)}{dx^2} = \frac{d^2g_3(x_3)}{dx^2}$$

Carrying out these derivatives one by one we find that:

$$b_1 + 2c_1h_1 + 3d_1h_1^2 = b_2 \qquad \text{(VII)}$$

$$b_2 + 2c_2h_2 + 3d_2h_2^2 = b_3 \qquad \text{(VIII)}$$

$$c_1 + 3d_1h_1 = c_2 \qquad \text{(IX)}$$

$$c_2 + 3d_2h_2 = c_3 \qquad \text{(X)}$$

Okay, we're almost there, we have 10/12 constraint equations. Before we move forward, let's tidy up the first 10 equations a little bit. Let's express the values for the a, b, and d coefficients in terms of the c coefficients and the things we know (h_i and y_i). **This way, all we have to do is solve the resulting equations for the c coefficients and then work backwards to get the a, b, and d coefficients.** This effectively reduces the problem from 12 unknowns to four.

Starting with Equations (IX) and (X) we have that:

$$d_1 = \frac{c_2 - c_1}{3h_1}$$

and

$$d_2 = \frac{c_3 - c_2}{3h_2}$$

Hold on, what about d_3? Let's assume there is some value for c_4 and relate it to d_3 in the same way:

$$d_3 = \frac{c_4 - c_3}{3h_3}$$

This step might seem a little confusing, how can we just add a fourth c coefficient? I thought there were only three regions. Just imagine it as another variable that we now have to solve for. Don't give attribute much significance to it, just know that we now have 13 unknowns, not 12, so we need 13 equations.

In the general case for $i = 1 \dots n$:

$$d_i = \frac{c_{i+1} - c_i}{3h_i}$$

Substituting this result into Equations (VII) and (VIII) we find that:

$$b_1 + 2c_1h_1 + h_1(c_2 - c_1) = b_2 \tag{8}$$

$$b_2 + 2c_2h_2 + h_2(c_3 - c_2) = b_3 \tag{9}$$

Substituting the same result into Equations (IV), (V), and (VI) we find that:

$$b_1h_1 + c_1h_1^2 + \frac{(c_2 - c_1)h_1^2}{3} = y_2 - y_1$$

$$b_2h_2 + c_2h_2^2 + \frac{(c_3 - c_2)h_2^2}{3} = y_3 - y_2$$

$$b_3h_3 + c_3h_3^2 + \frac{(c_4 - c_3)h_3^2}{3} = y_4 - y_3$$

After multiplying both sides by three and doing some rearranging, the above three equations can be written as:

$$b_1 = \frac{y_2 - y_1}{h_1} - \frac{(2c_1 + c_2)h_1}{3}$$

$$b_2 = \frac{y_3 - y_2}{h_2} - \frac{(2c_2 + c_3)h_2}{3}$$

$$b_3 = \frac{y_4 - y_3}{h_3} - \frac{(2c_3 + c_4)h_3}{3}$$

At this point we have equations for all three a, b, and d coefficients in terms of the c coefficients and the things we know (h_i and y_i). To recap:

These equations are true regardless of boundary conditions:

***a* coefficients:**

$$a_1 = y_1$$

$$a_2 = y_2$$

$$a_3 = y_3$$

***b* coefficients:**

$$b_1 = \frac{y_2 - y_1}{h_1} - \frac{(2c_1 + c_2)h_1}{3}$$

$$b_2 = \frac{y_3 - y_2}{h_2} - \frac{(2c_2 + c_3)h_2}{3}$$

$$b_3 = \frac{y_4 - y_3}{h_3} - \frac{(2c_3 + c_4)h_3}{3}$$

***d* coefficients:**

$$d_1 = \frac{c_2 - c_1}{3h_1}$$

$$d_2 = \frac{c_3 - c_2}{3h_2}$$

$$d_3 = \frac{c_4 - c_3}{3h_3}$$

That’s 9/13 equations (remember that we added that pesky c_4 coefficient into the mix). We only have four unknowns left, c_1, c_2, c_3, and c_4. We have reduced the original problem of 12 unknowns into one of only four unknowns.

We still haven’t used Equations 8 and 9. This will tick off two of the remaining four:

Substituting the b coefficients equations (page 5) into Equation 8 we find that:

$$\frac{y_2 - y_1}{h_1} - \frac{(2c_1 + c_2)h_1}{3} + 2c_1h_1 + h_1(c_2 - c_1) = \frac{y_3 - y_2}{h_2} - \frac{(2c_2 + c_3)h_2}{3}$$

which, after some rearranging, can be rewritten as:

$$h_1c_1 + 2(h_1 + h_2)c_2 + h_2c_3 = 3\left[\frac{y_3 - y_2}{h_2} - \frac{y_2 - y_1}{h_1}\right] \tag{10}$$

It can easily be shown that we get a similar result for Equation 9:

$$h_2c_2 + 2(h_2 + h_3)c_3 + h_3c_4 = 3\left[\frac{y_4 - y_3}{h_3} - \frac{y_3 - y_2}{h_2}\right] \tag{11}$$

Okay, we're close. All we need is two more equations with only c coefficients and the things we know (h_i and y_i). Where do these last two equations come from?

This is the part where you should really start paying attention…

Boundary conditions. We forced g_1 to pass through the first data point and g_3 to pass through the last data point but said nothing about their derivatives at the first and last points. **The aim of Tasks A2 and A3 is to show you that after all this work (five + pages) the choice of boundary conditions can have a significant impact on the shape of the cubic fits.** In the lab handout I listed a few types of boundary conditions:

1. Natural Spline:

$$g_1^{(2)}(x_1) = 0, \qquad g_n^{(2)}(x_{n+1}) = 0$$

2. End Slope Spline:

$$g_1^{(1)}(x_1) = y_0^{(1)}, \qquad g_n^{(1)}(x_{n+1}) = y_n^{(1)}$$

   where $y_0^{(1)}$ and $y_n^{(1)}$ are known.

3. Periodic Spline:

$$g_1^{(1)}(x_1) = g_n^{(1)}(x_{n+1}), \qquad g_1^{(2)}(x_1) = g_n^{(2)}(x_{n+1})$$

4. Not-a-Knot Spline (MATLAB defaults to this alternative):

$$g_1^{(3)}(x_2) = g_2^{(3)}(x_2), \qquad g_{n-1}^{(3)}(x_n) = g_n^{(3)}(x_n)$$

The math you carried out in Task A2 corresponded to the natural spline. I then asked you to consider the periodic spline conditions instead.

The important point to note here is that in my derivation of Equations 10 and 11 I did **not** use the boundary conditions. In other words, **Equations 10 and 11 hold for any boundary conditions**.

So, what do these boundary conditions mean? In the natural spline case, I can substitute my general expressions for $g_1(x)$ and $g_3(x)$:

$$g_1^{(2)}(x_1) = 0 = \frac{d^2}{dx^2}(a_1 + b_1(x - x_1) + c_1(x - x_1)^2 + d_1(x - x_1)^3)@(x_1)$$

$$= \left(2c_1 + 6d_1(x - x_1)\right) @(x = x_1) = 2c_1 \Rightarrow \boldsymbol{c_1 = 0}$$

The second boundary condition tells me that:

$$g_n^{(2)}(x_{n+1}) = 0 = \frac{d^2}{dx^2}(a_3 + b_3(x - x_3) + c_3(x - x_3)^2 + d_3(x - x_3)^3)@(x_{n+1})$$

$$= 2c_3 + 6d_3(x_4 - x_3)$$

We also have an expression for $d_3$from our d coefficient equations (page 5), substituting this into the above equation we find that:

$$2c_3 + 6d_3(x_4 - x_3) = 2c_3 + 6\left(\frac{c_4 - c_3}{3h_3}\right)h_3 = 2c_4$$

Finally, based on:

$$g_n^{(2)}(x_{n+1}) = 0$$

we have that: $\boldsymbol{c_4 = 0}$. This gives us a system of equations:

Natural spline boundary conditions:

$$c_1 = 0$$

$$h_1c_1 + 2(h_1 + h_2)c_2 + h_2c_3 = 3\left[\frac{y_3 - y_2}{h_2} - \frac{y_2 - y_1}{h_1}\right]$$

$$h_2c_2 + 2(h_2 + h_3)c_3 + h_3c_4 = 3\left[\frac{y_4 - y_3}{h_3} - \frac{y_3 - y_2}{h_2}\right]$$

$$c_4 = 0$$

which can be written in matrix form as:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ h_1 & 2(h_1+h_2) & h_2 & 0 \\ 0 & h_2 & 2(h_2+h_3) & h_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3\left[\frac{y_3-y_2}{h_2} - \frac{y_2-y_1}{h_1}\right] \\ 3\left[\frac{y_4-y_3}{h_3} - \frac{y_3-y_2}{h_2}\right] \\ 0 \end{bmatrix}$$

If you can't see that these two representations are equivalent, remember your matrix multiplication rules:

$$\begin{bmatrix} A & B & C & D \\ E & F & G & H \\ I & J & K & L \\ M & N & O & P \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} A*c_1 + B*c_2 + C*c_3 + D*c_4 \\ E*c_1 + F*c_2 + G*c_3 + H*c_4 \\ I*c_1 + J*c_2 + K*c_3 + L*c_4 \\ M*c_1 + N*c_2 + O*c_3 + P*c_4 \end{bmatrix}$$

I then got you to solve this matrix equation for the vector of c coefficients. From these c coefficients you could solve for the remaining coefficients using the equations we developed earlier (page 5).

In Task A3, I changed the boundary conditions from natural spline to periodic spline. Remember **Equations 10 and 11 hold for any boundary conditions.** What you need to do is use the periodic boundary conditions:

$$g_1^{(1)}(x_1) = g_n^{(1)}(x_{n+1}), \qquad g_1^{(2)}(x_1) = g_n^{(2)}(x_{n+1})$$

to generate two new equations, following the same steps as above:

1. Substitute the general expressions for $g_1(x)$ and $g_n(x)$:

   $$g_1(x) = a_1 + b_1(x-x_1) + c_1(x-x_1)^2 + d_1(x-x_1)^3$$

   $$g_n(x) = a_n + b_n(x-x_n) + c_n(x-x_n)^2 + d_n(x-x_n)^3$$

   and carry out the derivatives. **I have actually done this step for you**.

2. This is where you need to do some math on your own. The next step is to substitute for any a, b, and d coefficients using the equations given on page 5 of this document. The end result should be two equations in terms of the c coefficients and the things we know (h_i and y_i).
3. Rearrange your two equations so that all the c coefficients terms are on one side and all the other terms (h_i and y_i) are on the other.
4. Represent these two new equations and Equations 10 and 11 in matrix form by generating a new H matrix and a new Y column vector
5. Solve for the c coefficients
6. Solve for the a, b, and d coefficients