Computational Mathematics/Information Technology

Dr Oliver Kerr

2009–10

Recall...

The basic problem is:

Given a set of data points $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$ how can we best construct a function, $f(x)$, that in some way approximates this information.

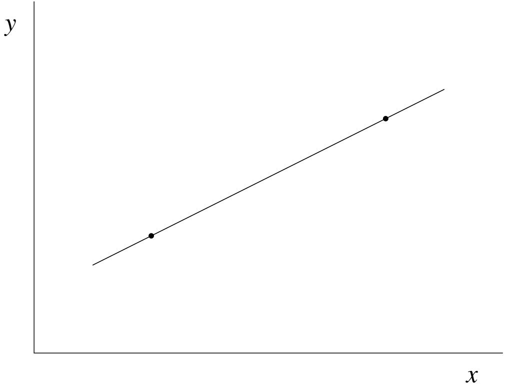
We will find that there are two basically different cases that we consider:

- The points (x_n, y_n) are accurate — we want a curve that goes through these points.
- There is some statistical scatter in the points (x_n, y_n) — we want some curve that approximates the underlying curve.

So far we have been looking at the first of these.

Linear Fit

If we only have two points



you can approximate by a straight line.

Higher Order Polynomial Fits

We can extend this to fitting a quadratic through 3 points, a cubic through 4 points, a quartic through 5 points, etc.

In general we can fit a polynomial of order n through $n+1$ data points.

Example: Given the points:

$$\{(-2,3),\ (-1,5),\ (0,4),\ (1,6),\ (3,7),\ (4,8)\}$$

construct a polynomial that passes through all the points.

Example: Given the points:

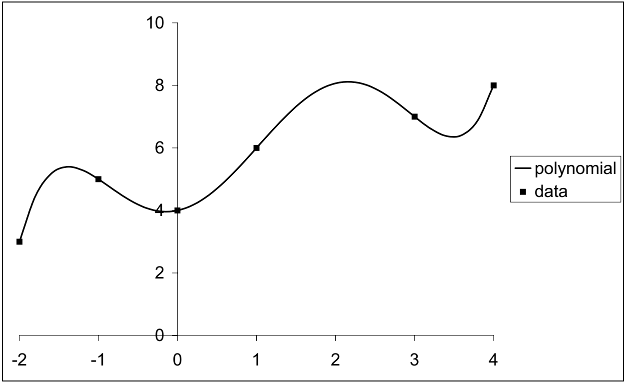
$$\{(-2,3),\ (-1,5),\ (0,4),\ (1,6),\ (3,7),\ (4,8)\}$$

construct a polynomial that passes through all the points.

We do some maths, including matrix manipulations, and find...

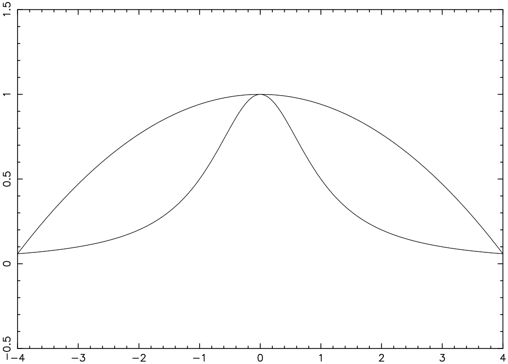
The polynomial through the six points is given by:

$$p(x) = 4 + 0.533x + 1.872x^2 - 0.106x^3 - 0.372x^4 + 0.072x^5$$

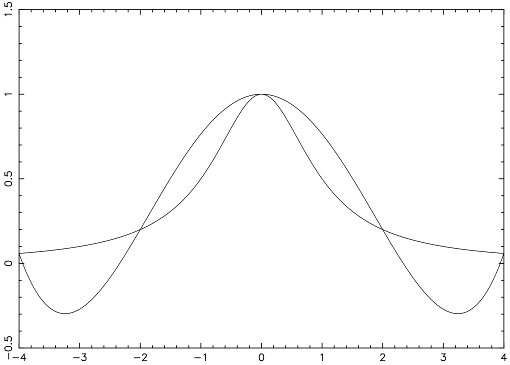


However, you don't always get what you expect!
If you approximate a perfectly sensible function with many points you don't always get the original curve.

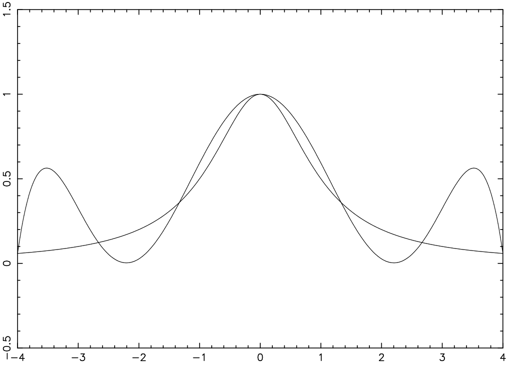
Approximating $y = 1/(1 + x^2)$ with a polynomial of order 2 between -4 and 4 with 3 evenly spaced data points:



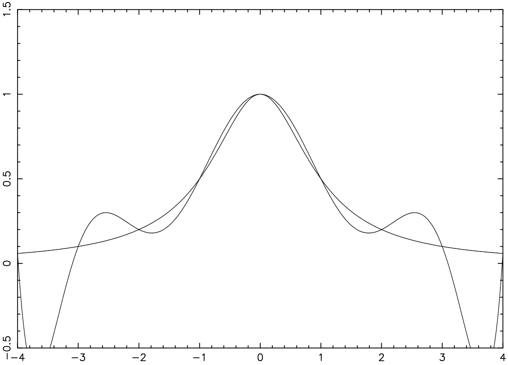
Approximating $y = 1/(1 + x^2)$ with a polynomial of order 4 between -4 and 4 with 5 evenly spaced data points:



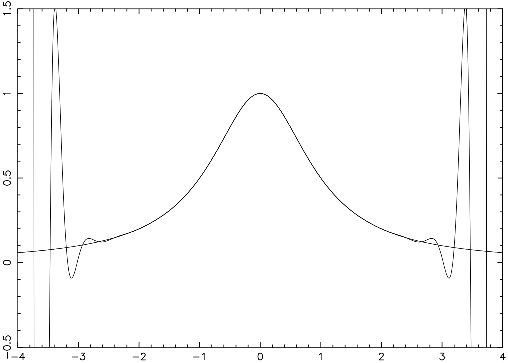
Approximating $y = 1/(1 + x^2)$ with a polynomial of order 6 between -4 and 4 with 7 evenly spaced data points:



Approximating $y = 1/(1+x^2)$ with a polynomial of order 8 between -4 and 4 with 9 evenly spaced data points:

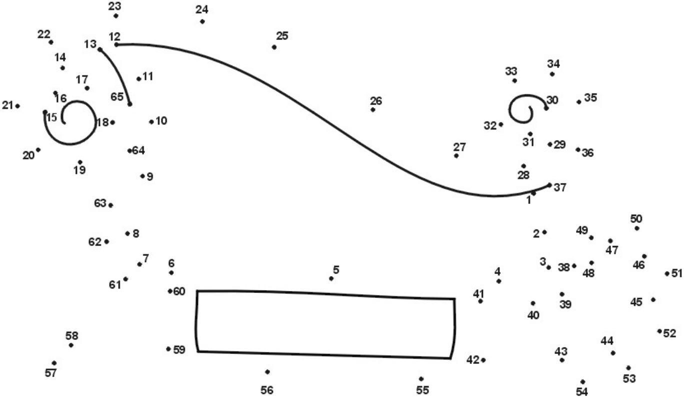


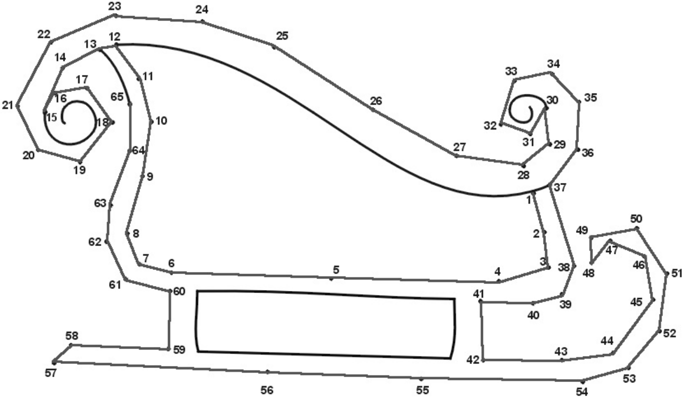
Approximating $y = 1/(1 + x^2)$ with a polynomial of order 30 between -4 and 4 with 31 evenly spaced data points:



If we were to sketch a graph through the end few points by hand we would do much better than this!

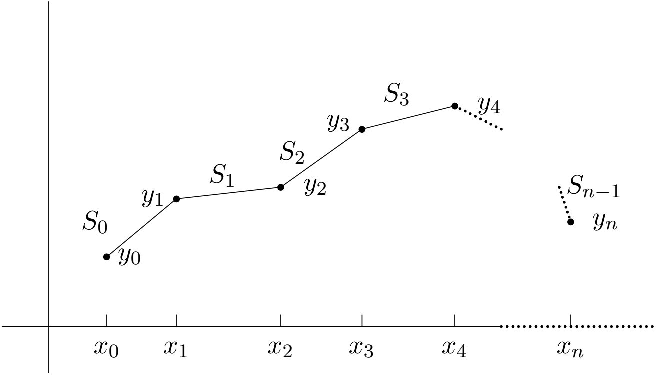
There is another approach you probably learned 10 years or so ago — the linear spline.

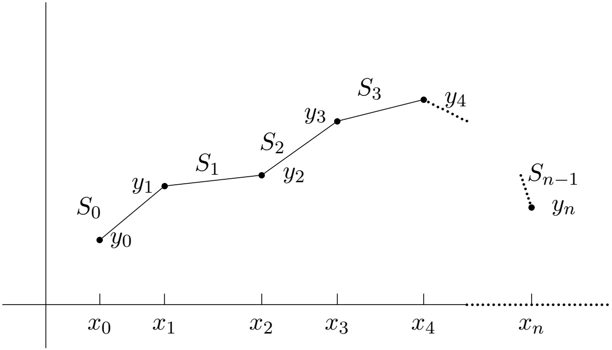

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Linear Spline

Maybe the simplest way to approximate the underlying function is to just join adjacent data points with straight lines.





Given the data points $\{(x_0, y_0), (x_1, y_1), \ldots (x_n, y_n)\}$ we join adjacent points with a straight line. The straight line between (x_k, y_k) and (x_{k+1}, y_{k+1}) is denoted by $y = S_k(x)$ and the complete set of all line segments by $S(x)$. It is this complete set of all line segments that is referred to as the **linear spline** through the points.

Advantages:

- Simple and robust.
- Looks OK if you have enough points.

Disadvantages:

- Curve has kinks in it — not differentiable.
- Approximation is only affected by the nearest points on either side.

New Stuff!

Cubic Spline

Instead of using a straight line between adjacent points we will improve things by having a cubic between adjacent points.
As well as passing through the points we will also require the curves to have continuous first and second derivatives at the data points.

A cubic

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

has three unknowns.

- To find 4 unknowns we need 4 equations.

A cubic

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

has three unknowns.

- To find 4 unknowns we need 4 equations.
- Requiring the cubic to go through the end-points supplies two of these equations.

A cubic

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

has three unknowns.

- To find 4 unknowns we need 4 equations.
- Requiring the cubic to go through the end-points supplies two of these equations.
- Requiring the first and second derivatives match at each end seems to give 4 more equations. Too many?

A cubic

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

has three unknowns.

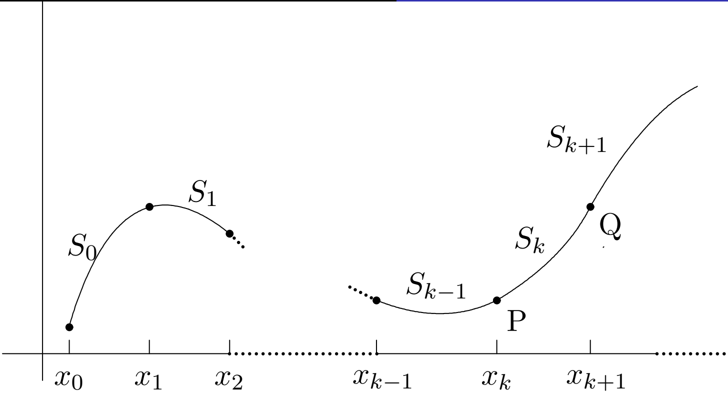
- To find 4 unknowns we need 4 equations.
- Requiring the cubic to go through the end-points supplies two of these equations.
- Requiring the first and second derivatives match at each end seems to give 4 more equations. Too many? But each of these is shared with the adjacent interval, and so effectively only provides two more equations as required.
- But we have to look at the end intervals more carefully.

The end intervals need one constraint applied at the unattached or free ends.
We impose the condition that the second derivative must be zero.

These end conditions are what classify the spline as a

natural cubic spline.

Other end conditions lead to other types of cubic spline.



We denote by $S_k(x)$ the cubic joining the point (x_k, y_k) to the point (x_{k+1}, y_{k+1}) (points P to Q in the above figure).

Spline conditions

(i) $S_k(x)$ must pass through P and Q, and so

$$S_k(x_k) = y_k \quad \text{and} \quad S_k(x_{k+1}) = y_{k+1}, \quad k = 0, \ldots, (n-1)$$

(ii) At P $\quad S'_k(x_k) = S'_{k-1}(x_k), \quad k = 1, \ldots, (n-1)$

(iii) At P $\quad S''_k(x_k) = S''_{k-1}(x_k), \quad k = 1, \ldots, (n-1)$

(iv) To produce a natural cubic spline we set the second derivatives equal to zero at the ends of the interval. Thus

$$S''_0(x_0) = S''_{n-1}(x_n) = 0$$

There are no conditions on the first derivatives at these point.

Condition (i): $S_k(x)$ must pass through P and Q, and so

$$S_k(x_k) = y_k \quad \text{and} \quad S_k(x_{k+1}) = y_{k+1}, \quad k = 0, \ldots, (n-1)$$

merely ensures that the completed spline passes through all the data points.

Condition (i): $S_k(x)$ must pass through P and Q, and so

$$S_k(x_k) = y_k \quad \text{and} \quad S_k(x_{k+1}) = y_{k+1}, \quad k = 0, \ldots, (n-1)$$

merely ensures that the completed spline passes through all the data points.

Condition (ii) At P $\quad S'_k(x_k) = S'_{k-1}(x_k), \quad k = 1, \ldots, (n-1)$
ensures that at a point where two cubics from adjacent sections meet, say at P, they do so smoothly in the sense that the two have the same tangent at the point of contact.

Condition (iii) At P $\quad S''_k(x_k) = S''_{k-1}(x_k), \quad k = 1, \dots, (n-1)$
ensures that at a point where two cubics from adjacent sections meet, say at P, they both have the same second derivative (and so same curvature) at the point of contact.

Condition (iii) At P $\quad S_k''(x_k) = S_{k-1}''(x_k), \quad k = 1, \ldots, (n-1)$
ensures that at a point where two cubics from adjacent sections meet, say at P, they both have the same second derivative (and so same curvature) at the point of contact.

Condition (iv) $S_0''(x_0) = S_{n-1}''(x_n) = 0$
has the effect of setting the curvature of the spline equal to zero at the end points.
In many problems this is a reasonable thing to do. There is no more obvious condition(?).

At first sight the construction of all the components $S_k(x)$ appears quite formidable — finding the coefficients for n cubics at the same time (i.e., solve $4n$ simultaneous equations).
However a great deal of ingenuity has been put into a method for making the problem quite straightforward.

The first trick is to assume that the cubic is written in a certain way: instead of simply writing the cubic in the standard form $a + bx + cx^2 + dx^3$ we express each $S_k(x)$ as:

$$S_k(x) = a_k + b_k(x - x_k) + c_k(x - x_k)^2 + d_k(x - x_k)^3, \quad k = 0, \ldots, (n-1)$$

- At P condition (i) implies

$$S_k(x_k) = y_k = a_k \qquad k = 0, \ldots, (n-1)$$

  Thus all the a coefficients are obtained immediately. This gives n equations.

The first trick is to assume that the cubic is written in a certain way: instead of simply writing the cubic in the standard form $a + bx + cx^2 + dx^3$ we express each $S_k(x)$ as:

$$S_k(x) = a_k + b_k(x - x_k) + c_k(x - x_k)^2 + d_k(x - x_k)^3, \quad k = 0, \ldots, (n-1)$$

- At P condition (i) implies

  $$S_k(x_k) = y_k = a_k \qquad k = 0, \ldots, (n-1)$$

  Thus all the a coefficients are obtained immediately. This gives n equations.
  At Q condition (i) implies

  $$S_k(x_{k+1}) = y_{k+1} = a_k + b_k(x_{k+1} - x_k) + c_k(x_{k+1} - x_k)^2 + d_k(x_{k+1} - x_k)^3$$

  This gives n equations.

- Condition (ii), after differentiating $S_k(x)$ and $S_{k-1}(x)$ gives:

$$S'_k(x_k) = S'_{k-1}(x_k)$$

and so

$$b_k = b_{k-1} + 2c_{k-1}(x_k - x_{k-1}) + 3d_{k-1}(x_k - x_{k-1})^2 \quad k = 1, \ldots, (n-1)$$

This gives $(n-1)$ equations.

- Condition (iii), after differentiating $S_k(x)$ and $S_{k-1}(x)$ twice gives:

$$S''_k(x_k) = S''_{k-1}(x_k)$$

and so

$$2c_k = 2c_{k-1} + 6d_{k-1}(x_k - x_{k-1}) \quad k = 1, \ldots, (n-1)$$

This gives $(n-1)$ equations.

- Imposing condition (iv):

$$S_0''(x_0) = 0 \quad \Rightarrow \quad 2c_0 = 0$$

and

$$S_{n-1}''(x_n) = 0 \quad \Rightarrow \quad 2c_{n-1} + 6d_{n-1}(x_n - x_{n-1}) = 0$$

This gives 2 equations.

We now have the $4n$ equations we require.

We note that our conditions have generated $4n$ equations, precisely the same as the number of coefficients. It should therefore be possible for us to solve these equations and hence construct $S(x)$. The following example outlines the general method for solving the equations. Since the equations always take the same form a systematic approach is possible for all spline problems; we do not have to solve the completely general simultaneous equation problem of $4n$ equations in $4n$ unknowns.

Example: Construct the natural cubic spline through the three points $\{(1,1),(2,-1),(4,3)\}$.
The spline $S(x)$ is given by:

$$S(x) = \begin{cases} S_0(x) = a_0 + b_0(x-1) + c_0(x-1)^2 + d_0(x-1)^3 & 1 \le x \le 2 \\ S_1(x) = a_1 + b_1(x-2) + c_1(x-2)^2 + d_1(x-2)^3 & 2 \le x \le 4 \end{cases}$$

Applying condition (i):

- $S_0(1) = 1 \quad \Rightarrow \quad a_0 = 1$
- $S_0(2) = -1 \quad \Rightarrow \quad -1 = a_0 + b_0 + c_0 + d_0$
- $S_1(2) = -1 \quad \Rightarrow \quad a_1 = -1$
- $S_1(4) = 3 \quad \Rightarrow \quad 3 = a_1 + 2b_1 + 4c_1 + 8d_1$

Applying condition (ii):

- $S_1'(2) = S_0'(2) \quad \Rightarrow \quad b_1 = b_0 + 2c_0 + 3d_0$

Applying condition (iii):

- $S_1''(2) = S_0''(2) \quad \Rightarrow \quad 2c_1 = 2c_0 + 6d_0$

Applying condition (iv)

- $S_0''(1) = 0 \quad \Rightarrow \quad 2c_0 = 0$
- $S_1''(4) = 0 \quad \Rightarrow \quad 2c_1 + 12d_1 = 0$

We now consider the solution of these eight equations in the following systematic fashion:

- In all cases the a_k coefficients are given straight away by $a_k = y_k$, and c_0 is always zero for the natural cubic spline. Thus in general we immediately have $(n+1)$ of the unknowns. In this example:

$$c_0 = 0 \qquad a_0 = 1 \qquad a_1 = -1$$

- From the conditions on $S_k''(x)$ we can always write the d coefficients in terms of the c coefficients. In this example:

$$d_0 = \frac{2c_1}{6} = \frac{c_1}{3} \qquad d_1 = -\frac{2c_1}{12} = -\frac{c_1}{6}$$

- From the results of condition (i) and substituting in the known a values and $c_0 = 0$, we can always write the b coefficients in terms of the c and d coefficients. If we now substitute for the ds in terms of the cs, from above, we obtain the bs in terms of the cs (and the as that we already know). In this example:

$$b_0 = -2 - c_0 - d_0 = -2 - \frac{c_1}{3}$$

$$b_1 = 2 - 2c_1 - 4d_1 = 2 - 2c_1 + \frac{2c_1}{3} = 2 - \frac{4c_1}{3}$$

- At this point we have all the bs and ds in terms of the cs. We now substitute for the bs and ds into the equations formed from condition (ii) (i.e., the constraints on $S_k'(x)$.) These equations will in general be solvable for the cs. By substituting back we can them calculate the bs and ds. In this example:

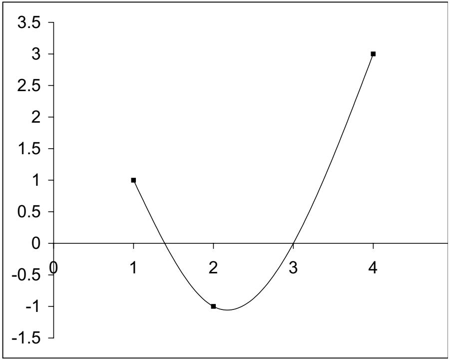
$$b_1 = b_0 + 2c_0 + 3d_0 \quad \Rightarrow \quad 2 - \frac{4c_1}{3} = -2 - \frac{c_1}{3} + 3\frac{c_1}{3} \quad \Rightarrow \quad c_1 = 2$$

- Finally substituting back with $c_1 = 2$ gives

$$b_1 = -\frac{2}{3} \qquad b_0 = -2 - \frac{2}{3} = -\frac{8}{3} \qquad d_1 = -\frac{1}{3} \qquad d_0 = \frac{2}{3}$$

The natural cubic spline $S(x)$ is given by:

$$S(x) = \begin{cases} S_0(x) = 1 - \frac{8}{3}(x-1) + \frac{2}{3}(x-1)^3 & 1 \le x \le 2 \\ S_1(x) = -1 - \frac{2}{3}(x-2) + 2(x-2)^2 - \frac{1}{3}(x-2)^3 & 2 \le x \le 4 \end{cases}$$



So the general plan of attack is:

1. Find all the *a*s (trivial)
2. Find expressions for the *d*s in terms of the *c*s using the second derivative expressions.
3. Find expressions for the *b*s in terms of *c*s and *a*s (which you know) using the expressions for continuity at the right of each interval and your expressions for the *d*s.
4. Substitute all this into the expressions for the continuity of the first derivative, giving $n-1$ expressions for the *c*s.
5. Solve the equations for the *c*s, remembering $c_0 = 0$.
6. Now find all the *b*s and *d*s — job done.