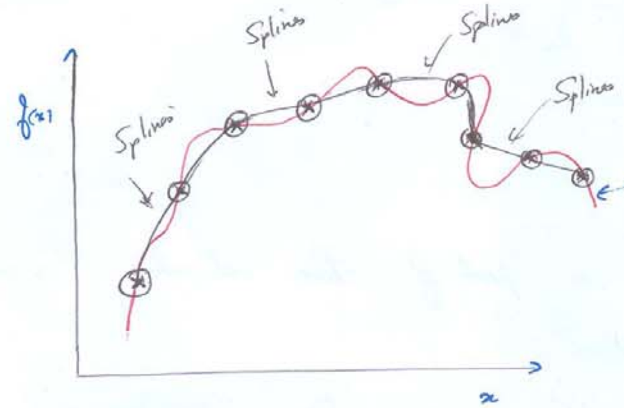
## Lecture 24: Cubic splines

**(26-Spt-2012)**

LECTURE 24
26 - SEPT - 2012

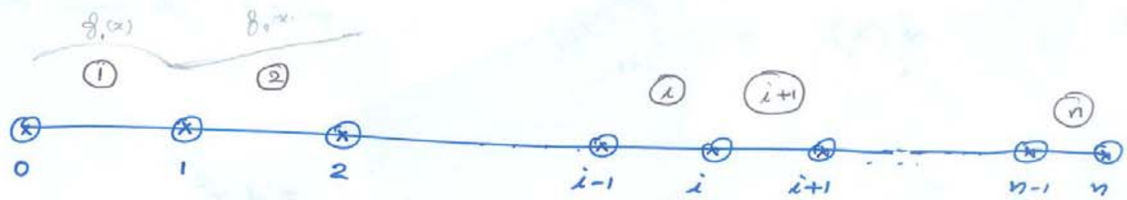
### Cubic Splines

In the last class we discussed that for a given set of data points, rather than going for higher order polynomial approximation, it is better to go for piece-wise approximation using splines.




Splines give piece-wise approximation

On a data line




There are $(n+1)$ data points
$n$ intervals.

→ We want to give cubic spline approximation to the function at each interval $i = 1, 2, 3, \ldots, n$.

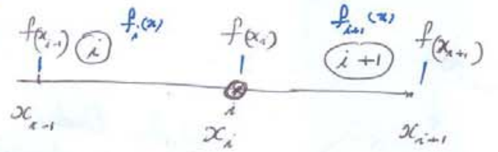
$$f_i(x) = a_i + b_i x + c_i x^2 + d_i x^3$$

$$x_{i-1} \le x \le x_i$$

The conditions for cubic ~~pol~~ spline approximations are:

(i) The function values at data points are known.

$\therefore$ $f(x_i)$ is available to you.




→ The point $x_i$ is part of two intervals ($i$) and ($i+1$).

→ $\therefore$ Whatever cubic spline approximations you are giving to intervals ($i$) and ($i+1$), it should satisfy

$$f_i(x_i) = f_{i+1}(x_i) = f(x_i) \text{ the known value}$$

(ii) Now $f_i'(x) = b_i + 2c_i x + 3d_i x^2$

→ The first derivative of cubic spline at the point $x_i$ should be such that

$$f_i'(x_i) = f_{i+1}'(x_i)$$

(iii) The second derivative of $f_i(x)$

$$f_i''(x) = 2c_i + 6d_i x$$

This is a linear function in $x$ in the interval $i$.

Now the condition is such that:

$$f_i''(x_i) = f_{i+1}''(x_i)$$

(iv) The first spline i.e. $f_1(x)$ should pass through $x = x_0$

The last spline i.e. $f_n(x)$ should pass through $x = x_n$

(v) The curvature or second derivative of $f(x)$ should be specified at $x = x_0$ and $x = x_n$.

i.e. $f_1''(x_0) = f''(x_0)$

and $f_n''(x_n) = f''(x_n)$

$\Rightarrow$ The second derivative of cubic spline at any interval $i$

$$f_i''(x) = 2c_i + 6d_i x$$

$x_{i-1} \leq x \leq x_i$

This linear function of $x$ can be represented in term of Lagrange polynomial ($1^{st}$ order).

i.e. $$f_i''(x) = \frac{x - x_i}{x_{i-1} - x_i} f_i''(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} f_i''(x_i)$$

Now actually in the intermediate intervals
$f_i''(x_{i-1})$ and $f_i''(x_i)$ may be unknown to you.

However, you know $f_i''(x_{i-1}) = f_{i-1}''(x_{i-1})$

$$\therefore\ f_i''(x) = \frac{x - x_i}{x_{i-1} - x_i} f_{i-1}''(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} f_i''(x_i)$$

or

$$f_i''(x) = \frac{x - x_i}{x_{i-1} - x_i} f_i''(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} f_{i+1}''(x_i)$$

If you begin with $i = 1$ i.e. first interval
$f_1''(x_0)$ is a known quantity to you.

Now $$\int f_i''(x)\,dx = f_i'(x) = \frac{\frac{x^2}{2} - x x_i}{x_{i-1} - x_i} f_{i-1}'' + \frac{\frac{x^2}{2} - x x_{i-1}}{x_i - x_{i-1}} f_i'' + C$$

and $$f_i(x) = \frac{\frac{x^3}{6} - \frac{x^2}{2} x_i}{x_{i-1} - x_i} f_{i-1}''(x_{i-1}) + \frac{\frac{x^3}{6} - \frac{x^2}{2} x_i}{x_i - x_{i-1}} f_i''(x_i) + Cx + D$$

Apply this expression for $f_i(x)$ at the known points $x = x_{i-1}$ and $x = x_i$, where the function values are $f(x_{i-1})$ and $f(x_i)$ (known to you)

→ That way you can eliminate C and D in the expression for $f_i(x)$.

→ You have to note that $f''_{i-1}(x_{i-1}) = f''_i(x_{i-1})$.

Similarly $f'_{i-1}(x_{i-1}) = f'_i(x_{i-1})$

Again $f'_i(x_i) = f'_{i+1}(x_i)$

Using all these relations one can finally arrive at a relation :

→ You get a cubic spline expression for interval (i) as:

$$\boxed{\begin{aligned} f_i(x) = {} & \frac{f''_{i-1}}{6(x_i - x_{i-1})}(x_i - x)^3 + \frac{f''_i}{6(x_i - x_{i-1})}(x - x_{i-1})^3 + (x_i - x)\left\{\frac{f_{i-1}}{(x_i - x_{i-1})} - \frac{(x_i - x_{i-1})f''_{i-1}}{6}\right\} \\ & + (x - x_{i-1})\left\{\frac{f_i}{(x_i - x_{i-1})} - \frac{(x_i - x_{i-1})f''_i}{6}\right\} \end{aligned}}$$

→ Here you have unknowns $f''_{i-1}(x_{i-1})$ and $f''_i(x_i)$

→ As suggested earlier use the relations $f''_{i-1}(x_{i-1}) = f''_i(x_{i-1})$, etc.

Differentiate $f_i(x)$ once to get $f_i'(x)$ and also considering $f'_{i-1}(x_{i-1}) = f'_i(x_{i-1})$ and going on: we finally get:

$$\boxed{(x_i - x_{i-1})\, f''_{i-1}(x_{i-1}) + 2(x_{i+1} - x_{i-1})\, f''_i(x_i) + (x_{i+1} - x_i)\, f''_{i+1}(x_{i+1}) = 6\,\frac{f_{i+1} - f_i}{x_{i+1} - x_i} - 6\,\frac{f_i - f_{i-1}}{x_i - x_{i-1}}}$$

⇒ So note that at each interval (i), you need to apply the above expression for cubic spline.

* A natural spline is one which have

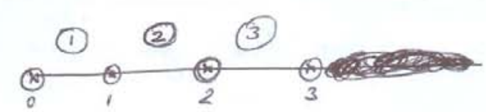
$$f_1''(x_0) = 0.0 \quad \text{and} \quad f_n''(x_n) = 0.0$$

Example

For the given data set provide cubic spline fit. (Hoffman 2001)

| $i$ | $x$ | $f(x)$ | $f''(x)$ |
|---|---|---|---|
| 0 | -0.50 | 0.731531 | 0.0 |
| 1 | 0.00 | 1.00000 | |
| 2 | 0.25 | 1.268400 | |
| 3 | 1.00 | 1.718282 | 0.0 |

Soln: There are 4 data points




→ There are three intervals. i.e. $f_1(x)$, $f_2(x)$, and $f_3(x)$ three cubic splines are to be formed.

Sol: Given $y_0 = 0.731531$, $y_1 = 1.0000$

$$y''(x_0) = 0.00$$

$\therefore$ For $y_1(x)$, $-0.50 \le x \le 0.00$

$$(x_1 - x_0) y_0'' + 2(x_2 - x_0) y_1'' + (x_2 - x_1) y_2'' = 6 \frac{y_2 - y_1}{x_2 - x_1} - 6 \frac{y_1 - y_0}{x_1 - x_0}$$

i.e. $0.5 \times 0.0 + 2 \times 0.75 \times y_1'' + 0.25 \times y_2'' = 3.219972$

Illy for $y_2(x)$

$$(x_2 - x_1) y_1'' + 2(x_3 - x_1) y_2'' + (x_3 - x_2) y_3'' = 6 \frac{y_3 - y_2}{x_3 - x_2} - 6 \frac{y_2 - y_1}{x_2 - x_1}$$

i.e. $0.25 \times y_1'' + 2(1.00) y_2'' + 0.75 \times 0.0 = -2.842544$

Solve for $y_1''$ and $y_2''$

$y_1'' = 2.434240$ and $y_2'' = -1.725552$

So now you will get a cubic spline for interval (1)

$$y_1(x) = y_0'' \times \frac{(\ )}{(\ )} \cdots + \frac{y_1''(x - x_0)^3}{6(x_1 - x_0)} + (x_1 - x)\left\{\frac{y_0}{(x_1 - x_0)} - \frac{(x_1 - x_0) y_0''}{6}\right\}$$

$$+ (x - x_0)\left\{\frac{y_1}{(x_1 - x_0)} - \frac{(x_1 - x_0)}{6} y_1''\right\}$$

---

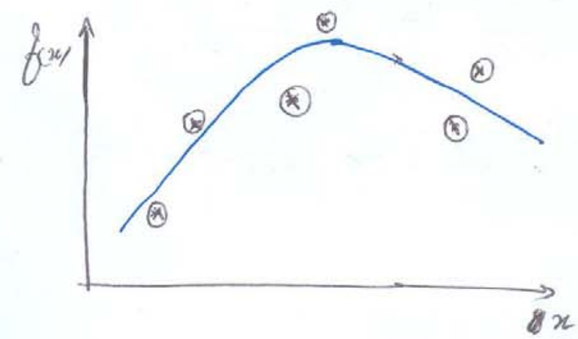
Similarly obtain $y_2(x)$ and $y_3(x)$ cubic splines as well.

## Approximate Fits

As discussed earlier the polynomial approximations you are using may pass through all the data points to give exact fits to the data-points.

However such polynomials may not be accurate as it may give erroneous results b/w the data points.

We had talked about approximate fits.

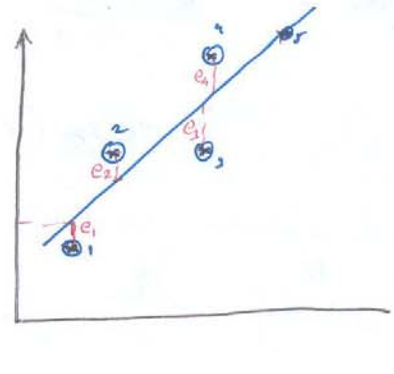



How will form approximate fits ?

A ~~partcle~~ particular method is Least Squares Approximation.

## Least Square Approximation

The objective of this method is to minimize the sum of of the squares of the deviations.



The deviation is given as the difference b/w fitted value and observed value.

$e_i =$

" For n data points $(x_i, \mathscr{Y}_i)$, you may approximate the functional form as $\mathscr{Y} \approx y$ and then minimize the square of the error

$$e_i = \mathscr{Y}_i - y_i$$

## Straight Line Approx.

If for given $(n+1)$ data points

| $i$ | $x$ | $\mathscr{Y}$ | $y$ | $e$ |
|---|---|---|---|---|
| 0 | $x_0$ | $\mathscr{Y}_0$ | $y_0$ | $e_0$ |
| 1 | $x_1$ | $\mathscr{Y}_1$ | $y_1$ | $e_1$ |
| 2 | $x_2$ | $\mathscr{Y}_2$ | $y_2$ | $e_2$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n$ | $x_n$ | $\mathscr{Y}_n$ | $y_n$ | $e_n$ |

→ We can fit a best straight line as

$$y = A + Bx$$

→ Now at each data point, you can relate

$$y_i = A + Bx_i$$

The error (or deviation) $e_i = \mathscr{Y}_i - y_i$

$$y_i = A + Bx_i$$
$$e_i = \mathscr{Y}_i - y_i$$

As you are not aware of $A$ and $B$

$$e_i \rightarrow e_i(A, B)$$

| $i$ | $x$ | $\mathscr{Y}$ | $y$ | $e$ | $e^2$ |
|---|---|---|---|---|---|
| 0 | $x_0$ | $\mathscr{Y}_0$ | $y_0$ | $\mathscr{Y}_0 - y_0$ | $(\mathscr{Y}_0 - y_0)^2$ |
| 1 | $x_1$ | $\mathscr{Y}_1$ | $y_1$ | $\mathscr{Y}_1 - y_1$ | $(\mathscr{Y}_1 - y_1)^2$ |
| 2 | $x_2$ | $\mathscr{Y}_2$ | $y_2$ | $\mathscr{Y}_2 - y_2$ | $(\mathscr{Y}_2 - y_2)^2$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n$ | $x_n$ | $\mathscr{Y}_n$ | $y_n$ | $\mathscr{Y}_n - y_n$ | $(\mathscr{Y}_n - y_n)^2$ |

As we have to minimise the sum of square of deviation

$$\cancel{S = (f_0 - y_0)^2 +} \qquad S = \sum_{i=0}^{n} (f_i - y_i)^2$$

Now $S \to S(A, B)$

$\therefore$ To minimise $S$ we can say

$$\frac{\partial S}{\partial A} = 0 \quad \text{and} \quad \frac{\partial S}{\partial B} = 0$$

As $y_i = A + Bx_i$

$$S = \sum_{i=0}^{n} (f_i - A - Bx_i)^2$$

$$\therefore \frac{\partial S}{\partial A} = \sum_{i=0}^{n} 2(f_i - A - Bx_i).(-1) = 0$$

$$\text{i.e.} \quad \sum_{i=0}^{n} f_i - A\sum_{i=0}^{n}(1) - B\sum_{i=0}^{n} x_i = 0$$

$$\text{III}^{ly} \quad \frac{\partial S}{\partial B} = \sum_{i=0}^{n} 2(f_i - A - Bx_i)(-x_i) = 0$$

$$\text{i.e.} \quad -\sum_{i=0}^{n} f_i x_i + A\sum_{i=0}^{n} x_i + B\sum_{i=0}^{n} x_i^2 = 0$$

or if $\sum_{i=0}^{n} \cancel{1} = \cancel{0 + 1 + 2 + 3 + \dots +} N = (n+1)$ the total data points.

$$\therefore \boxed{\begin{aligned} AN + B\sum_{i=0}^{n} x_i &= \sum_{i=0}^{n} f_i \\ A\sum_{i=0}^{n} x_i + B\sum_{i=0}^{n} x_i^2 &= \sum_{i=0}^{n} f_i x_i \end{aligned}}$$

These are normal equations for a straight line approximation.

On solving normal eqns. you will get $A$ and $B$.