NUMERICAL METHODS
Week-7
25.03.2014

Approximation by Spline functions & Smoothing of Data

Asst. Prof. Dr. Berk Canberk

Approximation by Spline functions

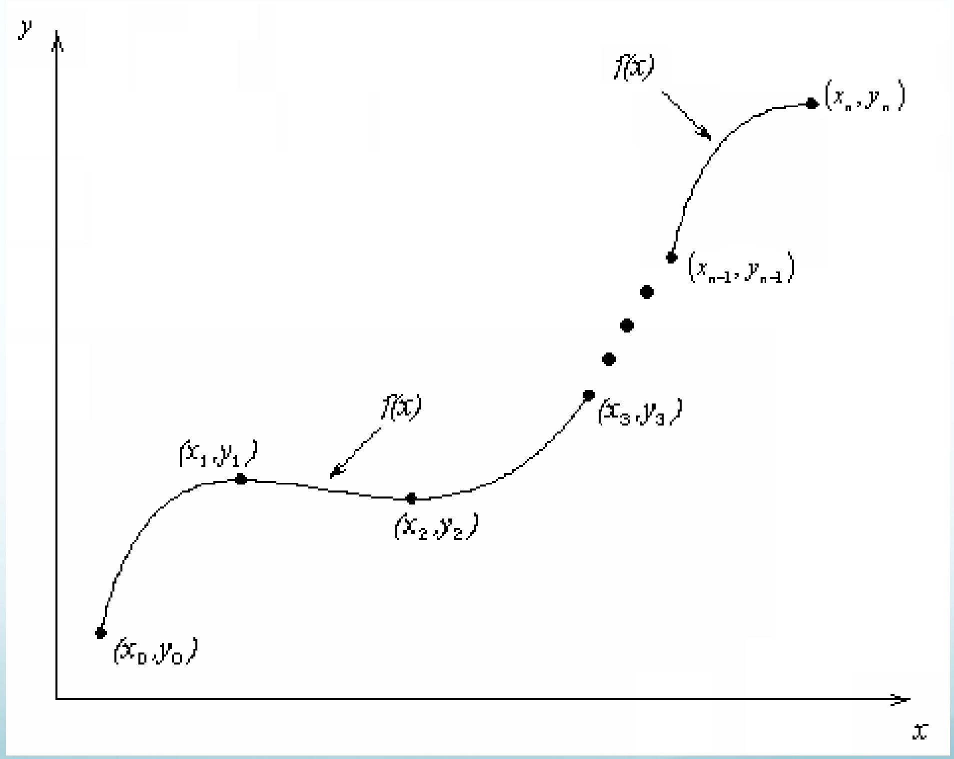
- **What** is a spline function?
- **Why** do we use approximation by splines?
- **How** do we solve spline equations?

What is a Spline?

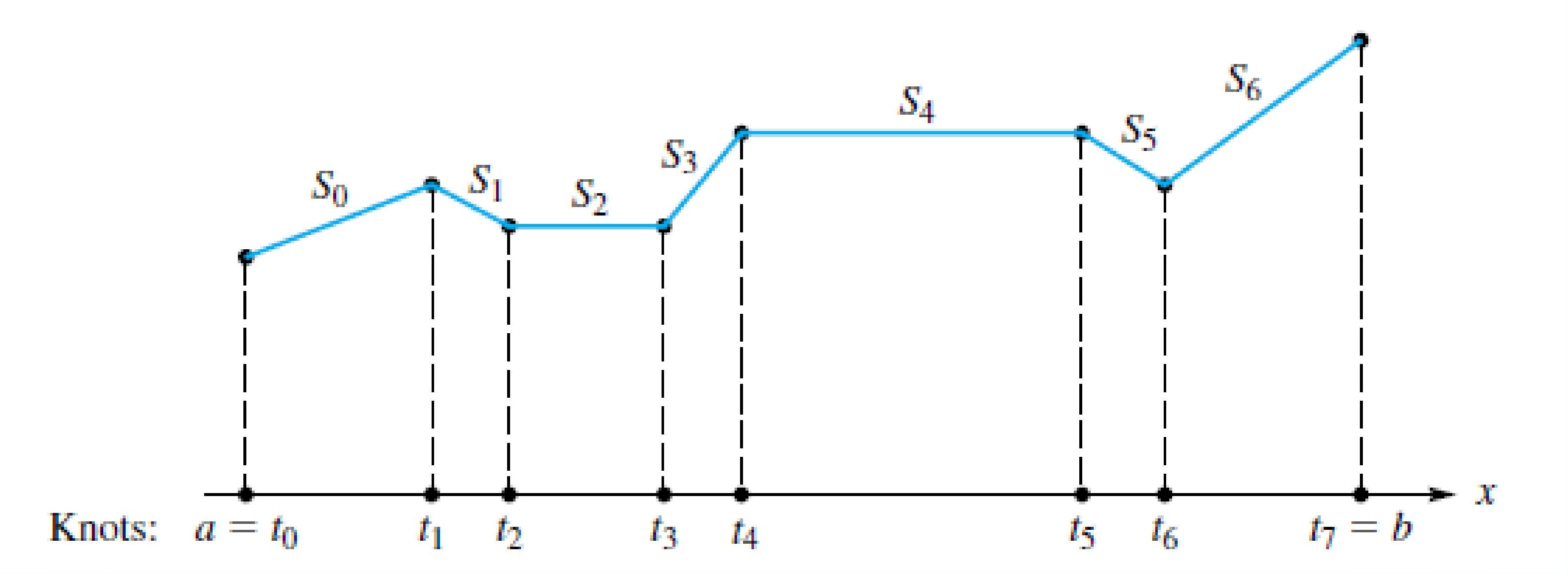
- A spline is a function that consisting of simple functions joined together.
- As with polynomial functions, splines are used to interpolate tabulated data as well as functions.
- A spline is different from a polynomial interpolation, which consists of a single well defined function that approximates a given shape; splines are normally **piecewise polynomial**.

What is a spline function ?

Given (x_0,y_0), (x_1,y_1), …… (x_n,y_n), find the value of ‘y’ at a value of ‘x’ that is not given.



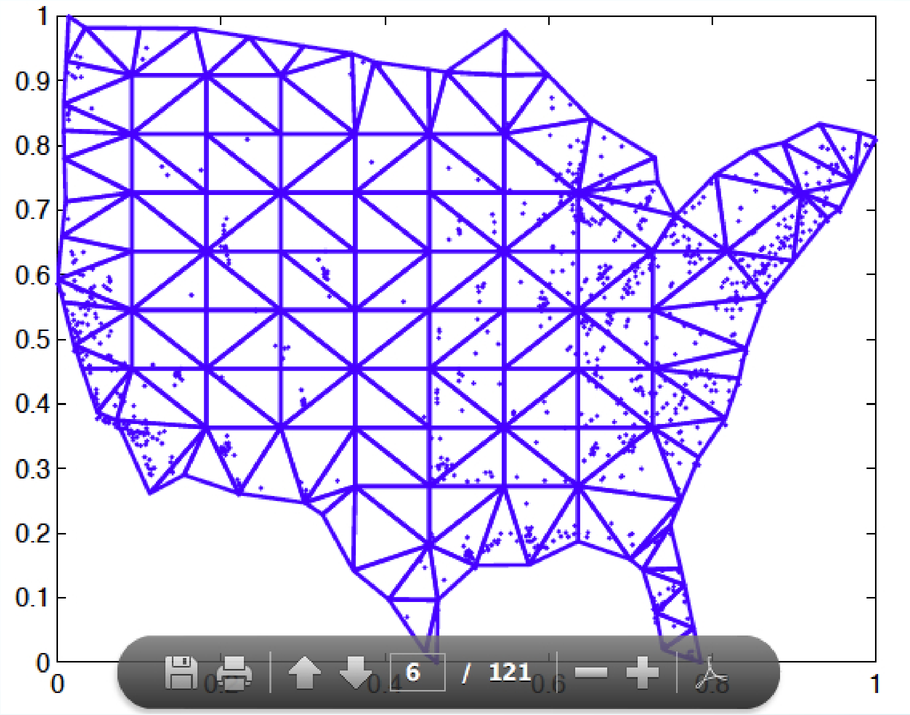
What is a spline function ?



- A simple example is the **polygonal** function (or spline of degree 1), whose pieces are linear polynomials joined together to achieve **continuity**, as in figure. The points $t_0, t_1, \ldots, t_n$ at which the function changes its character are termed **knots** in the theory of splines.

Some Applications

- Ship, aircraft building
- Detailed designs
- Computer Graphics
- Routing
- Movement Estimation



Why Splines ?

- Splines are used to approximate complex functions and shapes.
- Drawbacks of higher order polynomials in interpolating functions.
- Splines are normally piecewise polynomials so provides better approximation then polynomial interpolatings.

Why Splines ?

$$f(x) = \frac{1}{1+25x^2}$$

Table : Six equidistantly spaced points in [-1, 1]

x	$y = \frac{1}{1+25x^2}$
-1.0	0.038461
-0.6	0.1
-0.2	0.5
0.2	0.5
0.6	0.1
1.0	0.038461

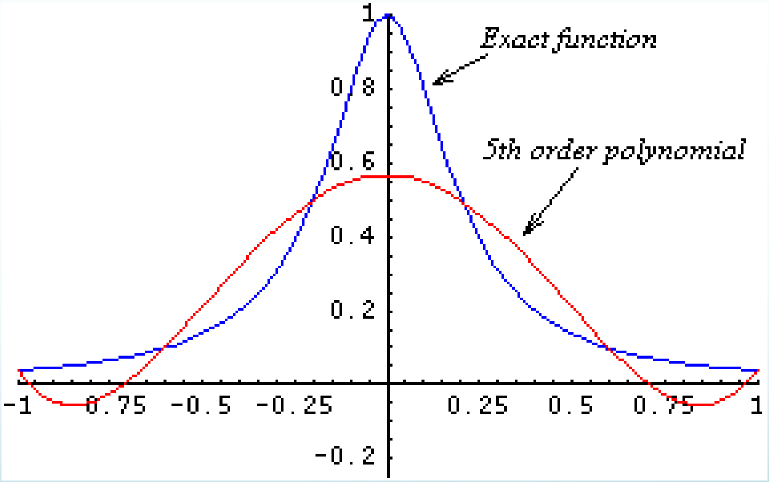


Figure : 5th order polynomial vs. exact function

Why Splines ?

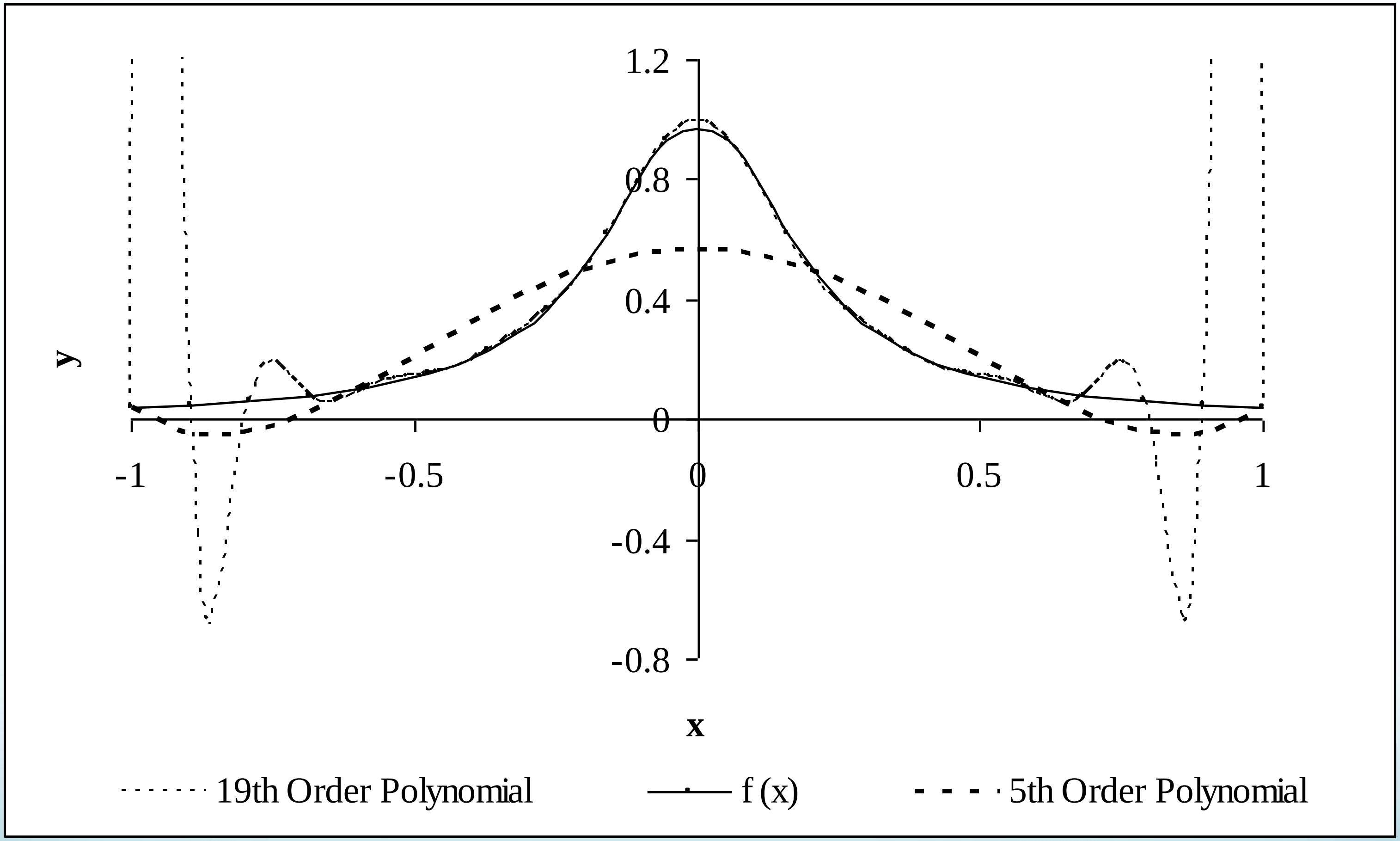
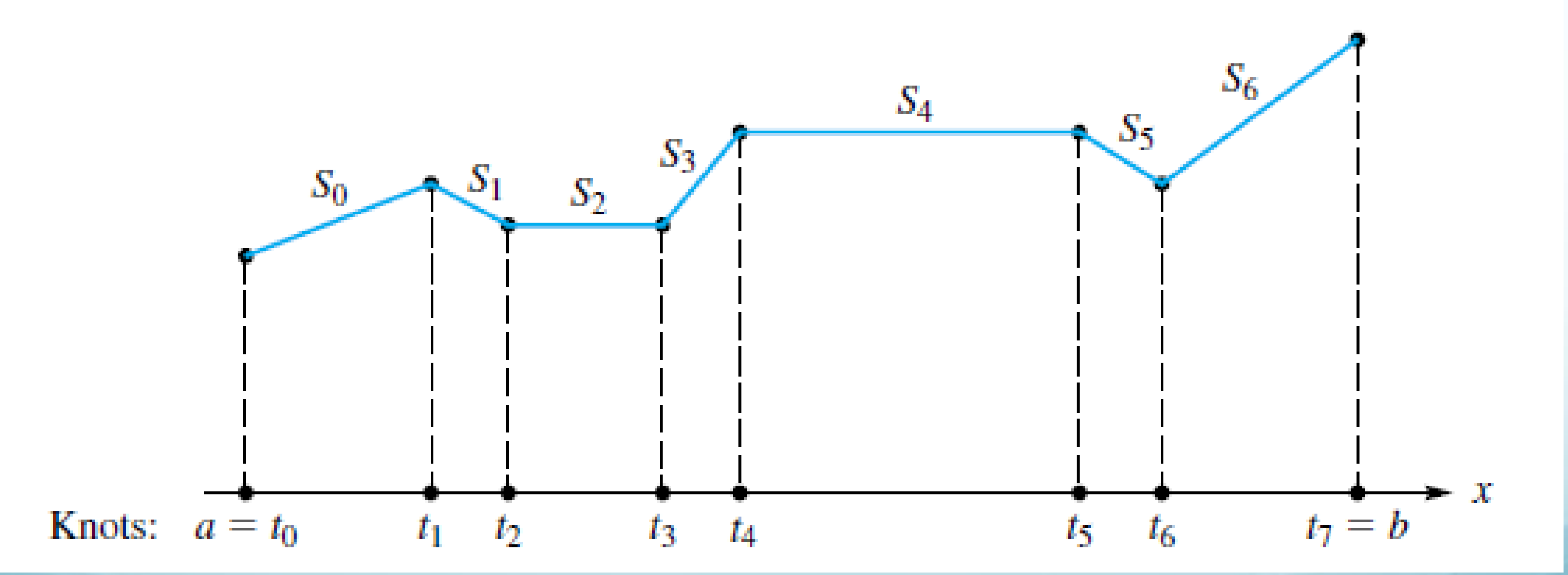


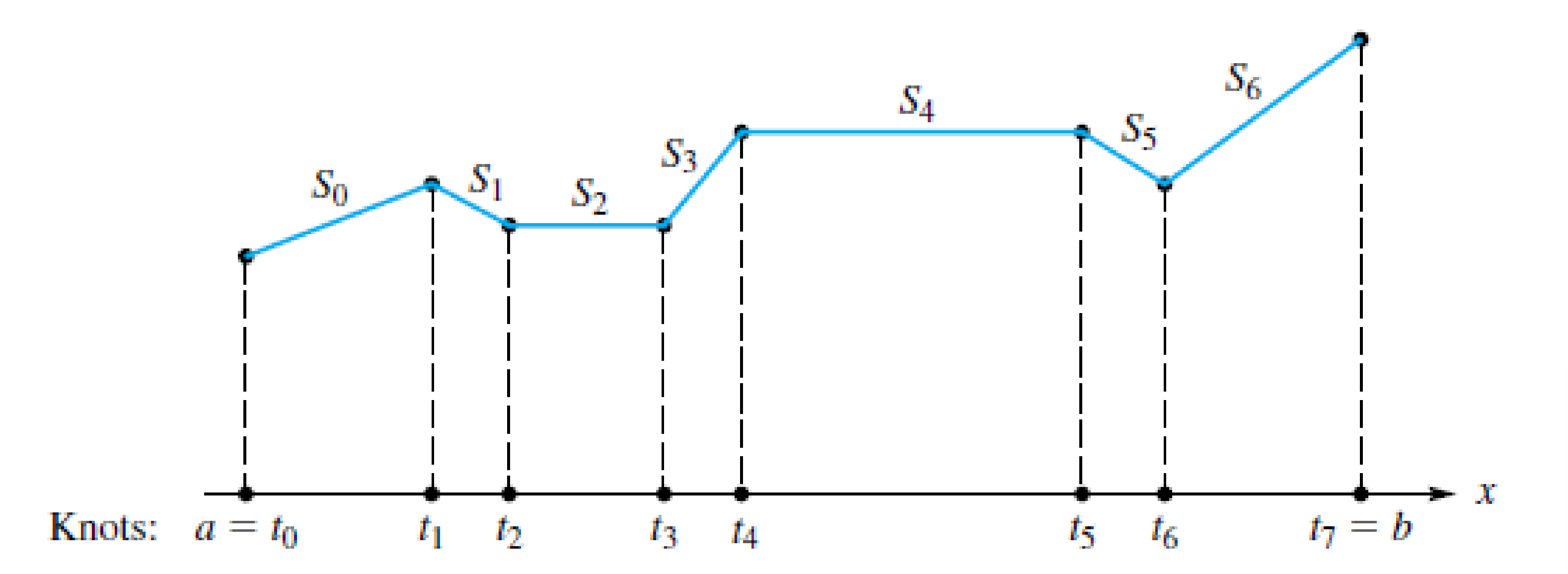
Figure : Higher order polynomial interpolation is a bad idea

First Degree Splines

- Splines make use of partitions, which are a way of cutting an interval into a number of subintervals.
- The spline functions of degree 1 can be used for interpolation



First Degree Splines



A function S is called a **spline of degree 1** if:

1. The domain of S is an interval $[a, b]$.
2. S is **continuous** on $[a, b]$.
3. There is a partitioning of the interval $a = t_0 < t_1 < \cdot \cdot \cdot < t_n = b$ such that S is a linear polynomial on each subinterval $[t_i , t_{i+1}]$.

First Degree Splines

$$f(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0), \qquad x_0 \le x \le x_1$$

$$= f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1), \qquad x_1 \le x \le x_2$$

$$.$$

$$.$$

$$.$$

$$= f(x_{n-1}) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x - x_{n-1}), \quad x_{n-1} \le x \le x_n$$

Note the terms of

$$\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

in the above function are simply slopes between x_{i-1} and x_i.

Example

The upward velocity of a rocket is given as a function of time in table.

Find the velocity at t=16 seconds using linear splines.

Table : Velocity as a function of time

t (s)	$v(t)$ (m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67

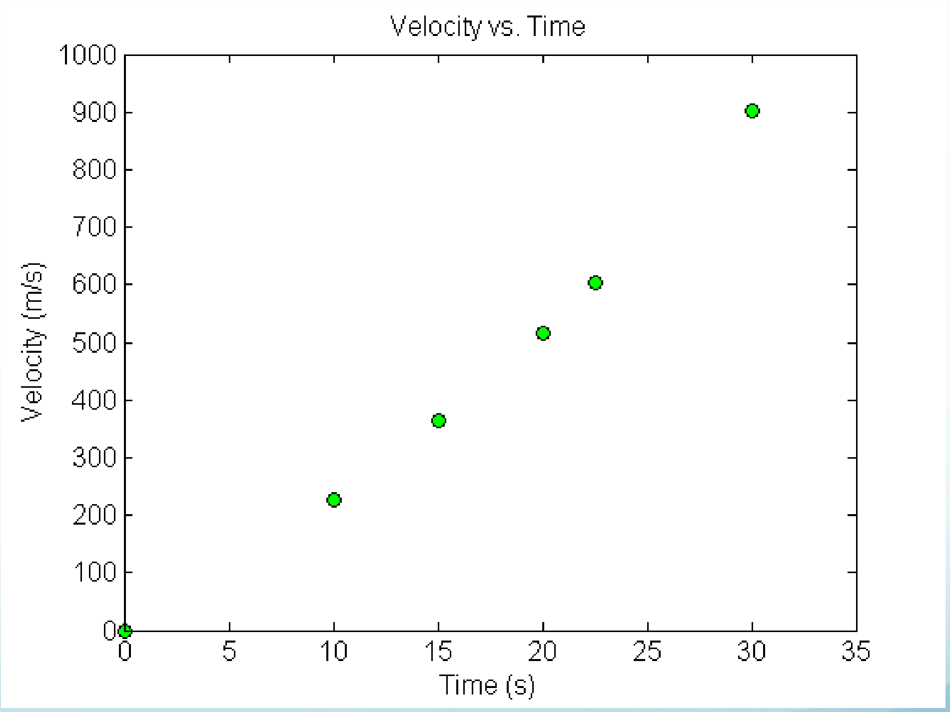


Figure : Velocity vs. time data for the rocket example

Linear Splines

$$t_0 = 15, \qquad v(t_0) = 362.78$$

$$t_1 = 20, \qquad v(t_1) = 517.35$$

$$v(t) = v(t_0) + \frac{v(t_1) - v(t_0)}{t_1 - t_0}(t - t_0)$$

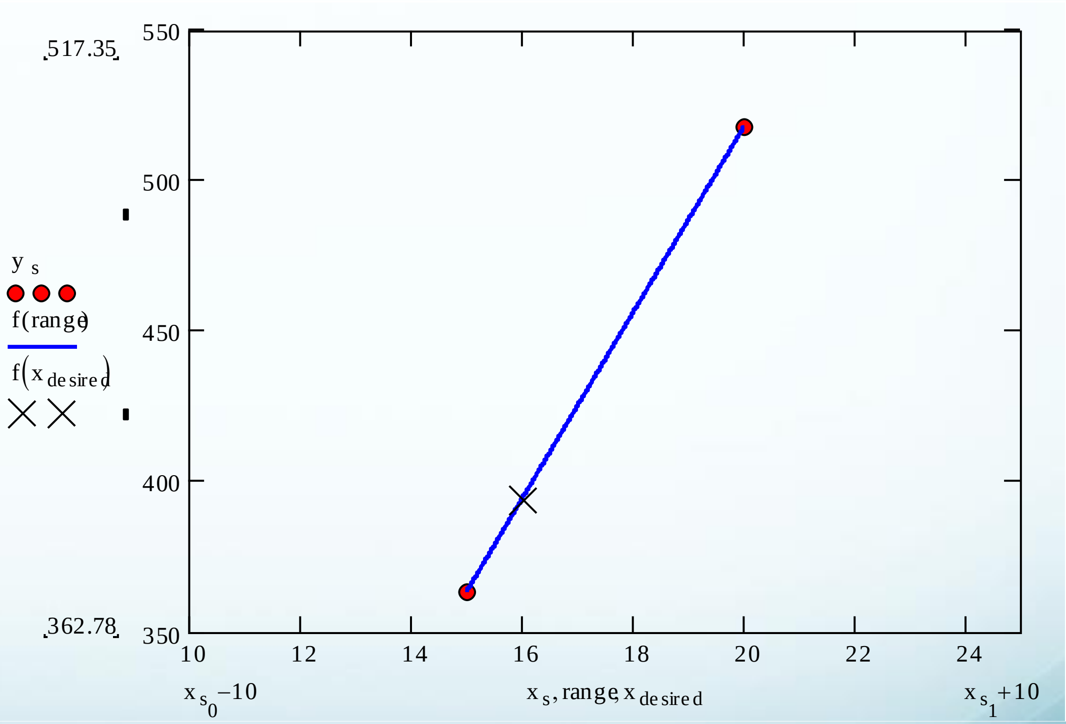
$$= 362.78 + \frac{517.35 - 362.78}{20 - 15}(t - 15)$$

$$v(t) = 362.78 + 30.913(t - 15)$$

At $t = 16$,

$$v(16) = 362.78 + 30.913(16 - 15)$$

$$= 393.7 \text{ m/s}$$



Linear Spline Algorithm

real function $\mathit{Spline1}(n, (t_i), (y_i), x)$
integer i, n; **real** x; **real array** $(t_i)_{0:n}, (y_i)_{0:n}$
for $i = n - 1$ **to** 0 **step** -1 **do**
  if $x - t_i \geqq 0$ **then** exit loop
end for
$\mathit{Spline1} \leftarrow y_i + (x - t_i)[(y_{i+1} - y_i)/(t_{i+1} - t_i)]$
end function $\mathit{Spline1}$

Quadratic Splines

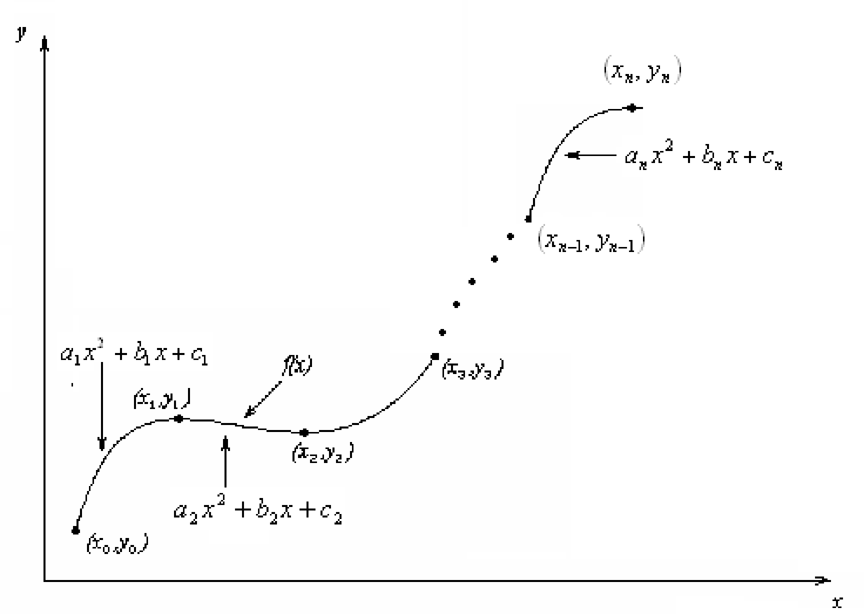
A function Q is a **second-degree spline** if it has the following properties

A function Q is called a **spline of degree 2** if:

1. The domain of Q is an interval **[*a, b*]**.
2. Q and Q' are continuous on **[*a, b*]**.
3. There are points t_i (called **knots**) such that $a = t_0 < t_1 < \cdot\cdot\cdot < t_n = b$ and Q is a polynomial of degree at most 2 on each subinterval $[t_i, t_{i+1}]$.

Quadratic Splines

Given $(x_0, y_0), (x_1, y_1), \ldots\ldots, (x_{n-1}, y_{n-1}), (x_n, y_n)$, fit quadratic splines through the data. The splines are given by



$$
\begin{aligned}
f(x) &= a_1 x^2 + b_1 x + c_1, & x_0 \le x \le x_1 \\
&= a_2 x^2 + b_2 x + c_2, & x_1 \le x \le x_2 \\
&. \\
&. \\
&. \\
&= a_n x^2 + b_n x + c_n, & x_{n-1} \le x \le x_n
\end{aligned}
$$

Find $a_i, b_i, c_i, i = 1, 2, \ldots, n$

Quadratic Splines

Each quadratic spline goes through two consecutive data points

$$a_1 x_0^{\;2} + b_1 x_0 + c_1 = f(x_0)$$

$$a_1 x_1^{\;2} + b_1 x_1 + c_1 = f(x_1) \qquad .$$

$$.$$

$$.$$

$$a_i x_{i-1}^{\;2} + b_i x_{i-1} + c_i = f(x_{i-1})$$

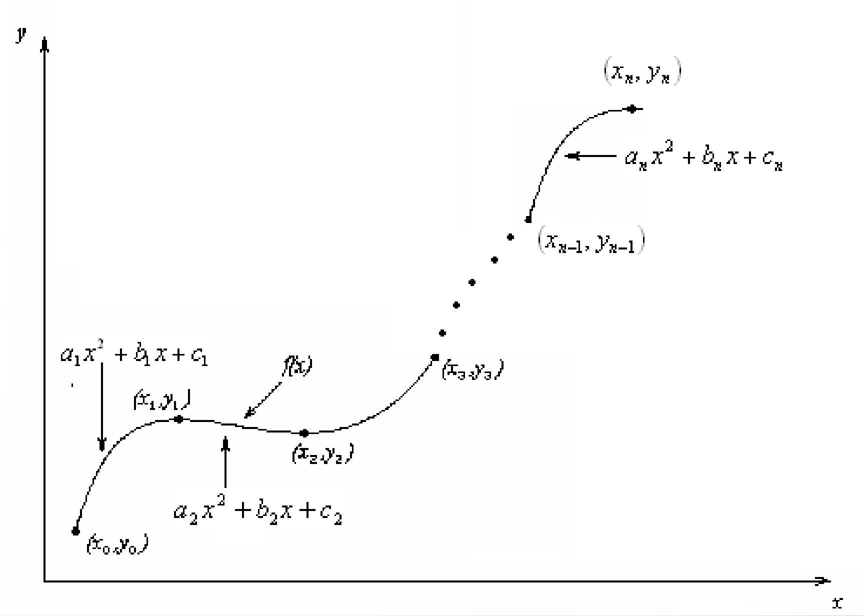
$$a_i x_i^{\;2} + b_i x_i + c_i = f(x_i) \qquad .$$

$$.$$

$$.$$

$$a_n x_{n-1}^{\;2} + b_n x_{n-1} + c_n = f(x_{n-1})$$

$$a_n x_n^{\;2} + b_n x_n + c_n = f(x_n)$$



This condition gives 2n equations

Quadratic Splines

The first derivatives of two quadratic splines are continuous at the interior points.

For example, the derivative of the first spline

$$a_1x^2 + b_1x + c_1 \quad \text{is} \quad 2a_1x + b_1$$

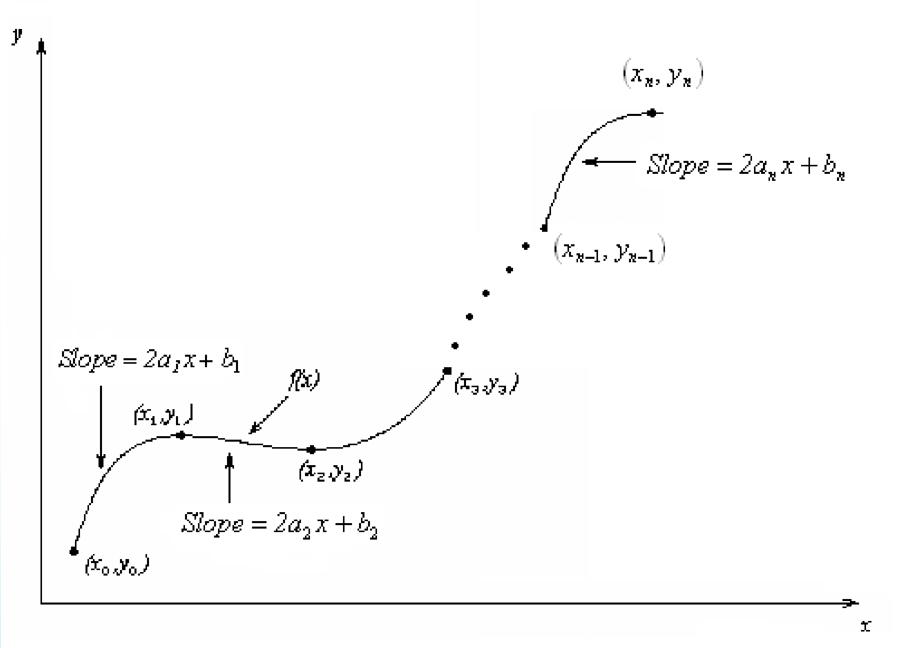
The derivative of the second spline

$$a_2x^2 + b_2x + c_2 \quad \text{is} \quad 2a_2x + b_2$$

and the two are equal at $x = x_1$ giving

$$2a_1x_1 + b_1 = 2a_2x_1 + b_2$$

$$2a_1x_1 + b_1 - 2a_2x_1 - b_2 = 0$$



Quadratic Splines

Similarly at the other interior points,

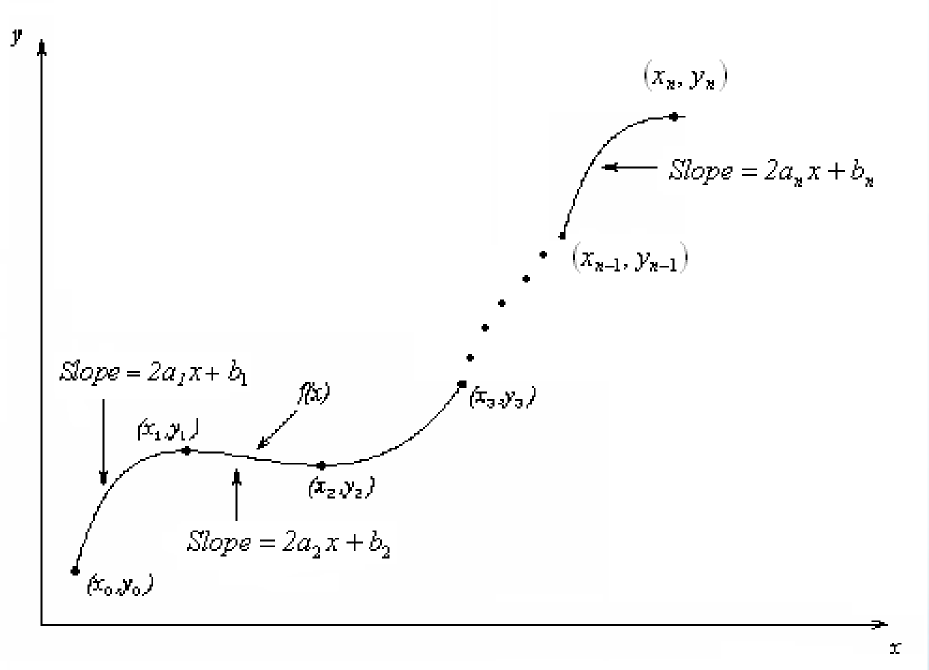
$$2a_2x_2 + b_2 - 2a_3x_2 - b_3 = 0$$

$$\vdots$$

$$2a_ix_i + b_i - 2a_{i+1}x_i - b_{i+1} = 0$$

$$\vdots$$

$$2a_{n-1}x_{n-1} + b_{n-1} - 2a_nx_{n-1} - b_n = 0$$



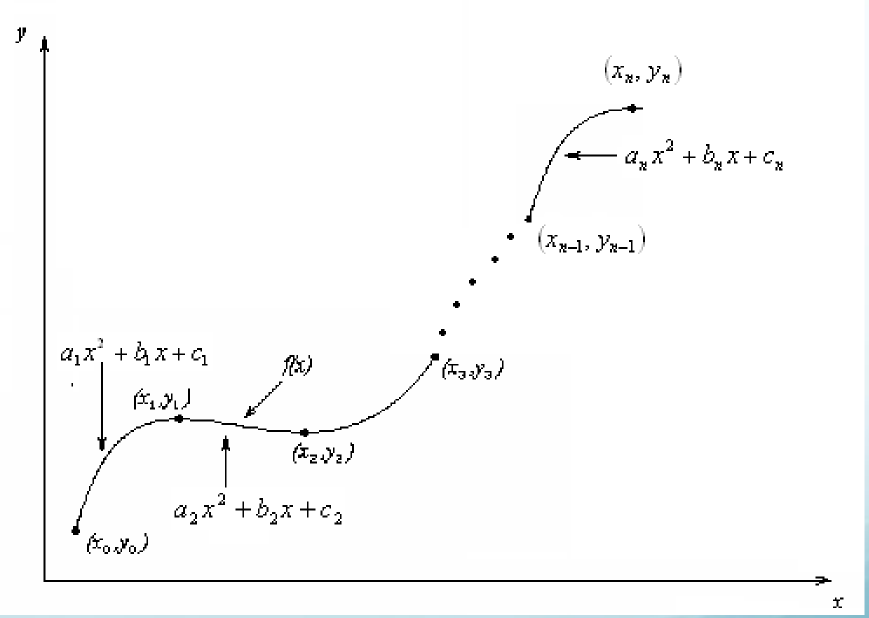
We have (n-1) such equations. The total number of equations is $(2n)+(n-1)=(3n-1)$.

We can assume that the first spline is linear, that is $a_1 = 0$

Quadratic Splines

This gives us '3n' equations and '3n' unknowns. Once we find the '3n' constants, we can find the function at any value of 'x' using the splines,

$$
\begin{aligned}
f(x) &= a_1x^2 + b_1x + c_1, && x_0 \le x \le x_1 \\
&= a_2x^2 + b_2x + c_2, && x_1 \le x \le x_2 \\
&\;\;. \\
&\;\;. \\
&\;\;. \\
&= a_nx^2 + b_nx + c_n, && x_{n-1} \le x \le x_n
\end{aligned}
$$



Quadratic Spline Example

The upward velocity of a rocket is given as a function of time. Using quadratic splines

a) Find the velocity at t=16 seconds

b) Find the acceleration at t=16 seconds

c) Find the distance covered between t=11 and t=16 seconds

Table : Velocity as a function of time

t (s)	$v(t)$ (m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67

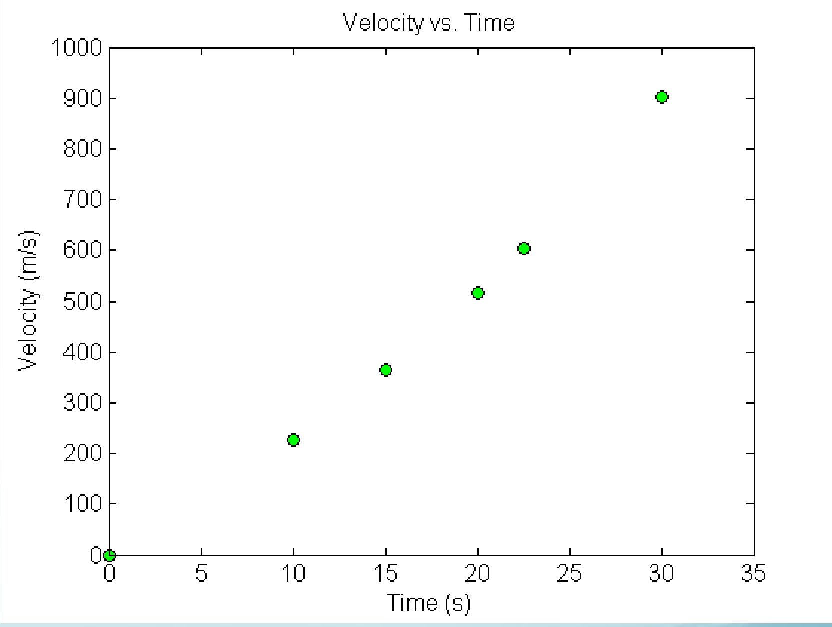


Figure : Velocity vs. time data for the rocket example

Solution

$$\begin{aligned} v(t) &= a_1 t^2 + b_1 t + c_1, & 0 \le t \le 10 \\ &= a_2 t^2 + b_2 t + c_2, & 10 \le t \le 15 \\ &= a_3 t^2 + b_3 t + c_3, & 15 \le t \le 20 \\ &= a_4 t^2 + b_4 t + c_4, & 20 \le t \le 22.5 \\ &= a_5 t^2 + b_5 t + c_5, & 22.5 \le t \le 30 \end{aligned}$$

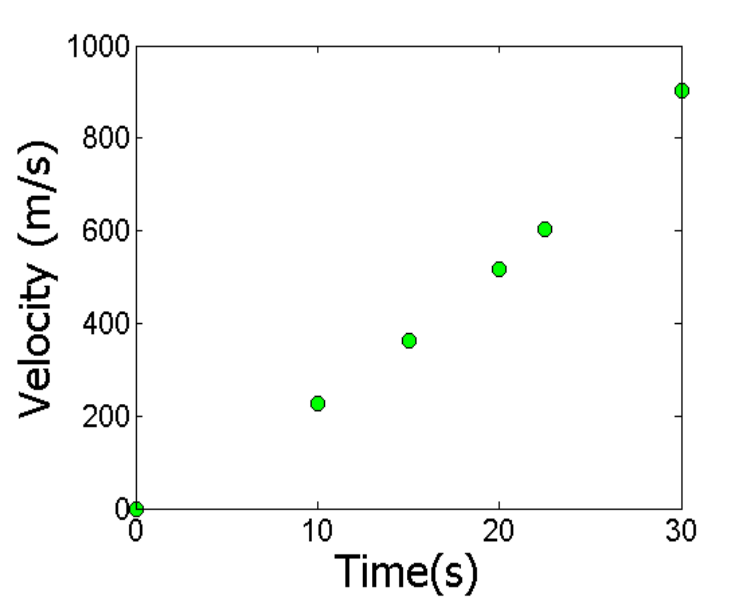
Let us set up the equations

Each Spline Goes Through Two Consecutive Data Points

$$v(t) = a_1 t^2 + b_1 t + c_1, \qquad 0 \le t \le 10$$

$$a_1(0)^2 + b_1(0) + c_1 = 0$$

$$a_1(10)^2 + b_1(10) + c_1 = 227.04$$



Each Spline Goes Through Two Consecutive Data Points

t	v(t)
s	m/s
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67

$$a_2(10)^2 + b_2(10) + c_2 = 227.04$$
$$a_2(15)^2 + b_2(15) + c_2 = 362.78$$
$$a_3(15)^2 + b_3(15) + c_3 = 362.78$$
$$a_3(20)^2 + b_3(20) + c_3 = 517.35$$
$$a_4(20)^2 + b_4(20) + c_4 = 517.35$$
$$a_4(22.5)^2 + b_4(22.5) + c_4 = 602.97$$
$$a_5(22.5)^2 + b_5(22.5) + c_5 = 602.97$$
$$a_5(30)^2 + b_5(30) + c_5 = 901.67$$

Derivatives are Continuous at Interior Data Points

$$v(t) = a_1 t^2 + b_1 t + c_1,\ 0 \le t \le 10$$

$$= a_2 t^2 + b_2 t + c_2, 10 \le t \le 15$$

$$\left.\frac{d}{dt}\left(a_1 t^2 + b_1 t + c_1\right)\right|_{t=10} = \left.\frac{d}{dt}\left(a_2 t^2 + b_2 t + c_2\right)\right|_{t=10}$$

$$\left.\left(2a_1 t + b_1\right)\right|_{t=10} = \left.\left(2a_2 t + b_2\right)\right|_{t=10}$$

$$2a_1(10) + b_1 = 2a_2(10) + b_2$$

$$20a_1 + b_1 - 20a_2 - b_2 = 0$$

Derivatives are Continuous at Interior Data Points

At t=10

$$2a_1(10)+b_1-2a_2(10)-b_2=0$$

At t=15

$$2a_2(15)+b_2-2a_3(15)-b_3=0$$

At t=20

$$2a_3(20)+b_3-2a_4(20)-b_4=0$$

At t=22.5

$$2a_4(22.5)+b_4-2a_5(22.5)-b_5=0$$

Last Equation $a_1=0$

Final Set of Equations

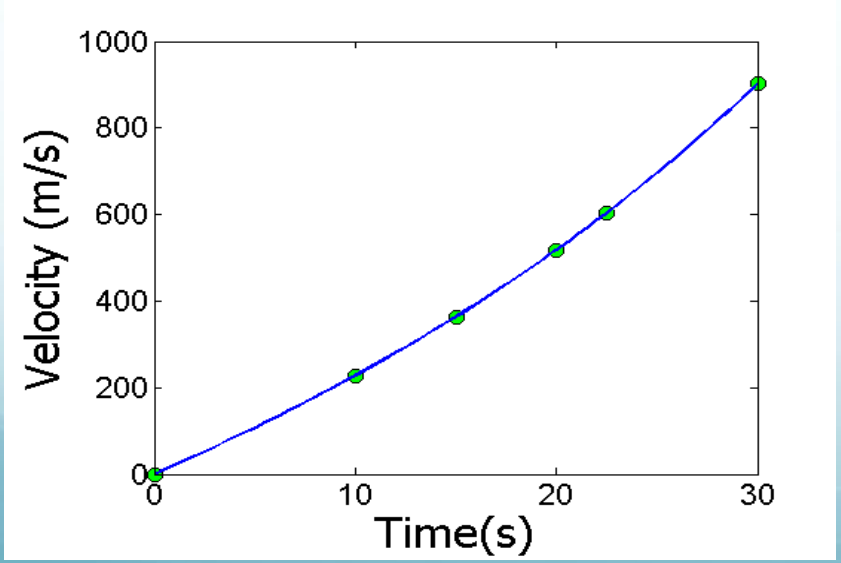
$$
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
100 & 10 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 100 & 10 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 225 & 15 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 225 & 15 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 400 & 20 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 400 & 20 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 506.25 & 22.5 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 506.25 & 22.5 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 900 & 30 & 1 \\
20 & 1 & 0 & -20 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 30 & 1 & 0 & -30 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 40 & 1 & 0 & -40 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 45 & 1 & 0 & -45 & -1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ a_3 \\ b_3 \\ c_3 \\ a_4 \\ b_4 \\ c_4 \\ a_5 \\ b_5 \\ c_5 \end{bmatrix}
=
\begin{bmatrix} 0 \\ 227.04 \\ 227.04 \\ 362.78 \\ 362.78 \\ 517.35 \\ 517.35 \\ 602.97 \\ 602.97 \\ 901.67 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

Coefficients of Spline

i	a_i	b_i	c_i
1	0	22.704	0
2	0.8888	4.928	88.88
3	−0.1356	35.66	−141.61
4	1.6048	−33.956	554.55
5	0.20889	28.86	−152.13

Final Solution

$$
\begin{aligned}
v(t) &= 22.704t, & 0 \le t \le 10 \\
&= 0.8888t^2 + 4.928t + 88.88, & 10 \le t \le 15 \\
&= -0.1356t^2 + 35.66t - 141.61, & 15 \le t \le 20 \\
&= 1.6048t^2 - 33.956t + 554.55, & 20 \le t \le 22.5 \\
&= 0.20889t^2 + 28.86t - 152.13, & 22.5 \le t \le 30
\end{aligned}
$$



Velocity at a Particular Point

a) Velocity at t=16

$$
\begin{aligned}
v(t) &= 22.704t, & 0 \le t \le 10 \\
&= 0.8888t^2 + 4.928t + 88.88, & 10 \le t \le 15 \\
&= -0.1356t^2 + 35.66t - 141.61, & 15 \le t \le 20 \\
&= 1.6048t^2 - 33.956t + 554.55, & 20 \le t \le 22.5 \\
&= 0.20889t^2 + 28.86t - 152.13, & 22.5 \le t \le 30
\end{aligned}
$$

$$
\begin{aligned}
v(16) &= -0.1356(16)^2 + 35.66(16) - 141.61 \\
&= 394.24 \text{ m/s}
\end{aligned}
$$

Smoothing of Data & Least Squares Method

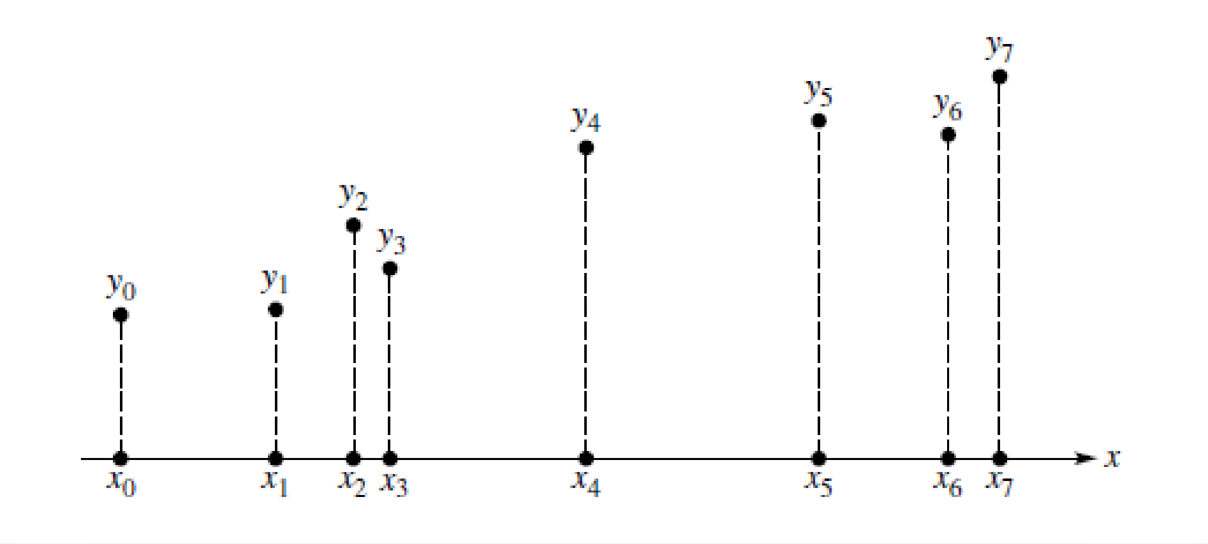
- **What** is data smoothing?

- **Why** do we use data smoothing & least squares?

- **How** do we solve method of least squares?

data smoothing

- Fitting a smooth curve to tabulated data by experiments.
- Draw a curve that defines the characteristic of data best.
- Obtaining a function that express data.



- Dispersed data points

Least Squares Method

- Using error function to fit into a curve. This is a method of **minimizing the error of function.**

Linear Form

f (*x*) = *a* +*bx*

In this case, the function to be minimized is

$$S(a,b) = \sum_{i=0}^{n} [y_i - f(x_i)]^2 = \sum_{i=0}^{n} (y_i - a - bx_i)^2$$

Linear Least Squares

$$S(a,b)=\sum_{i=0}^{n}\left[y_i-f(x_i)\right]^2=\sum_{i=0}^{n}\left(y_i-a-bx_i\right)^2$$

$$\frac{\partial S}{\partial a}=\sum_{i=0}^{n}-2(y_i-a-bx_i)=2\left[a\,(n+1)+b\sum_{i=0}^{n}x_i-\sum_{i=0}^{n}y_i\right]=0$$

$$\frac{\partial S}{\partial b}=\sum_{i=0}^{n}-2(y_i-a-bx_i)x_i=2\left(a\sum_{i=0}^{n}x_i+b\sum_{i=0}^{n}x_i^2-\sum_{i=0}^{n}x_iy_i\right)=0$$

Dividing both equations by 2(n+1), rearrange terms;

$$a+\bar{x}b=\bar{y}\qquad \bar{x}a+\left(\frac{1}{n+1}\sum_{i=0}^{n}x_i^2\right)b=\frac{1}{n+1}\sum_{i=0}^{n}x_iy_i$$

Linear Least Squares

$$a + \bar{x}b = \bar{y} \qquad \bar{x}a + \left(\frac{1}{n+1}\sum_{i=0}^{n} x_i^2\right) b = \frac{1}{n+1}\sum_{i=0}^{n} x_i y_i$$

Where;

$$\bar{x} = \frac{1}{n+1}\sum_{i=0}^{n} x_i \qquad \bar{y} = \frac{1}{n+1}\sum_{i=0}^{n} y_i$$

are the mean values of x and y data. The solution for parameters is;

$$a = \frac{\bar{y}\sum x_i^2 - \bar{x}\sum x_i y_i}{\sum x_i^2 - n\bar{x}^2} \qquad b = \frac{\sum x_i y_i - \bar{x}\sum y_i}{\sum x_i^2 - n\bar{x}^2}$$

Linear Least Squares

Data obtained by an experiment;

x	4	7	11	13	17
y	2	0	2	6	7

System of two equations;

$$\begin{cases} 644a + 52b = 227 \\ 52a + 5b = 17 \end{cases}$$

Corresponding values;

$$a = 0.4864$$

$$b = -1.6589$$