Numerical Method(Math-2073/53)
Lecture note: Chapter-IV

"Mathematics is the most beautiful and most powerful creation of the human spirit."

STEFAN BANAC

Dejen Ketema
Department of Mathematics
dejen.ketema@amu.edu.et

April, 2019

Contents

Chapter 1

Finite Differences

Lot of operators are used in numerical analysis/computation. Some of the frequently used operators, viz. forward difference (Δ), backward difference (∇), central difference (δ), shift (E) and mean (μ) are discussed in this chapter.

Assume that we have a table of values $f(x_i), i = 1, 2, 3, \cdots, n$ of any function $y = f(x)$ where the value of x being equally spaced, that is, $x_i = x_0 + ih$. When $x = x_i$, the value of y is denoted by y_i and is defined by $y_i = f(x_i)$. The values of x and y are called **arguments** and **entries** respectively. To determine the value of $f(x)$ and $f'(x)$ for some intermediate values of x, is based on the principle of finite difference. Which requires three types of differences.

1.1 Forward difference operators

Suppose that a function $y = f(x)$ is tabulated for the equally spaced arguments $x_0, x_0 + h, x_0 + 2h \cdots, x_0 + nh$ giving the functional values $y_0, y_1, y_2, \cdots, y_n$. The constant difference between two consecutive values of x is called the **interval of differences** and is denoted by **h**.

The operator Δ defined by

$$\begin{aligned}
\Delta y_0 &= y_1 - y_0 \\
\Delta y_1 &= y_2 - y_1 \\
\Delta y_2 &= y_3 - y_2 \\
&\vdots \\
\Delta y_{n-1} &= y_n - y_{n-1}
\end{aligned}$$

is called **Forward difference** operator.

The first forward difference is $\Delta y_n = y_{n+1} - y_n$. The second forward difference are defined as the difference of the first differences.

$$\begin{aligned}
&\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = (y_2 - y_1) - (y_1 - y_0) = y_2 - 2y_1 + y_0 \\
&\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = (y_3 - y_2) - (y_2 - y_1) = y_3 - 2y_2 + y_1 \\
&\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = (y_3 - 2y_2 + y_1) - (y_2 - 2y_1 + y_0) = y_3 - 3y_2 + 3y_1 - y_0 \\
&\Delta^3 y_1 = y_4 - 3y_3 + 3y_2 - y_1 \\
&\qquad \vdots
\end{aligned}$$

Similarly, higher order differences can be defined. In general,general,

$$\Delta^{n+1}f(x) = \Delta[\Delta^n f(x)],\ i.e.\Delta^{n+1}y_i = \Delta[\Delta^n y_i], n = 0, , 2, \cdots$$

Again, $\Delta^{n+1}f(x) = \Delta^n[f(x+h) - f(x)] = \delta^n f(x+h) - \Delta^n f(x)$ and

$$\Delta^{n+1}y_i = \Delta^n y_{i+1} - \Delta^n y_i, \quad n = 0, 1, 2, \cdots$$

It must be remembered that $\Delta^0 \equiv$ identity operator, i.e. $\Delta^0 f(x) = f(x)$ and $\Delta^1 \equiv \Delta$. All the forward differences can be represented in a tabular form, called the forward difference or diagonal difference table.

Let $x_0, x_1, \cdots, x_4$ be four arguments. All the forwarded differences of these arguments are shown in the table below.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
x_0	y_0					
		Δy_0				
x_1	y_1		$\Delta^2 y_0$			
		Δy_1		$\Delta^3 y_0$		
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$	
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$	
		Δy_3		$\Delta^3 y_2$		
x_4	y_4		$\Delta^2 y_3$			
		Δy_4				
x_5	y_5					

The operator Δ satisfies the following properties:

1. $\Delta[f(x) - g(x)] = \Delta f(x) - \Delta g(x)$
2. $\Delta[cf(x)] = c\Delta f(x)$
3. $\Delta^m \Delta^n f(x) = \Delta^{m+n} f(x)$, m, n are positive integers
4. Since $\Delta^n y_n$ is a constant, $\Delta^{n+1}y_n = 0$, $\Delta^{n+2}y_n = 0, \cdots$ i.e. $(n+1)^{th}$ and higher differences are zero.

1.2 Backward difference operators

Suppose that a function $y = f(x)$ is tabulated for the equally spaced arguments $x_0, x_0 + h, x_0 + 2h, \cdots, x_0 + nh$ giving the functional values $y_0, y_1, y_2, \cdots, y_n$. The operator ∇ defined by

$$\begin{aligned}\nabla y_1 &= y_1 - y_0\\ \nabla_y 2 &= y_2 - y_1\\ &\vdots\\ \nabla y_n &= y_n - y_{n-1}\end{aligned}$$

is called **Backward difference** operator.

The first backward difference is $\nabla y_n = y_n - y_{n-1}$ The second backward difference are obtain by the difference of the first differences.

$$\begin{aligned}
\nabla^2 y_2 &= \nabla(\nabla y_2) = \nabla(y_2 - y_1) = \nabla y_2 - \nabla y_1 \\
&= y_2 - y_1 - (y_1 - y_0) = y_2 - 2y_1 + y_0 \\
\nabla^2 y_3 &= \nabla y_3 - \nabla y_2 \\
\nabla^2 y_n &= \nabla y_n - \nabla y_{n-1}
\end{aligned}$$

In general nth backward difference of f is defined by

$$\nabla^n y_i = \nabla^{n-1} y_i - \nabla^{n-1} y_{i-1}$$

Backward Difference Table:

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$	$\nabla^5 y$
x_0	y_0					
		∇y_1				
x_1	y_1		$\nabla^2 y_2$			
		∇y_2		$\nabla^3 y_3$		
x_2	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$	
		∇y_3		$\nabla^3 y_4$		$\nabla^5 y_5$
x_3	y_3		$\nabla^2 y_4$		$\nabla^4 y_5$	
		∇y_4		$\nabla^3 y_5$		
x_4	y_4		$\nabla^2 y_5$			
		∇y_5				
x_5	y_5					

1.3 Central difference operators

The operator δ defined by

$$\begin{aligned}
\delta y_{1\frac{1}{2}} &= y_1 - y_0 \\
\delta y_{1\frac{3}{2}} &= y_2 - y_1 \\
&\vdots \\
\delta y_{n-1\frac{1}{2}} &= y_n - y_{n-1}
\end{aligned}$$

is called **Central difference** operator.

Similarly, higher order central differences are defined as

$$\begin{aligned}
\delta^2 y_1 &= y_{1\frac{3}{2}} - y_{1\frac{1}{2}} \\
\delta^2 y_2 &= y_{1\frac{5}{2}} - y_{1\frac{3}{2}} \\
&\vdots \\
\delta^n y_i &= \delta^{n-1} y_{i+1/2} - \delta^{n-1} y_{i-1/2}
\end{aligned}$$

Central Difference Table:

x	y	δy	$\delta^2 y$	$\delta^3 y$	$\delta^4 y$	$\delta^5 y$
x_0	y_0					
		δy_0				
x_1	y_1		$\delta^2 y_0$			
		δy_1		$\delta^3 y_0$		
x_2	y_2		$\delta^2 y_1$		$\delta^4 y_0$	
		δy_2		$\delta^3 y_1$		$\delta^5 y_0$
x_3	y_3		$\delta^2 y_2$		$\delta^4 y_1$	
		δy_3		$\delta^3 y_2$		
x_4	y_4		$\delta^2 y_3$			
		δy_4				
x_5	y_5					

NOTE:

- From all three difference tables, we can see that only the notations changes not the differences.
$$y_1 - y_0 = \Delta y_0 = \nabla y_1 = \delta y_{\frac{1}{2}}$$
- Alternative notations for the function $y = f(x)$. For two consecutive values of x differing by h.
$$\begin{aligned}\Delta y_x &= y_{x+h} - y_x = f(x+h) - f(x)\\ \nabla y_x &= y_x - y_{x-h} = f(x) - f(x-h)\\ \delta y_x &= y_{x+h/2} - y_{x-h/2} = f(x+h/2) - f(x-h/2)\end{aligned}$$

1.4 Other operators

1. **Shift Operator** E:
   E does the operation of increasing the argument x by h so that
$$\begin{aligned}Ef(x) &= f(x+h);\\ E^2 f(x) &= E(Ef(x)) = Ef(x+h) = f(x+2h);\\ E^3 f(x) &= f(x+3h);\\ E^n f(x) &= f(x+nh);\end{aligned}$$

   The inverse operator E^{-1} is defined as
$$\begin{aligned}E^{-1} f(x) &= f(x-h);\\ E^{-n} f(x) &= f(x-nh);\end{aligned}$$

   If y_x is the function $f(x)$, then
$$\begin{aligned}Ey_x &= y_x + h;\\ E^{-1} y_x &= y_{x-h};\\ E^n y_x &= y_x + nh;\end{aligned}$$

2. **Averaging Operator μ:**
   It is defined as

$$\mu f(x) = \frac{1}{2}\left[f(x+\frac{h}{2}) + f(x-\frac{h}{2})\right]$$
$$i.e\ \mu f(x) = \frac{1}{2}\left[y_{x+\frac{h}{2}} + y_{x-\frac{h}{2}}\right]$$

3. **Differential Operator D:**
   It is defined as

$$Df(x) = \frac{d}{dx}f(x) = f'(x);$$

1.5 Relation between the operators

1. $\Delta = E - 1$
2. $\nabla = 1 - E^{-1}$
3. $\delta = E^{1/2} - E^{-1/2}$
4. $\mu = \frac{1}{2}\{E^{1/2} - E^{-1/2}\}$
5. $\delta = E\nabla = \nabla E = \delta E^{\frac{1}{2}}$
6. $E = e^{hD}$
7. $(1+\Delta)(1-\nabla) = 1$
8. $\Delta - \nabla = \Delta\nabla = \delta^2$
9. $1 + \mu^2\delta^2 = \left(1 + \frac{1}{2}\delta^2\right)^2$
10. $\mu^2 = 1 + \frac{1}{4}\delta^2$
11. $E^{1/2} = \mu + \frac{1}{2}\delta$
12. $E^{-1/2} = \mu - \frac{1}{2}\delta$
13. $\Delta = \frac{1}{2}\delta^2 + \delta\sqrt{1 + \frac{\delta^2}{4}}$
14. $\mu\delta = \frac{1}{2}\Delta E^{-1} + \frac{1}{2}\Delta$

Example 1.1

Construct a forward difference table for the following values:

x	0	5	10	15	20	25
f(x)	7	11	14	18	24	32

Solution: *Forward difference table for given data is:*

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
0	7					
		4				
5	11		-1			
		3		2		
10	14		1		-1	
		4		1		0
15	18		2		-1	
		6		0		
20	24		2			
		8				
25	32					

Example 1.2

Given $f(0) = 3, f(1) = 12, f(2) = 81, f(3) = 200, f(4) = 100$ and $f(5) = 8$. From the difference table and find $\Delta^5 f(0)$.

Solution:
The difference table for given data is as follows:

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$	$\Delta^5 f(x)$
0	3					
		9				
5	12		60			
		69		-10		
10	81		50		-259	
		119		-269		755
15	200		-219		496	
		-100		227		
20	100		8			
		-92				
25	8					

Hence, $\Delta^5 f(0) = 755$.

Exercise 1.1

a) *Obtain the missing term in the following table:*

x	2.0	2.1	2.2	2.3	2.4	2.5	2.6
f(x)	0.135	?	0.111	0.100	?	0.082	0.074

*[**Ans:**f(2.1) = 0.123, f(2.4) = 0.0900]*

b) *Estimate the production for the year 1964 and 1966 from the following data:*

Year	1961	1962	1963	1964	1965	1966	1967
product	200	220	260	?	350	?	430

*[**Ans:** f(1964) = 306, f(1966) = 390]*

Exercise 1.2

Evaluate the following. The interval of difference being h.

a) $\Delta^n e^x$

b) $\Delta \log f(x)$

c) $\Delta(\tan^{-1} x)$

d) $\Delta^2 \cos 2x$

Chapter 2

Interpolation

It is often needed to estimate the value of a function $y = f(x)$ at certain point x based on the known values of the function $f(x_0), \cdots, f(x_n)$ at a set of $n + 1$ node points $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$ in the interval $[a, \ b]$. This process is called **interpolation** if $a < x < b$ or **extrapolation** if either $x < a$ or $b < x$. Theorem by Weierstrass in 1885, "Every continuous function in an interval (a, b) can be represented in that interval to any desired accuracy by a polynomial. "

Definition 2.1

Interpolation is the process of deriving a simple function from a set of discrete data points so that the function passes through all the given data points (i.e. reproduces the data points exactly) and can be used to estimate data points in-between the given ones.

It is necessary because in science and engineering we often need to deal with discrete experimental data. Interpolation is also used to simply complicated functions by sampling data points can interpolating them using a simpler function. Polynomials are commonly used for interpolation because they are easier to evaluate, differentiate, and integrate - known as polynomial interpolation.

It can be proven that given $n + 1$ data points it is always possible to find a polynomial of order/degree n to pass through/reproduce the $n+1$ points. Purpose of numerical Interpolation

1. Compute intermediate values of a sampled function
2. Numerical differentiation - foundation for Finite Difference and Finite Element methods
3. Numerical Integration

2.1 Interpolations with Equal Interval

2.1.1 Gregory - Newton Forward Interpolation Formula

To estimate the value of a function near the **beginning** a table, the forward difference interpolation formula in used.

Let $y_x = f(x)$ be a function which takes the values $y_{x_0}, y_{x_0+h}, y_{x_0+2h}, \cdots$ corresponding to the values $x_0, x_0 + h, x_0 + 2h, \cdots$of x. Suppose we want to evaluate y_x when $x = x_0 + ph$, where p is any real number.

Let it be y_p. For any real number n, we have defined operator E such that $E^n f(x) = f(x+nh)$.

$$\begin{aligned}\therefore y_x = y_{x_0} + ph = f(x_0+ph) = E^p y_{x_0} = (1+\Delta)^p y_0 \\ = \left[1 + p\Delta + \frac{p(p-1)}{2!}\Delta^2 + \frac{p(p-1)(p-2)}{3!}\Delta^3 + \cdots\right] y_0\end{aligned}$$

$$y_x = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots$$

is called **Newton's forward interpolation formula**.

Example 2.1

Ordinates $f(x)$ of a normal curve in terms of standard deviation x are given as:

x	1.00	1.02	1.04	1.06	1.08
f(x)	0.2420	0.2371	0.2323	0.2275	0.2227

.

Find the ordinate for standard deviation $x = 1.025$.
Solution:*Let us first form the difference table:*

x	y	Δ	Δ^2	Δ^3	Δ^4
1.00	0.2420				
		-0.0049			
1.02	0.2371		0.0001		
		-0.0048		-0.0001	
1.04	0.2323		0		0.001
		-0.0048		0	
1.06	0.2275		0		
		-0.0048			
1.08	0.2227				

Here,$h = 0.02$, $x_0 = 1.00$, $x = 1.025$

$$\therefore p = \frac{x - x_0}{h} = \frac{1.025 - 1.00}{0.02} = 1.25$$

$$y_x = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \frac{p(p-1)(p-2)p(p-3)}{4!}\Delta^4 y_0$$

Now on putting values of various functions, we get

$$\begin{aligned}y_{1.025} &= 0.2420 + 1.25 \times (-0.0049) + \frac{1.25(1.25-1)}{2!} \times (0.0001) \\ &+ \frac{1.25(1.25-1)(1.25-2)}{3!} \times (-0.0001) + \frac{1.25(1.25-1)(1.25-2)p(1.25-3)}{4!} \times (0.01) \\ &= 0.2420 - 0.006125 + 0.000015625 + 0.000003906 + 0.000001708 \\ &= 0.242021239 - 0.006125 = 0.235896239(Approx.)\end{aligned}$$

Example 2.2

Find the cubic polynomial which takes the following data:

x	0	1	2	3
f(x)	1	0	1	10

Solution: *Let us first form the difference table:*

x	y	Δ	Δ^2	Δ^3
0	1			
		-1		
1	0		2	
		1		6
2	1		8	
		9		
3	10			

$$y_x = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0$$

Now. $x = x_0 + ph$

$$p = \frac{x - x_0}{h} = \frac{x-0}{1} = x$$

$$\begin{aligned} y &= y_0 + x\Delta y_0 + \frac{x(x-1)}{2!}\Delta^2 y_0 + \frac{x(x-1)(x-2)}{3!}\Delta^3 y_0 \\ &= 1 + x(-1) + \frac{x(x-1)}{2} \times (2) + \frac{x(x-1)(x-2)}{6} \times (6) \\ &= 1 - x + x^2 - x + x^3 - 2x2 - x^2 + 2x \\ &= x^3 - 2x^2 + 1 \end{aligned}$$

Example 2.3

Find the number of students from the following data who secured marks not more than 45.

Marks range	30-40	40-50	50-60	60-70	70-80
No. of students	35	48	70	40	22

Construct a table Here, $h = 10, x_0 = 40, x = 45$

$$p = \frac{45 - 40}{10} = 0.5$$

No. of students who secured not more than 45 *marks are* 51.

Exercise 2.1

The population of the town in decennial census was as given below estimate the population for the year 1895.

Year:	1891	1901	1911	1921	1931
Population(in million):	46	66	81	93	101

Exercise 2.2

Following are the marks obtained by 492 candidates in a certain examination:

Marks	0-40	40-45	45-50	50-55	55-60	60-65
No.of Candidates	210	43	54	74	32	79

Find out

(a) *No. of candidates, if they secure more than 48 but less than 50 marks.*

(b) *Less than 48 but not less than 45 marks*

2.1.2 Gregory -Newton Backward Interpolation Formula

To estimate the value of a function near the **end** a table, the backward difference interpolation formula in used.

Let $y_x = f(x)$ be a function which takes the values $y_{x_0}, y_{x_0+h}, y_{x_0+2h}, \cdots$ corresponding to the values $x_0, x_0+h, x_0+2h, \cdots$of x. Suppose we want to evaluate y_x when $x = x_n + ph$, where p is any real number.
Let it be y_p. For any real number n, we have defined operator E such that $E^n f(x) = f(x+nh)$.

$$\therefore y_x = y_{x_n} + ph = f(x_n+ph) = E^p y_{x_n} = (1-\nabla)^{-p} y_n$$
$$= \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \cdots\right] y_n$$
$$y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \cdots$$

is called **Newton's Backward interpolation formula**.

Example 2.4

Using Newton's backward difference formula find the value of $e^{-1.9}$ from the following table of value of e^{-x}.

x	1	1.25	1.50	1.75	2.00
e^{-x}	.3679	0.2865	0.2231	0.1738	0.1353

Difference table for the given data as follows:

x	e^{-x}	∇	∇^2	∇^3	∇^4
1	0.3679				
		-0.0814			
1.25	0.2865		0.0180		
		-0.0634		-0.0039	
1.50	0.2231		0.0141		0.0006
		-0.0493		- 0.0033	
1.75	0.1738		0.0108		
		-0.0385			
2.00	0.1353				

$$p = \frac{x - x_n}{h} = \frac{1.9-2}{0.25} = -0.4$$

Example

using Newton's backward difference formula

$$y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_n$$

On putting the subsequent values, we get

$$\begin{aligned} y_{1.9} &= 0.1353 + (-0.4) \times (-0.0385) + \frac{-0.4(-0.4+1)}{2!}(0.0108) \\ &+ \frac{-0.4(-0.4+1)(-0.4+2)}{3!} \times (-0.0033) + \frac{-0.4(-0.4+1)(-0.4+2)(-0.4+3)}{4!} \times (0.0006) \\ &= 0.1353 + 0.0154 - 0.001296 + 0.0002112 + 0.000024 \\ &= 0.14959 \end{aligned}$$

Example 2.5

The area A of a circle of diameter d is given for the following values:

d	80	85	90	95	100
A	5026	5674	6362	7088	7854

Find A for 105.
First of all we form the difference table as follow:

d	A	∇	∇^2	∇^3	∇^4
80	5026				
		648			
85	5674		40		
		688		-2	
90	6362		38		4
		726		2	
95	7088		40		
		766			
100	7854				

Here, $h = 5, x_n = 100, x = 105$

$$p = \frac{x - x_n}{h} = \frac{105 - 100}{5} = 1$$

Now on applying Newton's backward difference formula, we have

$$\begin{aligned} y_x &= y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_n \\ y_{105} &= 785 + 1(766) + \frac{1(1+1)}{2!}(40) + \frac{1(1+1)(1+2)}{3!}(2) + \frac{1(1+1)(1+2)(1+3)}{4!}(4) \\ &= 7854 + 766 + 46 \\ &= 8666 \end{aligned}$$

Which is the required area for the given diameter of circle.

Exercise 2.3

Using Newtons backward difference interpolation, interpolate at $x = 1$ from the following data.

x	0.1	0.2	0.3	0.4	0.5	0.6
f(x)	1.699	1.073	0.375	0.443	1.429	2.631

Exercise 2.4

The table gives the distance in nautical miles of the visible horizon for the given heights in feet above the earth's surface. Find the value of y when $x = 390ft$.

Height(x):	100	150	200	250	300	350	400
Distance(y):	10.63	13.03	15.04	16.81	18.42	19.90	21.47

2.1.3 Stirling's Interpolation Formula

To estimate the value of a function near the middle a table, the central difference interpolation formula in used.
Let $y_x = f(x)$ be a functional relation between x and y. If x takes the values $x_0 - 2h, x_0 - h, x_0, x_0 + h, x_0 + 2h, \cdots$ and the corresponding values of y are $y_{-2}, y_{-1}, y_0, y_1, y_2 \cdots$ then we can form a central difference table as follows:

x	y	1^{st} $difference$	2^{nd} $difference$	3^{rd} $difference$	4^{th} $difference$
$x_0 - 2h$	y_{-2}				
		$\Delta y_{-2}(=\delta y_{-3/2})$			
$x_0 - h$	y_{-1}		$\Delta^2 y_{-2}(=\delta^2 y_{-1})$		
		$\Delta y_{-1}(=\delta y_{-1/2})$		$\Delta^3 y_{-2}(=\delta^3 y_{-1/2})$	
x_0	y_0		$\Delta^2 y_{-1}(=\delta^2 y_0)$		$\Delta^4 y_{-2}(=\delta^4 y_0)$
		$\Delta y_0(=\delta y_{1/2})$		$\Delta^3 y_{-1}(=\delta^3 y_{1/2})$	
$x + h$	y_1		$\Delta^2 y_0(=\delta^2 y_1)$		
		$\Delta y_1(=\delta y_{3/2})$			
$x + 2h$	y_2				

The Stirling's formula in forward difference notation is

$$
\begin{aligned}
y_p = y_0 + p\left[\frac{\Delta y_0 - \Delta y_{-1}}{2}\right] + \frac{p^2}{2!}\Delta^2 y_{-1} \\
+ \frac{p(p^2 - 1^2)}{3!}\left[\frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2}\right] + \frac{p^2(p^2 - 1^2)}{4!}\Delta^4 y_{-2} \cdots
\end{aligned}
$$

Example 2.6

Apply Stirling's formula to find a polynomial of degree three which takes the following values of x and y:

x	2	4	6	8	10
y	-2	1	3	8	20

Solution: *Let* $p = \dfrac{x-6}{2}$. *Now, we construct the following difference table:*

x	p	y_p	Δy_p	$\Delta^2 y_p$	$\Delta^3 y_p$	$\Delta^4 y_p$
2	-2	-2				
			3			
4	1	1		-1		
			2		4	
6	0	3		3		0
			5		4	
8	1	8		7		
			12			
10	2	20				

Stirling's formula is

$$\begin{aligned}
y_p &= y_0 + p\left[\frac{\Delta y_0 - \Delta y_{-1}}{2}\right] + \frac{p^2}{2!}\Delta^2 y_{-1} \\
&\quad + \frac{p(p^2-1^2)}{3!}\left[\frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2}\right] + \frac{p^2(p^2-1^2)}{4!}\Delta^4 y_{-2} \\
y_p &= 3 + p\left[\frac{2+5}{2}\right] + \frac{p^2}{2!}(3) + \frac{p(p^2-1^2)}{6}\left[\frac{4+4}{2}\right] + 0 \\
&= 3 + 2/3p^3 + 3/2u^2 + 17/6u \\
&= 3 + 2/3(\frac{x-6}{2})^3 + 3/2(\frac{x-6}{2})^2 + 17/6(\frac{x-6}{2}) \\
&= 0.0833x^3 - 1.125x^2 + 8.9166x - 19.
\end{aligned}$$

Exercise 2.5

Using Stirling's formula find y_{35}

x:	10	20	30	40	50
y:	600	512	439	346	243

Exercise 2.6

The function y is given in the table below: Find y for $x = 0.0341$

x:	0.01	0.02	0.03	0.04	0.05
y:	98.4342	48.4392	31.7775	23.4492	18.4542

2.2 Interpolations with Unequal Interval

The interpolation formula derived before for forward interpolation, Backward interpolation and central interpolation have the disadvantages of being applicable only to equally spaced argument values. So it is required to develop interpolation formulae for unequally spaced argument values of x. Therefore, when the values of the argument are not at equally spaced then we use two such formulae for interpolation.

1. Lagrange's Interpolation formula
2. Newton's Divided difference formula.
3. Spline Interpolation

2.2.1 Polynomial Interpolation

Let us assign polynomial P_n of degree n (or less) that assumes the given data values

$$P_n(x_0) = y_0, P_n(x_1) = y_1, \cdots, P_n(x_n) = y_n$$

This polynomial P_n is called **interpolation polynomial.** $x_0, x_1, \cdots, x_n$ is called the nodes (tabular points, pivotal points or arguments).

One way to carry out these operations is to approximate the function $f(x)$ by an n^{th} degree polynomial:

$$f(x) \approx P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0 = \sum_{j=0}^{n} a_j x^j$$

where the $n+1$ coefficients $a_0, \cdots, a_n$ can be obtained based on the $n+1$ given points. Once $P_n(x)$ becomes available, any operation applied to the function $f(x)$, such as differentiation, intergration, and root finding, can be carried out approximately based on $P_n(x) \approx f(x)$. This is particulary useful if the function $f(x)$ is non-elementary and therefore difficult to manipulate, or it is only available as a set of discrete samples without a closed-form expression.

2.2.2 Lagrange polynomial

In numerical analysis, Lagrange polynomials are used for polynomial interpolation. For a given set of distinct points x_j and numbers y_j, the Lagrange polynomial is the polynomial of lowest degree that assumes at each point x_j the corresponding value y_j (i.e. the functions coincide at each point). The interpolating polynomial of the least degree is unique, however, and since it can be arrived at through multiple methods, referring to "the Lagrange polynomial" is perhaps not as correct as referring to "the Lagrange form" of that unique polynomial.

Linear interpolation

Suppose that we have two points (x_0, y_0) and (x_1, y_1) where $x_0 \neq x_1$. We will define the linear Lagrange interpolating polynomial to be the straight line that passes through both of these points. Let's construct this straight line. We first note that the slope of this line will be $\frac{y_1 - y_0}{x_1 - x_0}$, and so in point-slope form we have that:

$$\begin{aligned}
y - y_1 &= \frac{y_1 - y_0}{x_1 - x_0}(x - x_1) \\
y &= y_1 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_1) \\
y &= \frac{y_1(x_1 - x_0) + (y_1 - y_0)(x - x_1)}{x_1 - x_0} \\
y &= \frac{y_1x_1 - y_1x_0 + y_1x - y_1x_1 - y_0x + y_0x_1}{x_1 - x_0} \\
y &= \frac{-y_1x_0 + y_1x - y_0x + y_0x_1}{x_1 - x_0} \\
y &= \frac{y_1(x - x_0) + y_0(x_1 - x)}{x_1 - x_0} \\
y &= y_0\frac{(x_1 - x)}{x_1 - x_0} + y_1\frac{(x - x_0)}{(x_1 - x_0)} \\
y &= y_0\left(\frac{x - x_1}{x_0 - x_1}\right) + y_1\left(\frac{x - x_0}{x_1 - x_0}\right)
\end{aligned}$$

If we let $L_0(x) = \frac{x-x_1}{x_0-x_1}$ and $L_1(x) = \frac{x-x_0}{x_1-x_0}$, then the polynomial above can be rewritten as $P_1(x) = y_0L_0(x) + y_1L_1(x)$. We note that indeed this function passes through the points (x_0, y_0) and (x_1, y_1) since $P_1(x_0) = y_0$ and $P_1(x_1) = y_1$. We formally define this polynomial below.

Definition 2.2

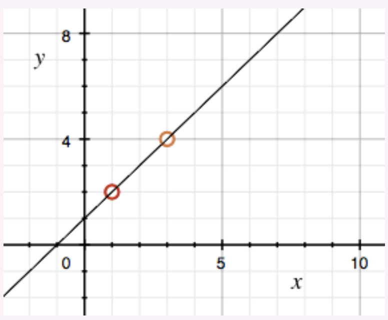
The Linear Lagrange Interpolating Polynomial that passes through the points (x_0, y_0) and (x_1, y_1) is $P_1(x) = y_0L_0(x) + y_1L_1(x)$.

which is the formula for linear interpolation in the interval (x_0, x_1). Outside this interval, the formula is identical to linear extrapolation. Computation reveals that $L_0(x_1) = 1, L_0(x_2) = 0, L_1(x_1) = 0,$ and $L_1(x_2) = 1$. The terms $L_0(x)$ and $L_1(x)$ are called **Lagrange coefficient polynomials** based on the nodes x_0 and x_1. Let's now look at some examples of applying linear Lagrange interpolating polynomials.

Example 2.7

Find the linear Lagrange interpolating polynomial, $P_1(x)$, that passes through the points $(1, 2)$ and $(3, 4)$. The function P_1 can be obtained directly by substituting the the points $(1, 2)$ and $(3, 4)$ into the formula above to get:

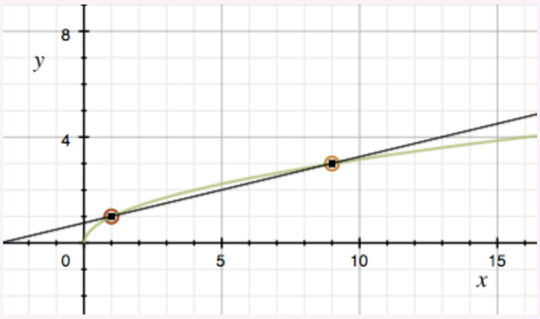
$$P_1(x) = \frac{4(x-1) + 2(3-x)}{3-1} = \frac{4x - 4 + 6 - 2x}{2} = \frac{2x+2}{2} = x + 1$$



Example 2.8

Estimate the value of $\sqrt{5}$ *using the linear Lagrange interpolating polynomial* $P_1(x)$ *that passes through the points* $(1,1)$ *and* $(9,3)$ *and evaluate the error of this approximation with the true value of* $\sqrt{5} \approx 2.23606\cdots$.
Note that $(1,1)$ *and* $(9,3)$ *are points on the function* $f(x) = \sqrt{x}$*. We first set up the linear Lagrange interpolating polynomial* $P_1(x)$ *as follows:*

$$P_1(x) = \frac{3(x-1)+1(9-x)}{9-1} = \frac{3x-3+9-x}{8} = \frac{2x+6}{8} = \frac{x+3}{4}$$



Now our approximation of $f(5) = \sqrt{5}$ *is given by* $P_1(5)$*:*

$$P_1(5) = \frac{5+3}{4} = 2$$

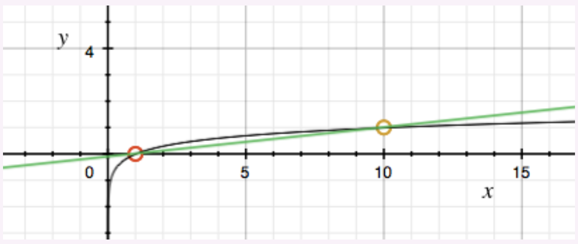
As we can see, our approximation is an underestimate of the true value of $\sqrt{5}$*. We only obtained one significant digit of accuracy.*

Example 2.9

Consider the function $y = \log(x)$*. Find the linear Lagrange polynomial* P_1 *that interpolates the points* $(1,0)$ *and* $(10,1)$*. Use* P_1 *to approximate the value of* $\log(2) \approx 0.301029\cdots$ *. Using the formula above we have that:*

$$P_1(x) = \frac{1(x-1)+0(10-x)}{10-1} = \frac{x-1}{9}$$

We have that $P_1(2) = \frac{1}{9} = 0.111\cdots$ *. As we can see, using* $P_1(2)$ *to approximate* $\log(2)$ *is not that accurate.*

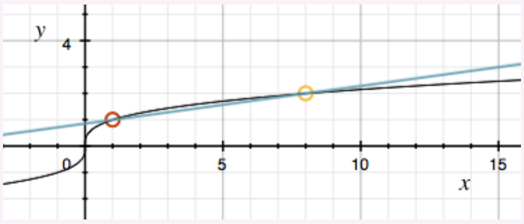


Example 2.10

Consider the function $y = \sqrt[3]{x}$. Find the linear Lagrange interpolating polynomial P_1 that interpolates the points $(1, 1)$ and $(8, 2)$. Use P_1 to approximate the value of $\sqrt[3]{3} \approx 1.44224$. Using the formula above we have that:

$$P_1(x) = \frac{2(x-1) + 1(8-x)}{8-1} = \frac{x+6}{7}$$

We have that $P_1(3) = 9/7 \approx 1.2857\cdots$, so using $P_1(3)$ to approximate $\sqrt[3]{3}$ is somewhat accurate.



In general, if $y_0 = f(x_0)$ and $y_1 = f(x_1)$ for some function f, then $P_1(x)$ is a linear approximation of $f(x)$ for all $x \in [x_0, x_1]$.

Quadratic Interpolation

If $(x_0, y_0), (x_1, y_1), (x_2, y_2)$, are given data points where $x_0 \neq x_1 \neq x_2$, then the quadratic polynomial passing through these points can be expressed as

$$\begin{aligned} P_2(x) &= y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) \\ &= y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \end{aligned}$$

where

$$\begin{aligned} L_0(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\ L_1(x) &= \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \\ L_2(x) &= \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \end{aligned}$$

Note that P_2 does in fact pass through all the points specified above since $P_2(x_0) = y_0, P_2(x_1) = y_1$, and $P_2(x_2) = y_2$. A formal definition of the polynomial above is given below.

Definition 2.3

*The **Quadratic Lagrange Interpolating Polynomial** through the points $(x_0, y_0), (x_1, y_1)$, and (x_2, y_2) where x_0, x_1, and x_2 are distinct is the polynomial $P_2(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$.*

It is important to note that while we define P_2 to be the "quadratic" Lagrange interpolating polynomial, it is possible that P_2 may have degree less than 2.

Remark: the functions L_0, L_1, L_2 are called **Lagrange's interpolating basis functions** and that

$$L_j(x_i) = \delta_{ji} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

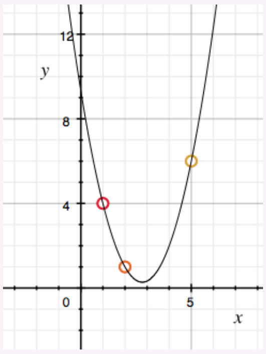
where δ_{ij} is the Kronecker delta. Let's now look at some examples of constructing a quadratic Lagrange interpolating polynomials.

Example 2.11

Construct the quadratic Lagrange interpolating polynomial $P_2(x)$ that interpolates the points $(1,4),(2,1)$, and $(5,6)$.
Applying the formula given above directly and we get that:

$$P_2(x) = 4\frac{(x-2)(x-5)}{(1-2)(1-5)} + 1\frac{(x-1)(x-5)}{(2-1)(2-5)} + 6\frac{(x-1)(x-2)}{(5-1)(5-2)}$$
$$P_2(x) = (x-2)(x-5) - \frac{1}{3}(x-1)(x-5) + \frac{1}{2}(x-1)(x-2)$$

The graph of $y = P_2(x)$ is given below:

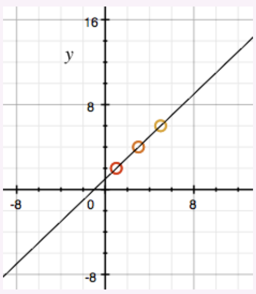


Example 2.12

Construct the quadratic Lagrange interpolating polynomial $P_2(x)$ that interpolates the points $(1,2),(3,4)$, and $(5,6)$.
Applying the formula given above directly and we get that:

$$P_2(x) = 2\frac{(x-3)(x-5)}{(1-3)(1-5)} + 4\frac{(x-1)(x-5)}{(3-1)(3-5)} + 6\frac{(x-1)(x-3)}{(5-1)(5-3)}$$
$$P_2(x) = \frac{1}{4}(x-3)(x-5) - (x-1)(x-5) + \frac{3}{4}(x-1)(x-3)$$

The graph of $y = P_2(x)$ is given below:

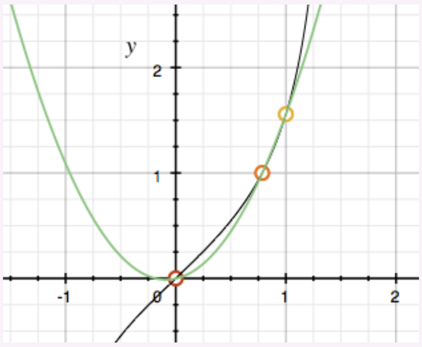


Note that in this example we shows that P_2 need not be quadratic and may be a polynomial of lesser degree.

Example 2.13

Find the quadratic Lagrange interpolating polynomial P_2 that interpolates the function $y = \tan x$ at the points $\left(\frac{\pi}{4}, 1\right)$, and $(1, \tan(1))$.
Applying the formula above and we have that:

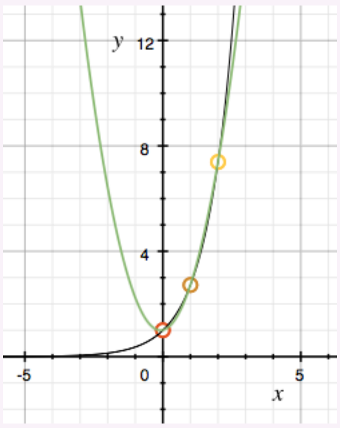
$$P_2(x) = \frac{(x-0)(x-1)}{\left(\frac{\pi}{4}-0\right)\left(\frac{\pi}{4}-1\right)} + \tan(1)\frac{(x-0)\left(x-\frac{\pi}{4}\right)}{(1-0)\left(1-\frac{\pi}{4}\right)}$$



Example 2.14

Find the quadratic Lagrange interpolating polynomial P_2 that interpolates the function $y = e^x$ at the points $(0, 1), (1, e)$, and $(2, e^2)$.
Applying the formula above and we have that:

$$P_2(x) = \frac{(x-1)(x-2)}{(0-1)(0-2)} + e\frac{(x-0)(x-2)}{(1-0)(1-2)} + e^2\frac{(x-0)(x-1)}{(2-0)(2-1)}$$



Higher-Degree Interpolation

Suppose now that we have $n+1$ points $(x_0, y_0), (x_1, y_1), \cdots, (x_n, y_n)$ where $x_0, x_1, \cdots, x_n$ are distinct numbers. Then we would need a polynomial with degree n or less to interpolate these points. To match with our definitions of linear and quadratic interpolating polynomials, we wish to obtain a polynomial function P_n in terms of functions $L_0, L_1, \cdots, L_n$ such that:

$$P_n(x) = y_0 L_0(x) + y_1 L_1(x) + \cdots + y_n L_n(x) \tag{2.1}$$

Where

$$
\begin{aligned}
L_0(x) &= \frac{(x-x_1)(x-x_2)(x-x_3)\cdots(x-x_n)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)\cdots(x_0-x_n)}\\
L_1(x) &= \frac{(x-x_0)(x-x_2)(x-x_3)\cdots(x-x_n)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)\cdots(x_1-x_n)}\\
L_2(x) &= \frac{(x-x_0)(x-x_1)(x-x_3)\cdots(x-x_n)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)\cdots(x_2-x_n)}\\
&\vdots\\
L_n(x) &= \frac{(x-x_0)(x-x_1)(x-x_2)\cdots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})}
\end{aligned}
$$

The numerator of $L_i(x)$ does not contain $(x-x_i)$.
The denominator of $L_i(x)$ does not contain (x_i-x_i).

$$P_n(x) = \sum_{k=0}^{n} y_k L_k(x)$$

Thus for each $k = 0, 1, \cdots, n$, define L_k as follows:

$$L_k(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)} = \prod_{i=0\, i\neq k}^{n} \frac{(x-x_i)}{(x_k-x_i)}$$

Note that from the definition of L_k above that we can easily obtain the general formulas for the linear Lagrange interpolating polynomials that pass through the points (x_0, y_0) and (x_1, y_1) (x_0 and x_1 distinct) and the quadratic linear Lagrange interpolating polynomial that passes through the points $(x_0, y_0), (x_1, y_1)$, and (x_2, y_2) (x_0, x_1, and x_2 distinct). We will now formally define higher order Lagrange interpolating polynomials.

Definition 2.4

The n^{th} Order Lagrange Interpolating Polynomial that passes through the $n+1$ points $(x_0, y_0), (x_1, y_1), \cdots (x_n, y_n)$ where $x_0, x_1, \cdots, x_n$ are distinct numbers is $P_n(x) = y_0L_0(x) + y_1L_1(x) + \cdots + y_nL_n(x)$.

For all $i \neq k$, $L_k(x)$ includes the term $(x - x_i)$ in the numerator, so the whole product will be zero at $x = x_i$:

$$L_{k\neq i}(x_i) = \prod_{m\neq k} \frac{x_i - x_m}{x_j - x_m} = \frac{(x_i-x_0)}{(x_j-x_0)} \cdots \frac{(x_i-x_i)}{(x_j-x_i)} \cdots \frac{(x_i-x_n)}{(x_k-x_n)} = 0.$$

On the other hand,

$$L_i(x_i) := \prod_{m\neq i} \frac{x_i - x_m}{x_i - x_m} = 1$$

In other words, all basis polynomials are zero at $x = x_i$, except $L_i(x)$, for which it holds that $L_i(x_i) = 1$, because it lacks the $(x - x_i)$ term.
It follows that $y_iL_i(x_i) = y_i$, so at each point x_i ,$L(x_i) = y_i + 0 + 0 + \cdots + 0 = y_i$, showing that L interpolates the function exactly.

Example 2.15

Approximate function $y = f(x) = x\sin(2x+\pi/4)+1$ *by a polynomial* p_n *of degree* $n = 3$*, based on the following* $n+1 = 4$ *points:*

i	0	1	2	3
x_i	-1	0	1	2
$y_i = f(x_i)$	1.937	1.000	1.349	-0.995

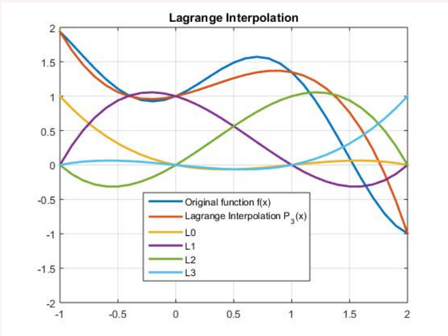
Based on these points, we construct the Lagrange polynomials as the basis functions of the polynomial space (instead of the power functions in the previous example):

$$L_0(x) = \frac{(x-0)(x-1)(x-2)}{-6} = \frac{x^3-3x^2+2x}{-6}$$
$$L_1(x) = \frac{(x+1)(x-1)(x-2)}{2} = \frac{x^3-2x^2-x+2}{2}$$
$$L_2(x) = \frac{(x+1)(x-0)(x-2)}{-2} = \frac{x^3-x^2-2x}{-2}$$
$$L_3(x) = \frac{(x+1)(x-0)(x-1)}{6} = \frac{x^3-x}{6}$$

Note that indeed $L_0(x)+L_1(x)+L_2(x)+L_3(x) = 1$*. The interpolating polynomial can be obtained as a weighted sum of these basis functions:*

$$L_3(x) = 1.937L_0(x) + 1.0L_1(x) + 1.349L_2(x) - 0.995L_3(x) = 1.0 + 0.369x + 0.643x^2 - 0.663x^3$$

which is the same as $P_3(x)$ *previously found based on the power basis functions, with the same error* $\epsilon = 0.3063$.



Example 2.16

Find the quartic Lagrange interpolating polynomial, $P_4(x)$, that interpolates the points $(-1,4),(1,1),(2,0),(6,4)$, and $(7,-1)$.
Applying the formula above directly and we get that:

$$P_4(x) = y_0L_0(x) + y_1L_1(x) + y_2L_2(x) + y_3L_3(x) + y_4L_4(x)$$

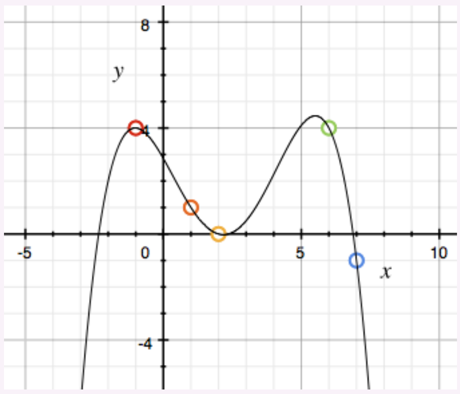
$$P_4(x) = y_0\frac{\prod_{j=0,j\neq 0}^{n}(x-x_j)}{\prod_{j=0,j\neq 0}^{n}(x_0-x_j)} + y_1\frac{\prod_{j=0,j\neq 1}^{n}(x-x_j)}{\prod_{j=0,j\neq 1}^{n}(x_1-x_j)} + y_2\frac{\prod_{j=0,j\neq 2}^{n}(x-x_j)}{\prod_{j=0,j\neq 2}^{n}(x_2-x_j)}$$
$$+ y_3\frac{\prod_{j=0,j\neq 3}^{n}(x-x_j)}{\prod_{j=0,j\neq 3}^{n}(x_3-x_j)} + y_4\frac{\prod_{j=0,j\neq 4}^{n}(x-x_j)}{\prod_{j=0,j\neq 4}^{n}(x_4-x_j)}$$

$$P_4(x) = 4\frac{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)(x_0-x_4)} + 1\frac{(x-x_0)(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)(x_1-x_4)}$$
$$+ 4\frac{(x-x_0)(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)(x_3-x_4)} - 1\frac{(x-x_0)(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_0)(x_4-x_1)(x_4-x_2)(x_4-x_3)}$$

$$P_4(x) = 4\frac{(x-1)(x-2)(x-6)(x-7)}{(-1-1)(-1-2)(-1-6)(-1-7)} + 1\frac{(x+1)(x-2)(x-6)(x-7)}{(1+1)(1-2)(1-6)(1-7)}$$
$$+ 4\frac{(x+1)(x-1)(x-2)(x-7)}{(6+1)(6-1)(6-2)(6-7)} - 1\frac{(x+1)(x-1)(x-2)(x-6)}{(7+1)(7-1)(7-2)(7-6)}$$

$$P_4(x) = \frac{1}{84}(x-1)(x-2)(x-6)(x-7) - \frac{1}{60}(x+1)(x-2)(x-6)(x-7)$$
$$- \frac{1}{35}(x+1)(x-1)(x-2)(x-7) - \frac{1}{240}(x+1)(x-1)(x-2)(x-6)$$

The graph of $y = P_4(x)$ is given below:



Theorem 2.1

Let f be a function that has $n+1$ continuous derivatives on the interval $[a,b]$, and let $(x_0,y_0),(x_1,y_1),\cdots,(x_n,y_n)$ be $n+1$points where $x_{0,x}1,\cdots,x_n$ are distinct numbers. Let P_n be the nth order Lagrange interpolating polynomial for these points. Then for each $x \in [a,b]$ there exists ξ_n between $m = \min\{x_0,x_1,\cdots,x_n,x\}$ and $M = \max\{x_0,x_1,\cdots,x_n,x\}$ such that $f(x) - P_n(x) = \frac{f^{(n+1)}(\xi_n)}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n)$.

Note that we can bound this error formula above by maximizing the error on the interval

$[a, b]$, that is find $x \in [a, b]$ such that:

$$\begin{aligned} \mid f(x) - P_n(x) \mid &= \left| \frac{f^{(n+1)}(\xi_n)}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n) \right| \\ &\leq \max_{m \leq x \leq M} \left| \frac{f^{(n+1)}(x)}{(n+1)!} \right| \max_{m \leq x \leq M} \mid (x-x_0)(x-x_1)\cdots(x-x_n) \mid \end{aligned}$$

2.2.3 Newton divided difference interpolation

Divided Difference

Let $f(x_0), f(x_1), \cdots, f(x_n)$ be the values of a function f corresponding to the arguments $x_0, x_1, \cdots, x_n$ where the intervals $x_1 - x_0, x_2 - x_1, \cdots, x_n - x_{n-1}$ are not necessarily equally spaced. Then the **first divided difference** of f for the arguments $x_0, x_1, \cdots, x_n$ are defined by ,

$$\begin{aligned} f[x_0, x_1] &= \frac{f[x_1] - f[x_0]}{x_1 - x_0} \\ f[x_1, x_2] &= \frac{f[x_2] - f[x_1]}{x_2 - x_1} \end{aligned}$$

The second divided difference of f for three arguments x_0, x_1, x_2 is defined by

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

and similarly the divided difference of order n is defined by

$$f[x_0, x_1, \cdots, x_n] = \frac{f[x_1, x_2, \cdots, x_n] - f[x_0, x_1, \cdots, x_{n-1}]}{x_n - x_0}$$

Properties:

- The divided differences are symmetrical in all their arguments; that is, the value of any divided difference is independent of the order of the arguments.
- The divided difference operator is linear.
- The nth order divided differences of a polynomial of degree n are constant, equal to the coefficient of x^n.

Newton's Divided Difference Interpolation

A major difficulty with the Lagrange Interpolation is that one is not sure about the degree of interpolating polynomial needed to achieve a certain accuracy. Thus, if the accuracy is not good enough with polynomial of a certain degree, one needs to increase the degree of the polynomial, and computations need to be started all over again. Furthermore, computing various Lagrangian polynomials is an expensive procedure. It is, indeed, desirable to have a formula which makes use of $P_{n-1}(x)$ in computing $P_n(x)$.

The following form of interpolation, known as Newton's interpolation allows us to do so. The idea is to obtain the interpolating polynomial $P_n(x)$ in the following form:

$$P_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \cdots + a_n(x-x_0)(x-x_1)\cdots(x-x_{n-1})$$

The constants a_0 through a_n can be determined as follows:

Newton's Divided Difference Interpolating Polynomial Or Newton's Form

Define

$$
\begin{aligned}
P_1(x) &= f(x_0) + f[x_0, x_1](x - x_0) \\
P_2(x) &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_1)(x - x_0) \\
&= P_1(x) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
P_3(x) &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_1)(x - x_0) \\
&+ f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) \\
&= P_2(x) + f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) \\
&\vdots \\
P_n(x) &= P_{n-1}(x) + f[x_0, x_1, \cdots, x_{n-1}](x - x_0)(x - x_1)\cdots(x - x_{n-1})
\end{aligned}
$$

The polynomial P_n is called Newton's divided deference formula for the interpolating polynomial or Newton's form for the interpolating polynomial. Note that $P_n(x_i) = f(x_i)$.

Newton's divided deference interpolating polynomial defined as

$$
\begin{aligned}
P_n(x) =& f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots \\
&+ f[x_0, x_1, x_2, \cdots, x_n](x - x_0)(x - x_1)(x - x_2)\cdots(x - x_{n-1})
\end{aligned}
$$

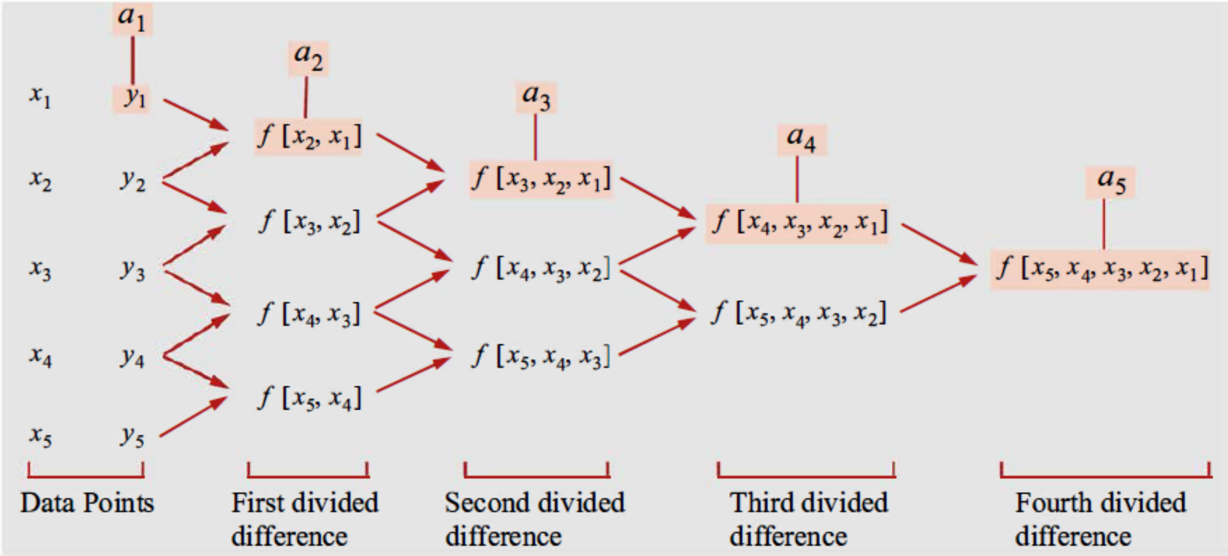
Where

$$
\begin{aligned}
a_0 &= f[x_0] = y_0 \quad \text{because } f[x_i] = y_i \quad \text{by definition} \\
a_1 &= f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0} \\
a_2 &= f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} \\
&\vdots \\
a_k &= f[x_0, x_1, x_2, \ldots, x_{k-1}, x_k] = \frac{f[x_1, x_2, \ldots, x_{k-1}, x_k] - f[x_0, x_1, x_2, \ldots, x_{k-1}]}{x_k - x_0}
\end{aligned}
$$

A difference table is again a convenient device for displaying differences, the standard diagonal form being used and thus the generation of the divided differences is outlined in Table below.

x	$f(x)$	First Divided Difference	Second Divided Difference	Third Divided Difference
x_0	$f[x_0]$			
		$f[x_0, x_1] = \frac{f[x_1]-f[x_0]}{x_1-x_0}$		
x_1	$f[x_1]$		$f[x_0, x_1, x_2] = \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}$	
		$f[x_1, x_2] = \frac{f[x_2]-f[x_1]}{x_2-x_1}$		$f[x_0, x_1, x_2, x_3] = \frac{f[x_1,x_2,x_3]-f[x_0,x_1,x_2]}{x_3-x_0}$
x_2	$f[x_2]$		$f[x_1, x_2, x_3] = \frac{f[x_2,x_3]-f[x_1,x_2]}{x_3-x_1}$	
		$f[x_2, x_3] = \frac{f[x_3]-f[x_2]}{x_3-x_2}$		$f[x_1, x_2, x_3, x_4] = \frac{f[x_2,x_3,x_4]-f[x_1,x_2,x_3]}{x_4-x_1}$
x_3	$f[x_3]$		$f[x_2, x_3, x_4] = \frac{f[x_3,x_4]-f[x_2,x_3]}{x_4-x_2}$	
		$f[x_3, x_4] = \frac{f[x_4]-f[x_3]}{x_4-x_3}$		
x_4	$f[x_4]$			

A table for solving the coefficients of a Newton's polynomial.

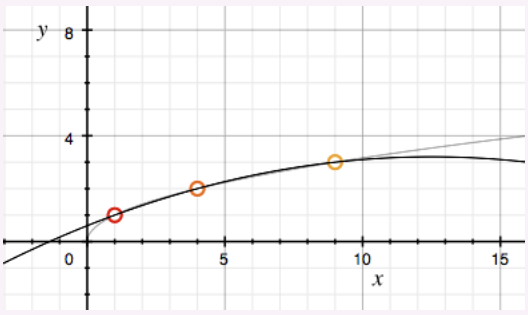


Example 2.17

Find the polynomial P_2 for the function $y = \sqrt{x}$ that interpolates the points $(1, 1), (4, 2)$, and $(9, 3)$ using Newton's divided difference formula.
Applying Newton's divided difference formula from above and we get that:

$$P_2(x) = f(x_0) + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2]$$
$$P_2(x) = 1 + (x - 1)\frac{f(x_1) - f(x_0)}{x_1 - x_0} + (x - 1)(x - 4)\frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$
$$P_2(x) = 1 + (x - 1)\frac{2 - 1}{4 - 1} + (x - 1)(x - 4)\frac{\frac{3-2}{9-4} - \frac{2-1}{4-1}}{9 - 1}$$
$$P_2(x) = 1 + \frac{1}{3}(x - 1) - \frac{1}{60}(x - 1)(x - 4)$$

The graph of $y = P_2(x)$ is given below.



Example 2.18

For example, given data points $(1, 6)$, $(2, 11)$, $(3, 18)$, and $(4, 27)$ we can draw the following table:

Example

i	x_i	y_i	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$
0	$x_0 = 1$	$y_0 = f[x_0] = 6$	$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{11-6}{2-1} = 5$	$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{7-5}{3-1} = 1$	$f[x_0, x_1, x_2, x_3] = \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0} = \frac{1-1}{4-1} = 0$
1	$x_1 = 2$	$y_1 = f[x_1] = 11$	$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{18-11}{3-2} = 7$	$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = \frac{9-7}{4-2} = 1$	
2	$x_2 = 3$	$y_2 = f[x_2] = 18$	$f[x_2, x_3] = \frac{f[x_3] - f[x_2]}{x_3 - x_2} = \frac{27-18}{4-3} = 9$		
3	$x_3 = 4$	$y_3 = f[x_3] = 27$			

The four data points lie on a polynomial of order 2, which is why the coefficient $a_3(f[x_0, x_1, x_2, x_3])$ is zero. Given $[a_0, a_1, a_2] = [6, 5, 1]$ the result polynomial is:

$$\begin{aligned} y &= a_0 + a_1(x - x_0) + a_2(x - x_1)(x - x_0) \\ &= 6 + 5 \times (x - 1) + 1 \times (x - 2)(x - 1) \\ &= x^2 + 2x + 3 \end{aligned}$$

Example 2.19

For a function f, the divided differences are given by

$$x_1 = 2 \quad f[x_1] = 2$$

$$x_0 = 1 \quad f[x_0] = -6$$

$$x_2 = 4 \quad f[x_2] = 12$$

find $f[x_0, x_1, x_2]$.
Solution:

x	$f(x)$	*First Divided Difference*	*Second Divided Difference*
$x_1 = 2$	$f[x_1] = 2$		
		$f[x_1, x_0] = \frac{(-6)-2}{1-2} = 8$	
$x_0 = 1$	$f[x_0] = -6$		$f[x_1, x_0, x_2] = \frac{6-8}{4-2} = -1$
		$f[x_0, x_2] = \frac{12-(-6)}{4-1} = 6$	
$x_2 = 4$	$f[x_2] = 12$		

Hence, $f[x_1, x_0, x_2] = -1$ and by symmetry property we know that $f[x_1, x_0, x_2] = f[x_0, x_1, x_2]$, Hence $f[x_0, x_1, x_2] = -1$.

Example 2.20

For a function f, the divided differences are given by

$$\begin{array}{llll} x_0 = 0.0 & f[x_0] & & \\ & & f[x_0, x_1] & \\ x_1 = 0.4 & f[x_1] & & f[x_0, x_1, x_2] = \frac{50}{7} \\ & & f[x_1, x_2] = 10 & \\ x_2 = 0.7 & f[x_2] = 6 & & \end{array}$$

Determine the missing entries in the table.
Solution: *We have the formula*

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

and substituting gives

$$\frac{50}{7} = \frac{(10 - f[x_0, x_1])}{0.7}.$$

Thus,

$$f[x_0, x_1] = -0.7 \cdot (\frac{50}{7}) + 10 = 5$$

Using the formula

$$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1}$$

and substituting gives

$$10 = \frac{(6 - f[x_1])}{0.3}.$$

Thus,

$$f[x_1] = 6 - 3 = 3.$$

Further,

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}.$$

So,

$$5 = \frac{(3 - f[x_0])}{0.4}.$$

Thus,

$$f[x_0] = 3 - 2 = 1.$$

Example 2.21

The set of the following five data points is given:

x	1	2	4	5	7
y	52	5	-5	-40	10

a) *Determine the fourth-order polynomial in Newton's form that passes through the points. Calculate the coefficients by using a divided difference table.*

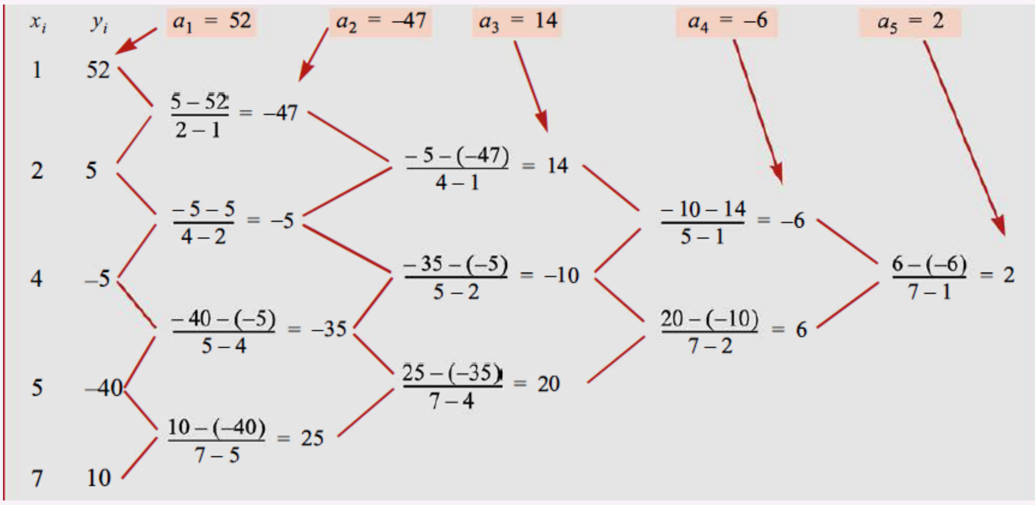
b) *Use the polynomial obtained in part (a) to determine the interpolated value for $x = 3$.*

Example: Solution

a) *Newton's polynomial for the given points has the form:*

$$f(x) = y = a_1 + a_2(x-1) + a_3(x-1)(x-2) + a_4(x-1)(x-2)(x-4) + a_5(x-1)(x-2)(x-4)(x-5)$$

The coefficients can be determined by the following divided difference table:



With the coefficients determined, the polynomial is:

$$f(x) = y = 52 - 47(x-1) + 14(x-1)(x-2) - 6(x-1)(x-2)(x-4) + 2(x-1)(x-2)(x-4)(x-5)$$

(b) *The interpolated value for x= 3 is obtained by substituting for x in the polynomial:*

$$f(3) = y = 52 - 47(3-1) + 14(3-1)(3-2) - 6(3-1)(3-2)(3-4) + 2(3-1)(3-2)(3-4)(3-5) = 6$$

2.3 Spline Interpolation

Many scientific and engineering phenomena being measured undergo a transition from one physical domain to another. Data obtained from these measurements are better represented by a set of piecewise continuous curves rather than by a single curve. One of the difficulties with polynomial interpolation is that in some cases the oscillatory nature of high-degree polynomials can induce large fluctuations over the entire range when approximating a set of data points. One way of solving this problem is to divide the interval into a set of subintervals and construct a lowerdegree approximating polynomial on each subinterval. This type of approximation is called piecewise polynomial interpolation.

Piecewise polynomial functions, especially spline functions, have become increasingly popular. Most of the interest has centered on cubic splines because of the ease of their applications to a variety of fields, such as the solution of boundary value problems for differential equations and the method of finite elements for the numerical solution of partial differential equations.

We start out with the general definition of spline functions. Let f be a real-valued function defined on some interval $[a, b]$ and let the set of data points $(a = x_1, f(x_1)), (x_2, f(x_2)), \cdots, (x_n.f(x_n))$ be given. For simplicity, assume that

$$a = x_1 < x_2 < \cdots < x_n = b.$$

We have the definition:

Definition 2.5:

function S is called a spline of degree k if it satisfies the following conditions:

1. *S is defined in the interval $[a, b]$.*
2. *$S^{(r)}$ is continuous on $[a, b]$ for $0 \leq r \leq k-1$.*
3. *S is a polynomial of degree $\leq k$ on each subinterval $[x_i, x_{i+1}], i = 1, 2, \cdots, n-1$.*

2.3.1 Linear Spline

The simplest connection between two points is a straight line. The first-order splines for a group of ordered data points can be defined as a set of linear functions by a series of straight lines as shown in Figure 2.1.

Using the formula of the equation of the line, it is easy to see that the function $S(x)$ is defined by

$$\begin{aligned} S1(x) &= f(x_1) + m_1(x - x_1),\ x_1 \leq x_2 \\ S_2(x) &= f(x_2) + m_2(x - x_2),\ x_2 \leq x_3 \\ S_3(x) &= f(x_3) + m_3(x - x_3),\ x_3 \leq x_4 \\ &\vdots \\ S_i(x) &= f(x_i) + m_i(x - x_i),\ x_{i-1} \leq x_i \end{aligned} \tag{2.2}$$

where m_i is the slope of the straight line connecting the points:

$$m_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = f[x_{i+1}, x_i]$$

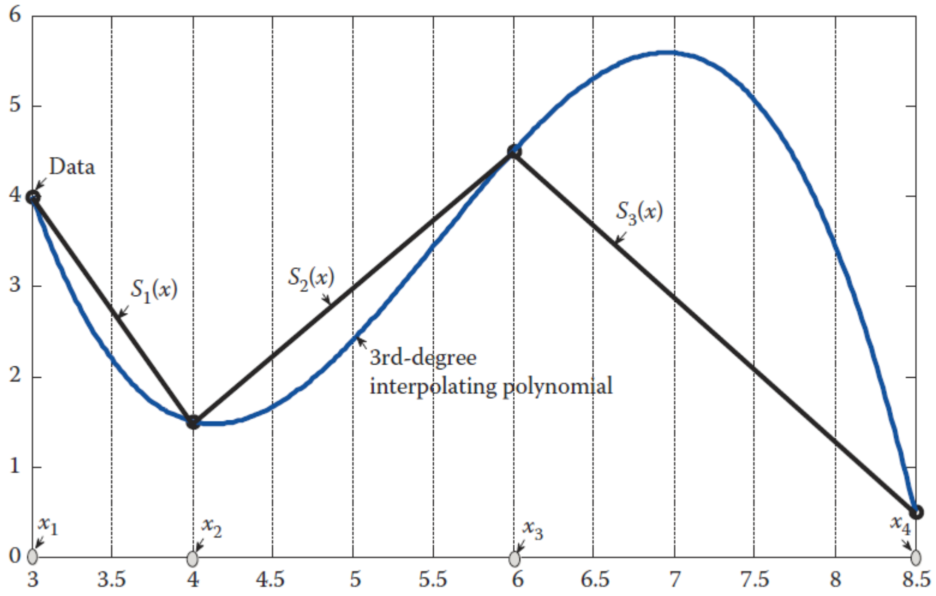


Figure 2.1: Linear Spline

Generally linear spline is defied as

$$S_i(x) = f(x_i) + f[x_{i+1}, x_i](x - x_i), \text{on each subinterval } [x_i, x_{i+1}].$$

Outside the interval $[a, b], S(x)$ is usually defined by

$$S(x) = \begin{cases} S_1(x), & \text{if } \ x < a \\ S_{n-1}(x) & \text{if } \ x > b \end{cases} .$$

The points $x_2, x_3, \cdots, x_{n-1}$, where $S(x)$ changes from one polynomial to another, are called the **breakpoints or knots**. Because $S(x)$ is continuous on $[a, b]$, it is called a **spline of degree 1**.

Example 2.22

The set of the following four data points is given:

x	8	11	15	18
f(x)	5	9	10	8

(a) *Determine the linear splines that fit the data.*

(b) *Determine the interpolated value for* $x = 12.7$.

SOLUTION:

(a) There are four points and thus three splines. Using Eq. (2.2) the equations of the splines are:

$$S_1(x) = f(8) + \frac{f(11) - f(8)}{11 - 8}(x - 8) = 8 + \frac{9 - 5}{11 - 8}(x - 8) = \frac{4x}{3} - \frac{8}{3}, \text{ for } 8 \le x \le 11$$
$$S_2(x) = f(11) + \frac{f(15) - f(11)}{15 - 11}(x - 11) = 9 + \frac{10 - 9}{15 - 11}(x - 11) = \frac{x}{4} + \frac{25}{4}, \text{ for } 11 \le x \le 15$$
$$S_3(x) = f(15) + \frac{f(18) - f(15)}{18 - 15}(x - 15) = 10 + \frac{8 - 10}{18 - 15}(x - 15) = -\frac{2x}{3}, \text{ for } 15 \le x \le 18$$

(b) The value $x = 12.7$ lies in $[11, 15]$, since the interpolated value of $f(x)$ for $x = 12.7$ is obtained by substituting the value of x in the equation for $S_2(x)$ above:

$$\frac{x}{4} + \frac{25}{4} = \frac{12.7}{4} + \frac{25}{4} = \frac{37.7}{4} = 9.425$$

2.3.2 Quadratic Spline

In many cases, linear piecewise polynomials are unsatisfactory when being used to interpolate the values of a function, which deviate considerably from a linear function. In such cases, piecewise polynomials of higher degree are more suitable to use to approximate the function. In this section, we shall discuss the simplest type of differentiable, piecewise polynomial functions, known as **quadratic splines**. As before, consider the subdivision

$$a = x_1 < x_2 < \cdots < x_n = b.$$

where $x_1, \cdots, x_n$ are given. For piecewise linear interpolation, we choose two points $(x_i, f(x_i))$ and $(x_{i+1}, f(x_{i+1}))$ in the subinterval $[x_i, x_{i+1}]$ and draw a line through those two points to interpolate the data. This approach is easily extended to construct the quadratic splines. Instead of choosing two points, we choose three points in the subinterval $[x_i, x_{i+1}]$ and pass a second-degree polynomial through these points as shown in Figure 2.2. We shall show that there is only one such polynomial. To construct a quadratic spline $S(x)$, we first define a

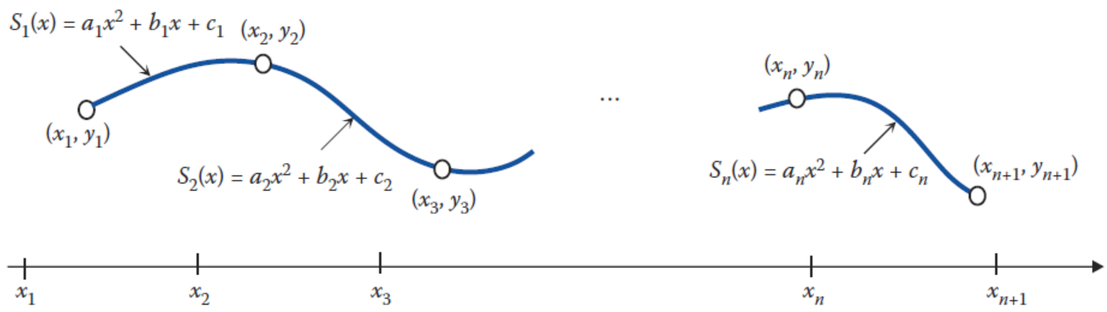


Figure 2.2: Quadratic spline

quadratic function in each subinterval $[x_i, x_{i+1}]$ by

$$S_i(x) = c_i + b_i(x - x_i) + a_i(x - x_i)^2 \tag{2.3}$$

where a_i, b_i, and c_i are constants to be determined.

Now by Definition 2.3, $S(x)$ **must satisfy the conditions**

$$S(x) = S_i(x) \text{ on } [x_i, x_{i+1}] \quad \text{for } i = 1, 2, \cdots, n-1. \tag{2.4}$$

$$S_i(x_i) = f(x_i) \tag{2.5}$$

$$S_i(x_{i+1}) = f(x_{i+1}). \tag{2.6}$$

$S'(x)$ **is continuous on** $[a, b]$ **if**

$$S_i'(x_i) = d_i \quad \text{and} \quad S_i'(x_{i+1}) = d_{i+1} \tag{2.7}$$

Here the values of d_i will be defined later. Using conditions (2.5) and (2.7), it is easy to see that$s_i(x)$ is uniquely defined on $[x_i, x_{i+1}]$ by

$$S_i(x) = f(x_i) + d_i(x - x_i) + \frac{d_{i+1} - d_i}{2(x_{i+1} - x_i)}(x - x_i)^2. \tag{2.8}$$

We now use condition (2.6) to obtain d_i from the recursive formula

$$d_{i+1} = -d_i + 2\left[\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}\right], \quad i = 1, 2, \cdots, n-1 \tag{2.9}$$

with d_1 arbitrary.

Thus, given the data $(x_i, f(x_i))$ and an arbitrary value for d_1, the quadratic spline $S(x)$ is uniquely determined by formulas (2.4), (2.8), and (2.9).

Example 2.23

The set of the following four data points is given:

x	8	11	15	18
f(x)	5	9	10	8

(a) *Determine the quadratic splines that fit the data.*

(b) *Determine the interpolated value for $x = 12.7$.*

SOLUTION:

(a) To determine the quadratics spline we flowing the above approach let $d_1 = 0$, then from (2.9) we have

$$d_2 = -d_1 + 2\left[\frac{f(x_2) - f(x_1)}{x_2 - x_1}\right] = 0 + 2\left[\frac{f(11) - f(8)}{11 - 8}\right] = 2\left[\frac{9-5}{3}\right] = \frac{8}{3}$$
$$d_3 = -d_2 + 2\left[\frac{f(x_3) - f(x_2)}{x_3 - x_2}\right] = -\frac{8}{3} + 2\left[\frac{f(15) - f(11)}{15 - 11}\right] = -\frac{8}{3} + 2\left[\frac{10-9}{4}\right] = \frac{-13}{6}$$
$$d_4 = -d_3 + 2\left[\frac{f(x_4) - f(x_3)}{x_4 - x_3}\right] = \frac{13}{6} + 2\left[\frac{f(18) - f(15)}{18 - 15}\right] = \frac{13}{6} + 2\left[\frac{8-10}{3}\right] = \frac{5}{6}$$

We now use (2.8) to get a quadratic spline $S(x)$ defined by

$$S_1(x) = f(x_1) + d_1(x - x_1) + \frac{d_2 - d_1}{2(x_2 - x_1)}(x - x_1)^2 = 5 + \frac{4}{9}(x-8)^2, \quad [8, 11]$$
$$S_2(x) = f(x_2) + d_2(x - x_2) + \frac{d_3 - d_2}{2(x_3 - x_2)}(x - x_2)^2 = \frac{-61}{3} + \frac{8x}{3} - \frac{29}{48}(x-11)^2, \quad [11, 15]$$
$$S_3(x) = f(x_3) + d_3(x - x_3) + \frac{d_4 - d_3}{2(x_4 - x_3)}(x - x_3)^2 = 10 - \frac{-13}{6}(x-15) - \frac{1}{2}(x-15)^2, \quad [15, 18]$$

(b) $S_(12.7) = 11.787$

2.3.3 Cubic Spline

In cubic splines, third-degree polynomials are used to interpolate over each interval between data points. Suppose there are $n+1$ data points $(x_1, f(x_1)), \cdots (x_{n+1}, f(x_{n+1}))$ so that there are n intervals and thus n cubic polynomials. Each cubic polynomial is conveniently expressed in the form

$$S_j(x) = a_j + bj(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3, \quad \forall j \in \{0, 1, \cdots, n-1\}$$

where $a_i, b_i, c_i, d_i (i = 1, 2, \cdots, n)$ are unknown constants to be determined. Since there are n such polynomials, and each has four unknown constants, there are a total of $4n$ unknown constants. Therefore, $4n$ equations are needed to determine all the unknowns. These equations are derived based on the same logic as quadratic splines, except that second derivatives of adjacent splines also agree at the interior knots and two boundary conditions are required.

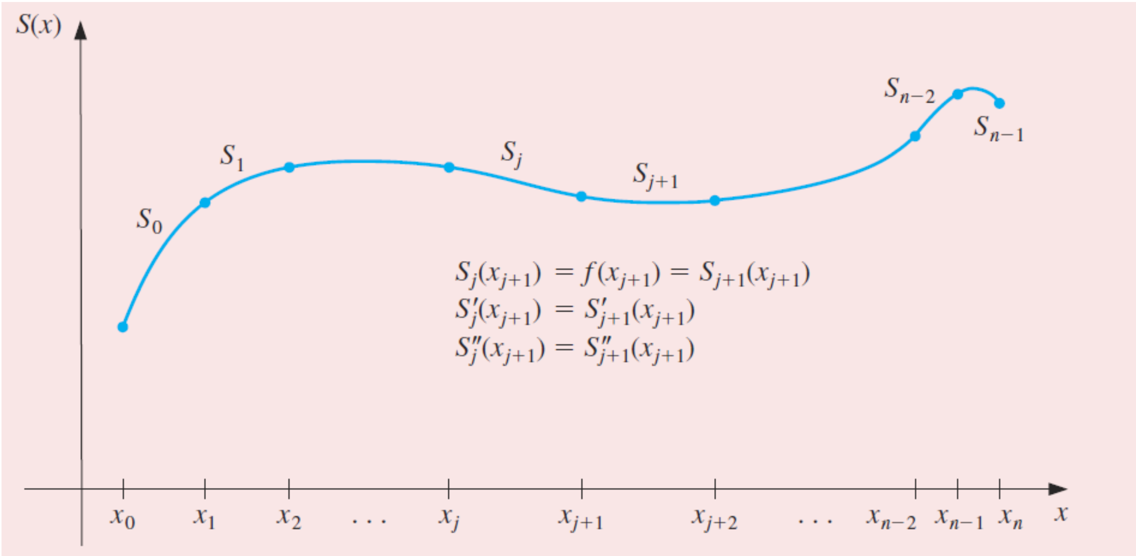


Figure 2.3: Cubic Spline

Definition 2.6

Given a function f defined on $[a, b]$ and a set of nodes $a = x_0 < x_1 < \cdots < x_n = b$, a cubic spline interpolant S for f is a function that satisfies the following conditions:

a. $S(x)$ *is a cubic polynomial, denoted* $S_j(x)$*, on the sub-interval* $[x_j, x_{j+1}]$ $\forall j = 0, 1, \cdots, n-1$.

b. $S_j(x_j) = f(x_j), \quad \forall j = 0, 1, \cdots, (n-1)$. *"Left" Interpolation*

c. $S_j(x_{j+1}) = f(x_{j+1}), \quad \forall j = 0, 1, \cdots, (n-1)$. *"Right" Interpolation*

d. $S_j'(x_{j+1}) = S_{j+1}'(x_{j+1}), \quad \forall j = 0, 1, \cdots, (n-2)$. *Slope-match*

e. $S_j''(x_{j+1}) = S_{j+1}''(x_{j+1}), \quad \forall j = 0, 1, \cdots, (n-2)$. *Curvature match*

One of the following sets of boundary conditions is satisfied:

(i) $S''(x_0) = S''(x_n) = 0$ *(natural (or free) boundary);*

(ii) $S'(x_0) = f'(x0)$ *and* $S'(x_n) = f'(x_n)$ *(clamped boundary).*

Building Cubic Splines. Applying the Conditions

We start with

$$S_j(x) = a_j + bj(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3, \quad \forall j \in \{0, 1, \cdots, n-1\} \tag{2.10}$$

and apply all the conditions to these polynomials. For convenience we introduce the notation

$$h_j = x_{j+1} - x_j.$$

b. $S_j(x_j) = a_j = f(x_j)$

c.
$$\begin{aligned} S_{j+1}(x_{j+1}) &= S_j(x_{j+1}) \\ f(x_{j+1}) &= a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 + d_j(x_{j+1} - x_j)^3 \\ a_{j+1} &= a_j + b_j h_j + c_j h_j^2 + d_j h_j^3 \end{aligned} \tag{2.11}$$

d. By differentiating Eqn. (2.10), we obtain

$$\begin{aligned} S_j'(x) &= b_j + 2c_j(x - x_j) + 3d_j(x - x_j)^2 \\ S_j''(x) &= 2c_j + 6d_j(x - x_j). \end{aligned}$$

The continuity condition on the first derivative implies

$$\begin{aligned} S_j'(x_j) &= S_{j-1}'(x_j) \\ b_j &= b_{j-1} + 2c_{j-1}(x_j - x_{j-1}) + 3d_{j-1}(x_j - x_{j-1})^2 \\ &= b_{j-1} + 2c_{j-1}h_{j-1} + 3d_{j-1}h_{j-1}^2,\ i = 2, 3, \cdots, n. \end{aligned} \tag{2.12}$$

Notice $S_j'(x_j) = b_j$, hence we get

$$b_{j+1} = b_j + 2c_j h_j + 3d_j h_j^2 \tag{2.13}$$

e. Similarly, imposing the continuity condition on the second derivative gives

$$\begin{aligned} S_j''(x_j) &= S_{j-1}''(x_j) \\ 2c_j &= 2c_{j-1} + 6d_{j-1}(x_j - x_{j-1}) \\ c_j &= c_{j-1} + 3d_{j-1}h_{j-1}, i = 2, 3, \cdots, n. \end{aligned} \tag{2.14}$$

Notice $S_j''(x_j) = 2c_j$, hence we get

$$c_{j+1} = c_j + 3d_j h_j. \tag{2.15}$$

We got a whole lot of equations to solve!!! (How many???)

Solving the Resulting Equations.
Now, by solving Eqn. (2.15) for d_j we have

$$d_j = \frac{c_{j+1} - c_j}{3h_j},$$

and substituting its value into Eqn. (2.12), we get,

$$a_{j+1} = a_j + b_j h_j + \frac{h_j^2}{3}(2c_j + c_{j+1}), \tag{2.16}$$

Similarly, solve Eqn.(2.15) for d_j and substitute its value into Eqn.(2.13) to get

$$b_{j+1} = b_j + h_j(c_j + c_{j+1}). \tag{2.17}$$

We solve for b_j in 2.16 and get

$$b_j = \frac{1}{h_j}(a_{j+1} - a_j) - \frac{h_j}{3}(2c_j + c_{j+1}). \tag{2.18}$$

Reduce the index by 1, to get

$$b_{j-1} = \frac{1}{h_{j-1}}(a_j - a_{j-1}) - \frac{h_{j-1}}{3}(2c_{j-1} + c_j). \tag{2.19}$$

Finally Plug 2.18 (lhs) and 2.19 (rhs) into the index-reduced-by-1 version of 2.17, i.e.

$$b_j = b_{j-1} + h_{j-1}(c_{j-1} + c_j). \tag{2.20}$$

After some ”massaging” we end up with the linear system of equations for $j \in \{1, 2, \cdots, n-1\}$ (the interior nodes).

$$h_{j-1}c_{j-1} + u_j c_j + h_j c_{j+1} = v_j, i = 2, 3, \cdots, n-1 \tag{2.21}$$

Where

$$u_j = 2(h_{j-1} + h_j),\ v_j = 3w_j - 3w_{j-1} \text{ and } w_j = \frac{1}{h_j}(a_{j+1} - a_j)$$

We are almost ready to solve for the coefficients $\{c_j\}_{j=0}^{n-1}$, but we only have $(n-1)$ equations for $(n+1)$ unknowns. We can complete the system in many ways, some common ones are:

1. **Natural boundary conditions:**

$$\begin{aligned} &[n_1] \quad 0 = S_0''(x_0) = 2c_0 \implies c_0 = 0 \\ &[n_2] \quad 0 = S_n''(x_n) = 2c_n \implies c_n = 0 \end{aligned}$$

From this we produce a linear, tridiagonal system of the form

$$Ax = b$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0+h_1) & h_1 & \ddots & \vdots & \\ 0 & h_1 & 2(h_1+h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ \dfrac{3(a_2-a_1)}{h_1} - \dfrac{3(a_1-a_0)}{h_0} \\ \vdots \\ \dfrac{3(a_n-a_{n-1})}{h_{n-1}} - \dfrac{3(a_{n-1}-a_{n-2})}{h_{n-2}} \\ 0 \end{bmatrix}, \quad x = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix}$$

x are the unknowns (the quantity we are solving for!)

2. **Clamped boundary conditions:** (Derivative known at endpoints).

$$[c_1] \quad S_0'(x_0) = b_0 = f'(x_0)$$
$$[c_2] \quad S_{n-1}'(x_n) = b_n = b_{n-1} + h_{n-1}(c_{n-1} + c_n) = f'(x_n)$$

$[c_1]$ and $[c_2]$ give the additional equations

$$[c_1'] \quad 2h_0c_0 + h_0c_1 = \frac{3}{h_0}(a_1 - a_0) - 3f'(x_0)$$
$$[c_2'] \quad h_{n-1}c_{n-1} + 2h_{n-1}c_n = 3f'(x_n) - \frac{3}{h_{n-1}}3(a_n - a_{n-1}).$$

From this we have a linear system
$$Ax = b$$

where
$$A = \begin{bmatrix} 2h_0 & h_0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0+h_1) & h_1 & \ddots & \vdots & \\ 0 & h_1 & 2(h_1+h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & h_{n-1} & 2h_{n-1} \end{bmatrix}$$

$$b = \begin{bmatrix} \frac{3(a_1-a_0)}{h_0} - 3f'(x_0) \\ \frac{3(a_2-a_1)}{h_1} - \frac{3(a_1-a_0)}{h_0} \\ \vdots \\ \frac{3(a_n-a_{n-1})}{h_{n-1}} - \frac{3(a_{n-1}-a_{n-2})}{h_{n-2}} \\ 3f'(x_n) - \frac{3(a_n-a_{n-1})}{h_{n-1}} \end{bmatrix}, \quad x = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix}$$

Boundary Terms: marked in red-bold.

This system is strictly diagonally dominant, so an elimination method can be used to find the solution without the need for pivoting. Having obtained the values of $\{c_j\}_{i=1}^{n}$, the remainder of the spline coefficients for

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$$

is obtained using the formulas

$$\begin{aligned} a_j &= f(x_j) \\ b_j &= \frac{a_{j+1} - a_j}{h_j} - \frac{h_j(2cj + c_{j+1})}{3} \\ d_j &= \frac{c_{j+1} - c_j}{3h_j} \end{aligned}. \tag{2.22}$$

for $i = 1, 2, \cdots n - 1$.

Example 2.24

Use the values given by $f(x) = x^3+2$ at $x = 0, 0.2, 0.4, 0.6, 0.8,$ and 1.0 to find an approximation of $f(x)$ at $x = 0.1, 0.3, 0.5, 0.7,$ and 0.9 using natural cubic spline interpolation.

We have the table

x	0.0	0.2	0.4	0.6	0.8	1.0
f(x)	2.0	2.008	2.064	2.216	2.512	3.0

$$u_1 = 0.8, u_2 = 0.8, u_3 = 0.8, u_4 = 0.8$$
$$v_1 = 0.72, v_2 = 1.44, v_3 = 2.16, v_4 = 2.88.$$

Using these values, we obtain the linear system of equations

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.8 & 0.2 & 0 & 0 & 0 \\ 0 & 0.2 & 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0.2 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.72 \\ 1.44 \\ 2.16 \\ 2.88 \\ 0 \end{bmatrix}.$$

The solution of this system is

$$c = [0.0, 0.586, 1.257, 1.585, 3.204, 0.0]^T.$$

Using Eqn. (2.22), we obtain the coefficients

$$a = \begin{pmatrix} 2 \\ 2.008 \\ 2.064 \\ 2.216 \\ 2.512 \end{pmatrix}, b = \begin{pmatrix} 0.001 \\ 0.118 \\ 0.487 \\ 1.055 \\ 2.013 \end{pmatrix}, d = \begin{pmatrix} 0.976 \\ 1.120 \\ 0.545 \\ 2.699 \\ -5.340 \end{pmatrix}.$$

Hence,

$$S(x) = \begin{cases} S_1(x), x \in [0.0, 0.2] \\ S_2(x), x \in [0.2, 0.4] \\ S_3(x), x \in [0.4, 0.6] \\ S_4(x), x \in [0.6, 0.8] \\ S_5(x), x \in [0.8, 1.0] \end{cases}$$

With

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3, \; i = 1, \cdots, 5.$$

That is,

$$S(x) = \begin{cases} 2 + 0.001x + 0.976x^3, x \in [0.0, 0.2] \\ 2.008 + 0.118(x - 0.2) + 0.586(x - 0.2)^2 + 1.120(x - 0.2)^3, x \in [0.2, 0.4] \\ 2.064 + 0.487(x - 0.4) + 1.257(x - 0.4)^2 + 0.545(x - 0.4)^3, x \in [0.4, 0.6] \\ 2.216 + 1.055(x - 0.6) + 1.585(x - 0.6)^2 + 2.699(x - 0.6)^3, x \in [0.6, 0.8] \\ 2.512 + 2.013(x - 0.8) + 3.204(x - 0.8)^2 - 5.340(x - 0.8)^3, x \in [0.8, 1.0] \end{cases}$$

For example, the value $x = 0.5$ lies in the interval $[0.4, 0.6]$, so $S_3(0.5) = 2.064+0.487(0.5-0.4) + 1.257(0.5 - 0.4)^2 + 0.545(0.5 - 0.4)^3 \approx$ 2.126.

Example 2.25

Construct the natural cubic spline interpolant for $f(x) = \ln(e^x + 2)$ with nodal values:

x	f (x)
-1.0	0.86199480
-0.5	0.95802009
0.0	1.0986123
0.5	1.2943767

Calculate the absolute error in using the interpolant to approximate $f(0.25)$ and $f'(0.25)$. In this case $n = 3$ and $h_0 = h_1 = h_2 = 0.5$ with

$$a_0 = 0.86199480, a_1 = 0.95802009,$$
$$a_2 = 1.0986123, a_3 = 1.2943767.$$

The linear system resembles,

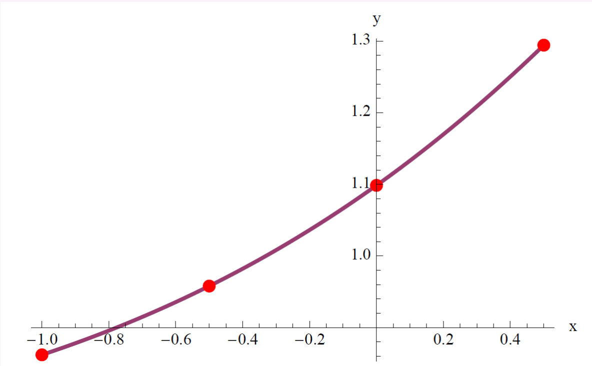
$$Ax = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 2.0 & 0.5 & 0.0 \\ 0.0 & 0.5 & 2.0 & 0.5 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0.0 \\ 0.267402 \\ 0.331034 \\ 0.0 \end{bmatrix} = b$$

The coefficients of the piecewise cubics:

j	a_j	b_j	c_j	d_j
0	0.861995	0.175638	0.0	0.0656509
1	0.95802	0.224876	0.0984763	0.028281
2	1.09861	0.344563	0.140898	-0.093918

The cubic spline:

$$S(x) = \begin{cases} 0.861995 + 0.175638(x+1) + 0.0656509(x+1)^3 & -1 \le x \le -0.5 \\ 0.95802 + 0.224876(x+0.5) + 0.0984763(x+0.5)^2, & -0.5 \le x \le 0 \\ +0.028281(x+0.5)^3 & \\ 1.09861 + 0.344563x + 0.140898x^2 - 0.093918x^3, & 0 \le x \le 0.5 \end{cases}$$



f(0.25)	S(0.25)	Abs. Err.	$f'(0.25)$	$S'(0.25)$	Abs. Err.
1.18907	1.19209	3.02154^{-3}	0.390991	0.3974	6.40839×10^{-3}

Example 2.26

Construct the clamped cubic spline interpolant for $f(x) = \ln(e^x + 2)$ *with nodal values:*

x	f (x)
-1.0	0.86199480
-0.5	0.95802009
0.0	1.0986123
0.5	1.2943767

Calculate the absolute error in using the interpolant to approximate $f(0.25)$ *and* $f'(0.25)$*.*
In this case $n = 3$ and $h_0 = h_1 = h_2 = 0.5$ *with*

$$a_0 = 0.86199480, a_1 = 0.95802009,$$
$$a_2 = 1.0986123, a_3 = 1.2943767.$$

Note that $f'(-1) \approx 0.155362$ *and* $f'(0.5) \approx 0.451863$ *The linear system resembles,*

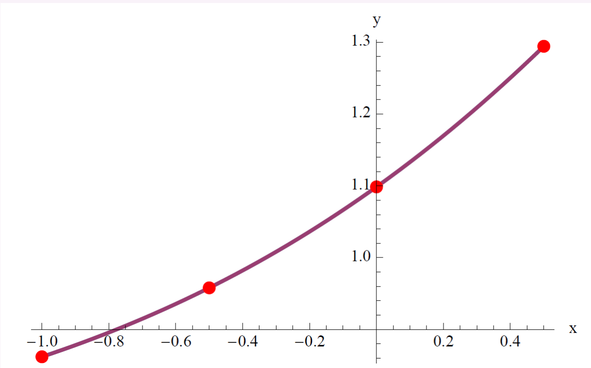
$$Ax = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 2.0 & 0.5 & 0.0 \\ 0.0 & 0.5 & 2.0 & 0.5 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0.110064 \\ 0.267402 \\ 0.331034 \\ 0.181001 \end{bmatrix} = b$$

The coefficients of the piecewise cubics:

j	a_j	b_j	c_j	d_j
0	0.861995	0.1553628	0.0653748	0.0160031
1	0.95802	0.23274	0.0893795	0.0150207
2	1.09861	0.333384	0.11191	0.00875717

The cubic spline:

$$S(x) = \begin{cases} 0.861995 + 0.1553628(x+1) + 0.0653748(x+1)^2 \\ +0.0160031(x+1)^3, \ -1 \le x \le -0.5 \\ 0.95802 + 0.23274(x+0.5) + 0.0893795(x+0.5)^2 \\ +0.0150207(x+0.5)^3, \ -0.5 \le x \le 0 \\ 1.09861 + 0.333384x + 0.11191x^2 + 0.00875717x^3, \ 0 \le x \le 0.5 \end{cases}$$



f(0.25)	S(0.25)	Abs. Err.	$f'(0.25)$	$S'(0.25)$	Abs. Err.
1.18907	1.18991	1.97037×10^{-5}	0.390991	0.390982	9.67677×10^{-6}

Chapter 3

Least Square Method

Many scientific and engineering observations are made by conducting experiments in which physical quantities are measured and recorded. The experimental records are typically referred to as data points.

An available set of data can be used for developing, or evaluating, mathematical formulas (equations) that represent the data. In some cases, the data is represented by a function, which in turn can be used for numerical differentiation or integration. Such function may be obtained through curve fitting, or approximation, of the data. **Curve fitting** is a procedure where a function is used to it a given set of data in the ”best” possible manner without having to match the data exactly. As a result, while the function does not necessarily yield the exact value at the data points, overall it its the set of data well. Curve fitting is normally used when the data has substantial inherent error, such as data gathered from experimental measurements.

The goal of Least-Squares Method is to find a good estimation of parameters that fit a function, $f(x)$, of a set of data, $x_1, x_2, \cdots x_n$. The Least-Squares Method requires that the estimated function has to deviate as little as possible from $f(x)$ in the sense of a 2-norm. Generally speaking, Least-Squares Method has two categories, linear and non-linear.

3.1 Least-Squares Regression

Sometimes we get a lot of data, many observations, and want to fit it to a simple model. Low dimensional models (e.g. low degree polynomials) are easy to work with, and are quite well behaved (high degree polynomials can be quite oscillatory.) However, when the data has substantial error, even if the size of data is small, this may no longer be appropriate. Consider Figure 3.1, which shows a set of seven data points collected from an experiment. The nature of the data suggests that, for the most part, the y values increase with the x values. A single interpolating polynomial goes through all of the data points, but displays large oscillations in some regions. As a result, the interpolated values near $x = 1.2$ and $x = 2.85$, for instance, will be well outside of the range of the original data. In these types of situations, it makes more sense to ind a function that does not necessarily go through all of the data points, but its the data well overall. One option, for example, is to it the ”best” straight line into the data. This line is not random and can be generated systematically via least-squares regression.

All measurements are noisy, to some degree. Often, we want to use a large number of measurements in order to ”average out” random noise.

Approximation Theory looks at two problems:

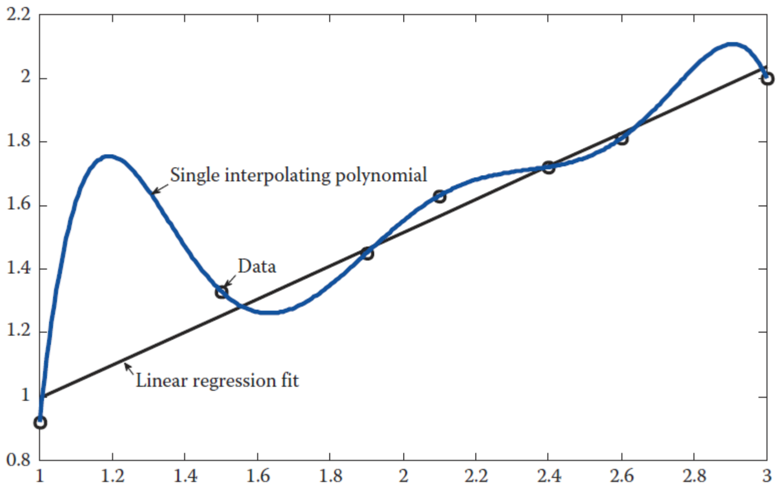


Figure 3.1: Interpolation by a single polynomial, and linear regression it of a set of data.

1 Given a data set, find the best fit for a model (i.e. in a class of functions, find the one that best represents the data.)

2 Find a simpler model approximating a given function.

3.1.1 Criteria for a "Best" Fit

A criterion that measures how good a fit is between given data points and an approximating linear function is a formula that calculates a number that quantifies the overall agreement between the points and the function. Such a criterion is needed for two reasons. First, it can be used to compare two different functions that are used for fitting the same data points. Second, and even more important, the criterion itself is used for determining the coefficients of the function that give the best fit.

We are going to relax the requirement that the approximating function must pass through all the data points. Now we need a measurement of how well our approximation fits the data.

If $f(x_i)$ are the measured function values, and $a(x_i)$ are the values of our approximating functions, we can define a function, $r(x_i) = f(x_i) - a(x_i)$ which measures the deviation (residual) at x_i. Notice that $\tilde{r} = \{r(x_0), r(x_1), \cdots, r(x_n)\}^T$ is a vector.
Notation: From now on, $f_i = f(x_i)$, $a_i = a(x_i)$, and $r_i = r(x_i)$. Further, $\tilde{f} = \{f_0, f_1, \cdots, f_n\}^T, \tilde{a} = \{a_0, a_1, \cdots, a_n\}^T$, and $\tilde{r} = \{r_0, r_1, \cdots, r_n\}^T$.

Different strategies can be considered for determining the best fit of a set of n data points $(x_1, y_1), \cdots, (x_n, y_n)$.

1. One strategy is to minimize the sum of all the individual errors,

$$E = \sum_{i=0}^{n} r_i = \sum_{i=0}^{n} [f_i - a_i]$$

This criterion, however, does not offer a good measure of how well the line its the data because, as shown in Figure3.2, it allows for positive and negative individual errors even very large errors to cancel out and yield a zero sum.

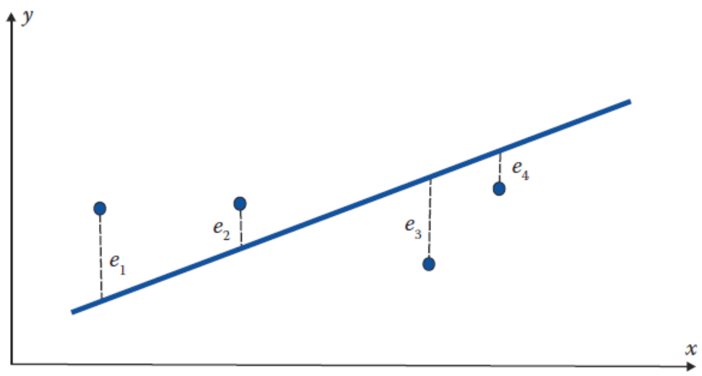


Figure 3.2: Zero total error based on the criterion

2. To minimize the sum of the absolute values of the individual errors,

$$E_1 = \sum_{i=0}^{n} |r_i| = \sum_{i=0}^{n} |f_i - a_i| \Leftrightarrow \|\tilde{r_i}\|_1$$

As a result, the individual errors can no longer cancel out and the total error is always positive. This criterion, however, is not able to uniquely determine the coefficients that describe the best line it because for a given set of data, several lines can have the same total error. Figure 3.3 shows a set of four data points with two line its that have the same total error.

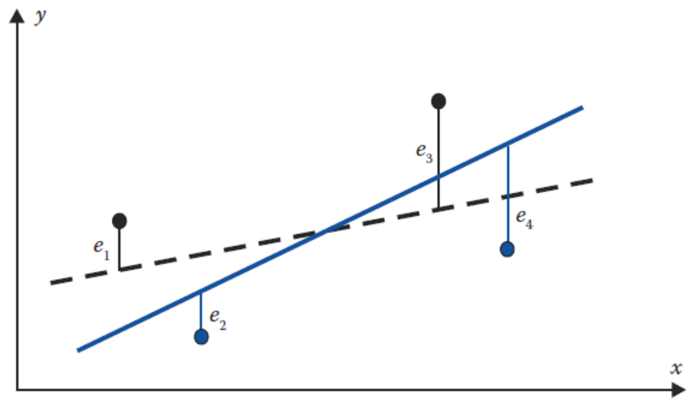


Figure 3.3: Two linear its with the same total error calculated

3. To minimize the sum of the squares of the individual errors,

$$E_1 = \sum_{i=0}^{n} r_i^2 = \sum_{i=0}^{n} (f_i, -a_i)^2 \Leftrightarrow \|\tilde{r_i}\|_1$$

This criterion uniquely determines the coefficients that describe the best line it for a given set of data. As in the second strategy, individual errors cannot cancel each other out and the total error is always positive. Also note that small errors get smaller and large errors get larger. This means that larger individual errors have larger contributions to the total error being minimized so that this strategy essentially minimizes the maximum distance that an individual data point is located relative to the line.

3.1.2 Discrete Least Squares Approximation

We have chosen the sum-of-squares measurement for errors. Lets find the constant that best fits the data, minimize

$$E(C) = \sum_{i=0}^{n} (f_i - C)^2.$$

If C^* is a minimizer, then $E'(C^*) = 0$ [derivative at a max/min is zero]

$$E'(C) = -\sum_{i=0}^{n} 2(f_i - C) = \underbrace{-2\sum_{i=0}^{n} fi + 2(n+1)C,}_{\text{Set =0, and solve for C}} \quad E''(C) = \underbrace{2(n+1)}_{\text{Positive}}$$

hence $C^* = \frac{1}{n+1}\sum_{i=0}^{n} f_i$ it is a min since $E''(C^*) = 2(n+1) > 0.$ is the constant that best the fits the data. (Note: C^* is the average.)

3.2 Linear Least Square

As decided above, the criterion to find the line $y = a_1x + a_0$ that best fits the data is to determine the coefficients a_1 and a_0. The error $E(a_0, a_1)$ we need to minimize is:

$$E(a_0, a_1) = \sum_{i=0}^{n} [(a_1x_i + a_0) - y_i]^2 \tag{3.1}$$

The first partial derivatives with respect to a_0 and a_1 better be zero at the minimum:

$$\frac{\partial}{\partial a_0} E(a_0, a_1) = 2\sum_{i=0}^{n} [(a_1x_i + a_0) - y_i] = 0$$
$$\frac{\partial}{\partial a_1} E(a_0, a_1) = 2\sum_{i=0}^{n} x_i[(a_1x_i + a_0) - y_i] = 0$$

Expanding and rearranging the above equations to get the Normal Equations

$$\begin{cases} a_0(n+1) + a_1\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} y_i \\ a_0\sum_{i=0}^{n} x_i + a_1\sum_{i=0}^{n} x_i^2 = \sum_{i=0}^{n} x_iy_i \end{cases}$$

Since everything except a_0 and a_1 is known, this is a 2-by-2 system of equations.

$$\begin{bmatrix} (n+1) & \sum_{i=0}^{n} x_i \\ \sum_{i=0}^{n} x_i & \sum_{i=0}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{n} y_i \\ \sum_{i=0}^{n} x_iy_i \end{bmatrix}.$$

By Cramer's rule, the solutions are found as

$$a_1 = \frac{(n+1)\left(\sum_{i=0}^{n} x_iy_i\right) - \left(\sum_{i=0}^{n} x_i\right)\left(\sum_{i=0}^{n} y_i\right)}{(n+1)\left(\sum_{i=0}^{n} x_i^2\right) - \left(\sum_{i=0}^{n} x_i\right)^2}$$

$$a_0 = \frac{\left(\sum_{i=0}^{n} x_i^2\right)\left(\sum_{i=0}^{n} y_i\right) - \left(\sum_{i=0}^{n} x_i y_i\right)\left(\sum_{i=0}^{n} x_i\right)}{(n+1)\left(\sum_{i=0}^{n} x_i^2\right) - \left(\sum_{i=0}^{n} x_i\right)^2}$$

Example 3.1

Find the best-fit values of a and b so that $y = a + bx$ fits the data given in the table.

x	0	1	2	3	4
y	1	1.8	3.3	4.5	6.3

Let the straight line is $y = a + bx$

x_i	y_i	$x_i y_i$	x_i^2
0	1	0	0
1	1.8	1.8	1
2	3.3	6.6	4
3	4.5	13.5	9
4	6.3	25.2	16
$\sum x_i = 10$	$\sum y_i = 16.9$	$\sum x_i y_i = 47.1$	$\sum x_i^2 = 30$

Putting these values in normal equations we get,

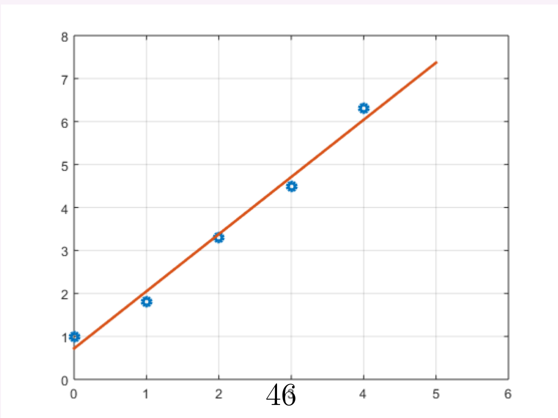
$$16.9 = 5a + 10b$$
$$47.1 = 10a + 30b$$

On solving these two equations we get, $a = 0.72, b = 1.33$. So required line $y = 0.72 + 1.33x$. Now, we can evaluate the sum of squares of errors as:

y_i	1	1.8	3.3	4.5	6.3
$p_1(x_i) = 0.72 + 1.33x_i$	0.72	2.05	3.38	4.71	6.04
$e_i = y_i - p_1(x_i)$	0.28	-0.25	-0.08	-0.21	0.26

$$\sum_{i=0}^{4} (e_i)^2 = (0.28)^2 + (-0.25)^2 + (-0.08)^2 + (-0.21)^2 + (0.26)^2 = 0.2590$$

```
x=0:4;
y=[1 1.8 3.3 4.5 6.3];
 a =     1.330000
 b =     0.720000

----------------------------------------------
 x        y          a*x+b          |y-(ax+b)|

----------------------------------------------
  0.00    1.00       0.720000       0.280000
  1.00    1.80       2.050000       0.250000
  2.00    3.30       3.380000       0.080000
  3.00    4.50       4.710000       0.210000
  4.00    6.30       6.040000       0.260000
 E(a,b) =     0.259000
```



Example 3.2

Using the method of least-squares, find the linear function that best fits the following data

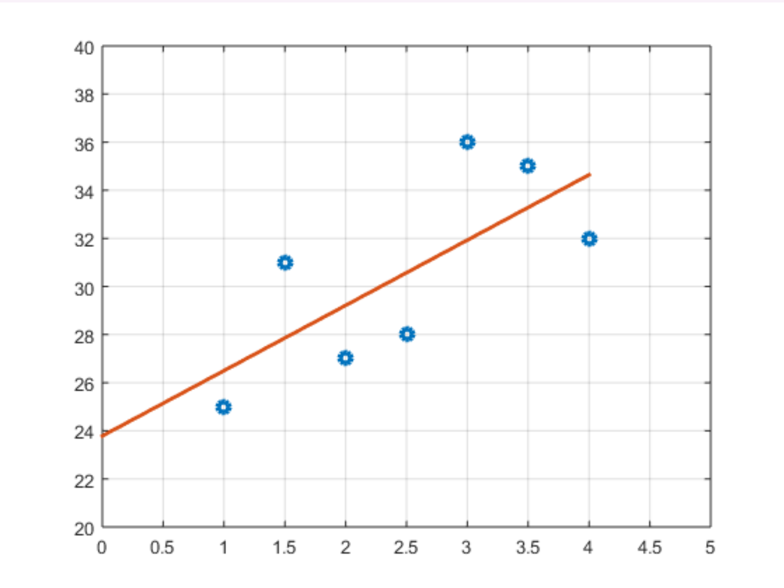
x	1	1.5	2	2.5	3	3.5	4
y	25	31	27	28	36	35	32

```
x=[1:0.5:4];
y=[25 31 27 28 36 35 32];
linlsqr(x,y)
```

```
 linear least squares

 a =     2.714286
 b =    23.785714

----------------------------------------------
 x       y          a*x+b         |y-(ax+b)|

----------------------------------------------
   1.00   25.00     26.500000      1.500000
   1.50   31.00     27.857143      3.142857
   2.00   27.00     29.214286      2.214286
   2.50   28.00     30.571429      2.571429
   3.00   36.00     31.928571      4.071429
   3.50   35.00     33.285714      1.714286
   4.00   32.00     34.642857      2.642857
 E(a,b) =    50.142857
```

Therefore, the least-squares line is

$$y = 2.71428571x + 23.78571429.$$



3.3 Non-Linear Least Square

3.3.1 Quadratic Model

For the quadratic polynomial $p_2(x) = a_0 + a_1x + a_2x_2$, the error is given by

$$E(a_0, a_1, a_2) = \sum_{i=0}^{n} \left[a_2x_i^2 + a_1x_i + a_0 - y_i\right]^2$$

At the minimum (best model) we must have

$$\begin{aligned}
\frac{\partial}{\partial a_0}E(a_0, a_1, a_2) &= 2\sum_{i=0}^{n}[(a_2x_i^2 + a_1x_i + a_0) - y_i] = 0 \\
\frac{\partial}{\partial a_1}E(a_0, a_1, a_2) &= 2\sum_{i=0}^{n}x_i[(a_2x_i^2 + a_1x_i + a_0) - y_i] = 0 \\
\frac{\partial}{\partial a_2}E(a_0, a_1, a_2) &= 2\sum_{i=0}^{n}x_i^2[(a_2x_i^2 + a_1x_i + a_0) - y_i] = 0
\end{aligned}$$

Similarly for the quadratic polynomial $p_2(x) = a_0 + a_1x + a_2x^2$, the normal equations are:

$$\begin{cases}
a_0(n+1) + a_1\sum_{i=0}^{n}x_i + a_2\sum_{i=0}^{n}x_i^2 = \sum_{i=0}^{n}y_i \\
a_0\sum_{i=0}^{n}x_i + a_1\sum_{i=0}^{n}x_i^2 + a_2\sum_{i=0}^{n}x_i^3 = \sum_{i=0}^{n}x_iy_i \\
a_0\sum_{i=0}^{n}x_i^2 + a_1\sum_{i=0}^{n}x_i^3 + a_2\sum_{i=0}^{n}x_i^4 = \sum_{i=0}^{n}x_i^2y_i
\end{cases}$$

Note: Even though the model is quadratic, the resulting (normal) equations are linear. The model is linear in its parameters, a_0, a_1, and a_2.
We rewrite the Normal Equations as:

$$\begin{bmatrix}
(n+1) & \sum_{i=0}^{n}x_i & \sum_{i=0}^{n}x_i^2 \\
\sum_{i=0}^{n}x_i & \sum_{i=0}^{n}x_i^2 & \sum_{i=0}^{n}x_i^3 \\
\sum_{i=0}^{n}x_i^2 & \sum_{i=0}^{n}x_i^3 & \sum_{i=0}^{n}x_i^4
\end{bmatrix}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}
=
\begin{bmatrix}
\sum_{i=0}^{n}y_i \\
\sum_{i=0}^{n}x_iy_i \\
\sum_{i=0}^{n}x_i^2y_i
\end{bmatrix}.$$

Example 3.3

Find the least square polynomial approximation of degree two to the data.

x	0	1	2	3	4
y	-4	-1	4	11	20

also compute the least error.

Solution: *Let the equation of the polynomial be* $y = a + bx + cx^2$

x_i	y_i	$x_i y_i$	x_i^2	$y_i x_i^2$	x_i^3	x_i^4
0	- 4	0	0	0	0	0
1	- 1	- 1	1	-1	1	1
2	4	8	4	16	8	16
3	11	33	9	99	27	81
4	20	80	16	320	64	256
$\sum x_i = 10$	$\sum y_i = 30$	$\sum x_i y_i = 120$	$\sum x_i^2 = 30$	$\sum y_i x_i^2 = 434$	$\sum x_i^3 = 100$	$\sum x_i^4 = 354$

From the normal equation we have

$$\begin{aligned} 30 &= 5a + 10b + 30c \\ 120 &= 10a + 30b + 100c \\ 434 &= 30a + 100b + 354c \end{aligned}$$

On solving these equations, we get $a = -4, b = 2, c = 1$. *Therefore required polynomial is* $y = -4 + 2x + x^2$ *, errors = 0 since* $p_2(x_i) = y_i$

3.3.2 Least-Squares Polynomial

The method of least-squares data fitting is not restricted to linear functions $f(x) = ax+b$ only. As a matter of fact, in many cases data from experimental results are not linear, so we need to consider some other guess functions.

$$p_m(x) = \sum_{k=0}^{m} a_k x_k \tag{3.2}$$

of degree $m \leq n-1$. So, according to the least-squares principle, we need to find the coefficients $a_0, a_1, \cdots, a_m$ that minimize

$$\begin{aligned} E(a_0, \cdots, a_m) &= \sum_{i=1}^{n} [p_m(x_i) - y_i]^2 \\ &= \sum_{i=1}^{n} \left[\sum_{k=0}^{m} (a_k x_i^k) - y_i \right]^2 \end{aligned} \tag{3.3}$$

As before,E is minimum if

$$\frac{\partial}{\partial a_j} E(a_0, \cdots, a_m) = 0, j = 0, 1, \cdots, m. \tag{3.4}$$

that is

$$\begin{aligned}
\frac{\partial E}{\partial a_0} &= \sum_{i=1}^{n} 2\left[\sum_{k=0}^{m}(a_k x_i^k) - y_i\right] = 0,\\
\frac{\partial E}{\partial a_1} &= \sum_{i=1}^{n} 2\left[\sum_{k=0}^{m}(a_k x_i^k) - y_i\right](x_i) = 0,\\
&\vdots\\
\frac{\partial E}{\partial a_m} &= \sum_{i=1}^{n} 2\left[\sum_{k=0}^{m}(a_k x_i^k) - y_i\right](x_i^m) = 0.
\end{aligned} \tag{3.5}$$

Rearranging Eqn. (3.5) gives the $(m+1)$ normal equations for the $(m+1)$ unknowns $a_0, a_1, \cdots, a_m$

$$\begin{aligned}
a_0 n + a_1 \sum x_i + \cdots + a_m \sum x_i^m &= \sum y_i\\
a_0 \sum x_i + a_1 \sum x_i^2 + \cdots + a_m \sum x_i^{m+1} &= \sum y_i x_i\\
a_0 \sum x_i^2 + a_1 \sum x_i^3 + \cdots + a_m \sum x_i^{m+2} &= \sum y_i x_i^2\\
&\vdots\\
a_0 \sum x_i^m + a_1 \sum x_i^{m+1} + \cdots + a_m \sum x_i^{2m} &= \sum y_i x_i^m
\end{aligned} \tag{3.6}$$

As before,$\sum$ denotes $\sum_{i=1}^{n}$.

3.3.3 Linearization of Nonlinear Data

If the relationship between the independent and dependent variables is not linear, curve-fitting techniques other than linear regression must be used.

Exponential Function

The exponential function is in the form

$$y = ae^{bx} \quad (a, b \ \text{ const}$$

Because differentiation of the exponential function returns a constant multiple of the exponential function, this technique applies to situations where the rate of change of a quantity is directly proportional to the quantity itself; for instance, radioactive decay. Conversion into linear form is made by taking the natural logarithm of Equation above to obtain

$$\ln y = bx + \ln a$$

Therefore, the plot of $\ln y$ versus x is a straight line with slope b and intercept $\ln a$; see Figure 3.4 a and d.

Power Function

Another example of a nonlinear function is the power function

$$y = ax^b \quad (a, b \ \text{ const}$$

Linearization is achieved by taking the standard (base 10) logarithm ,

$$\log y = b \log x + \log a$$

so that the plot of $\log y$ versus $\log x$ is a straight line with slope b and intercept $\log a$; see Figure 3.4 b and e.

Example 3.4

Fit a curve $y = ab^x$ to the following data:

x	2	3	4	5	6
y	144	172.8	207.4	248.8	298.5

Given equation $y = ab^x$ reduced to $\ln y = \ln a + x \ln b$ The normal equations are:

$$(\sum_{i=0}^{n} 1) \ln a + (\sum_{i=0}^{n} x_i) \ln b = \sum_{i=0}^{n} \ln y_i$$
$$(\sum_{i=0}^{n} x_i) \ln a + (\sum_{i=0}^{n} x_i^2) \ln b = \sum_{i=0}^{n} x_i \ln y_i$$

i	x_i	y_i	x_i^2	$\ln y_i$	$x_i \ln y_i$
0	2	144	4	4.9698	9.9396
1	3	172.8	9	5.1521	15.4564
2	4	207.4	16	5.3346	21.3386
3	5	248.8	25	5.5166	27.5832
4	6	298.5	36	5.6988	34.1926
$\sum$	20		90	26.6719	108.5104

$$5 \ln a + 20 \ln b = 26.6719$$
$$20 \ln a + 90 \ln b = 108.5104$$

On solving the system of equation, we get $\ln a = 4.6053$, & $\ln b = 0.1823 \implies e^{\ln a} = a = e^{4.6053} = 100.0130 \& e^{lnb} = b = e^{0.1823} = 1.2$. *Thus, the curve is* $y = 100.0130 \times 1.2^x$

Example 3.5

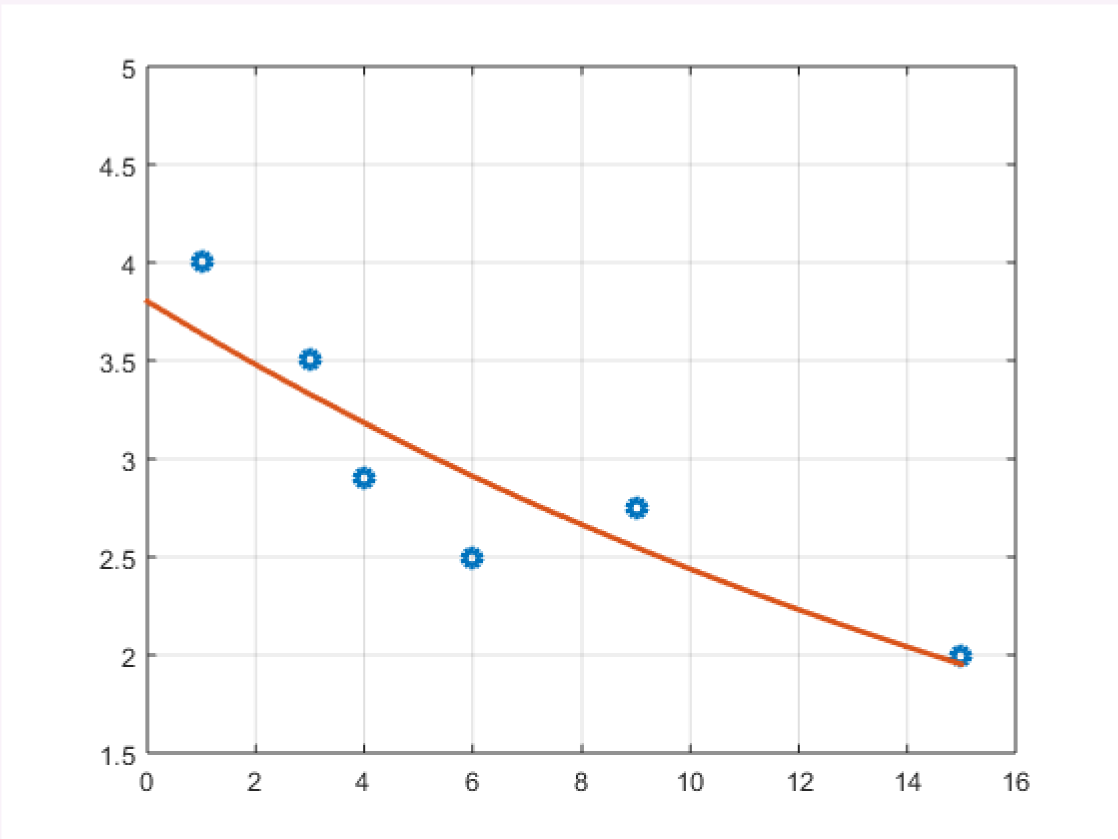
Find the least-squares exponential that best fits the following data:

x	1	3	4	6	9	15
y	4.0	3.5	2.9	2.5	2.75	2.0

```
x=[1 3 4 6 9 15];
 y=[4 3.5 2.9 2.5 2.75 2];


 Exponential least squares

 a =     3.801404
 b =   -0.044406
-----------------------------------------------
    xi    yi     ln(yi)    a*exp(b*xi)  |yi-[a*exp(b*xi)]|

-----------------------------------------------
  1.00    4.00    1.39      3.636293      0.363707
  3.00    3.50    1.25      3.327274      0.172726
  4.00    2.90    1.06      3.182757      0.282757
  6.00    2.50    0.92      2.912280      0.412280
  9.00    2.75    1.01      2.549046      0.200954
 15.00    2.00    0.69      1.952841      0.047159
```

The equation of the least-squares exponential is

$$y = 3.801404e^{-0.044406x}$$



Saturation Function

The saturation function is in the form

$$y = \frac{x}{ax+b} \quad (a, b \ \text{ const}$$

Inverting the Equation yields

$$\frac{1}{y} = b\left(\frac{1}{x}\right) + a$$

so that the plot of $1/y$ versus $1/x$ is a straight line with slope b and intercept a; see Figure 3.4 c and f.

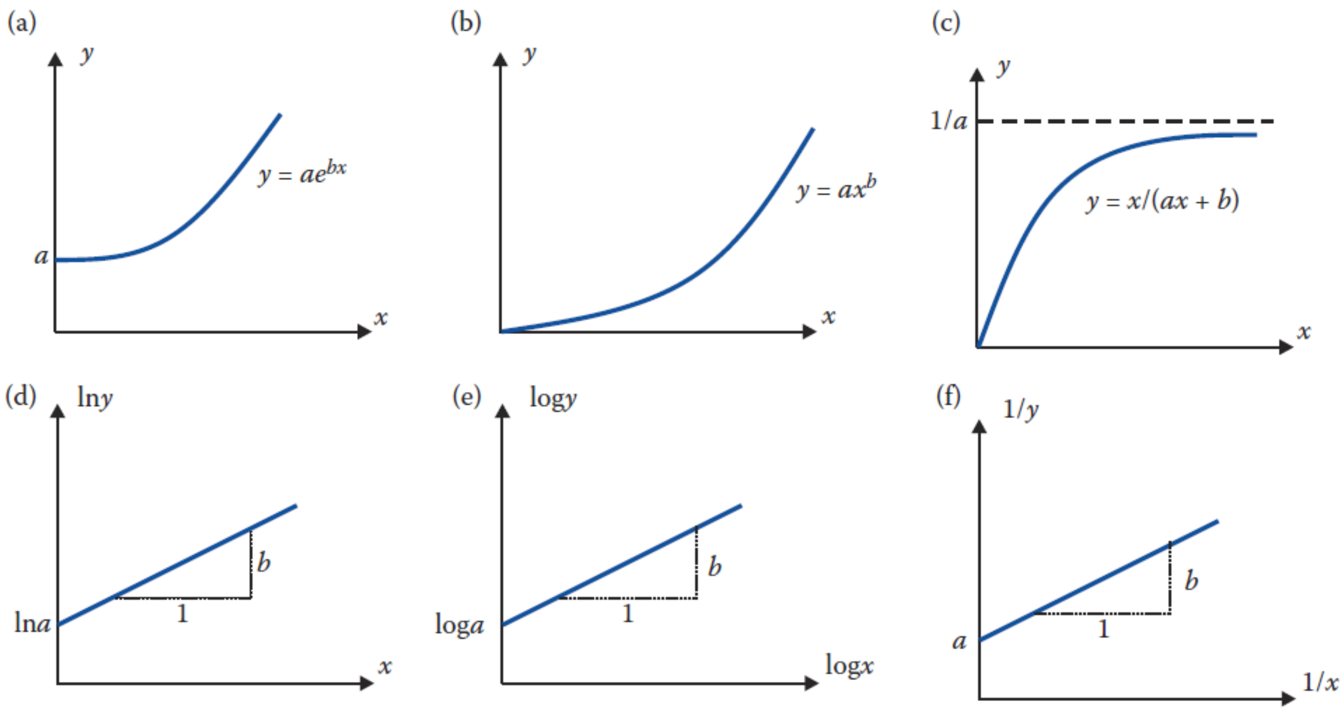


Figure 3.4: Linearization of three nonlinear functions for curve itting. (a, d) Exponential function, (b, e) Power function, (c, f) Saturation function.

Example 3.6

Find the least-squares saturation function that best fits the data in above example

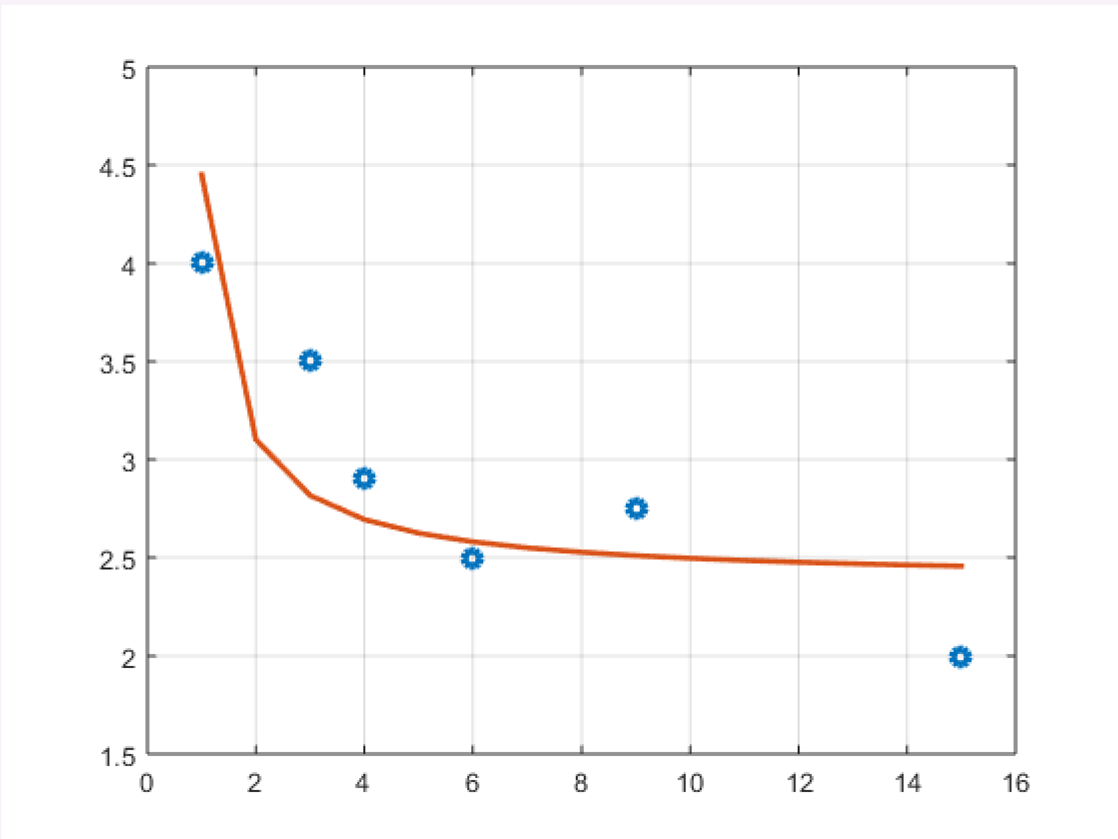
$1/x$	1.0000	0.3333	0.2500	0.1667	0.1111	0.0667
$1/y$	0.2500	0.2857	0.3448	0.4000	0.3636	0.5000

```
x=[1 3 4 6 9 15];
 y=[4 3.5 2.9 2.5 2.75 2];
```

```
 hyperbolic least squares

 a =    0.420179
 b =   -0.195506

-----------------------------------------------------
    x      y     1/xi  x(i)/(a*x(i)+b) |y-x(i)/(a*x(i)+b)|

-----------------------------------------------------
  1.00    4.00   1.00     4.450930      0.450930
  3.00    3.50   0.33     2.816824      0.683176
  4.00    2.90   0.25     2.693226      0.206774
  6.00    2.50   0.17     2.580019      0.080019
  9.00    2.75   0.11     2.509690      0.240310
 15.00    2.00   0.07     2.456129      0.456129
```

The equation of the least-squares hyperbolic is

$$y = x/(0.420179x - 0.195506)$$



3.4 Continuous Least-Squares Approximation

In the previous section, we have described least-squares approximation to fit a set of **discrete data**. Here we first describe **continuous least-square approximations** of a function $f(x)$ by using polynomials and later in the subsequent sections using orthogonal polynomials and Fourier series.

3.4.1 Approximation by Polynomials

First, consider approximation by a polynomial with **monomial basis**: $\{1, x, x^2, \cdots, x^n\}$.

Least-Square Approximations of a Function Using Monomial Polynomials
Given a function $f(x)$, continuous on $[a, b]$, find a polynomial $P_n(x)$ of degree at most n:

$$P_n(x) = a_0 + a_1x + a_2x^2 + \cdots a_nx^n \quad (n \leq m),$$

such that the integral of the square of the error is minimized. That is,

$$E(a_0, a_1, a_2, \cdots, a_n) = \int_a^b [f(x) - P_n(x)]^2 \, dx,$$

is minimized.

The polynomial $P_n(x)$ is called the **Least-Squares Polynomial.** For minimization, we must have

$$\frac{\partial E}{\partial a_i} = 0, \quad i = 0, 1, \cdots, n.$$

As before, these conditions will give rise to a system of $(n+1)$ **normal equations** in $(n+1)$ unknowns: $a_0, a_1, \cdots, a_n$. Solution of these equations will yield the unknowns: $a_0, a_1, \cdots, a_n$.

Setting up the Normal Equations

Since

$$E = \int_a^b \left[f(x) - \left(a_0 + a_1x + a_2x^2 + \cdots a_nx^n\right)\right]^2 dx,$$

differentiating E with respect to each a_i results in

$$\begin{aligned}
\frac{\partial E}{\partial a_0} &= -2\int_a^b \left[f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n\right] dx, \\
\frac{\partial E}{\partial a_1} &= -2\int_a^b x\left[f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n\right] dx, \\
&\vdots \\
\frac{\partial E}{\partial a_n} &= -2\int_a^b x^n\left[f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n\right] dx.
\end{aligned}$$

Thus, we have

$$\frac{\partial E}{\partial a_0} = 0 \quad \Longrightarrow \quad a_0\int_a^b 1dx + a_1\int_a^b xdx + a_2\int_a^b x^2dx + \cdots + a_n\int_a^b x^ndx = \int_a^b f(x).$$

Similarly,

$$\frac{\partial E}{\partial a_i} = 0 \quad \Longrightarrow \quad a_0\int_a^b x^idx + a_1\int_a^b dx^{i+1}xdx + a_2\int_a^b x^{i+2}dx + \cdots + a_n\int_a^b x^{i+n}dx = \int_a^b x^if(x).$$

$$i = 0, 1, 2, \cdots n.$$

So, the $(n+1)$ normal equations in this case are:

$$
\begin{aligned}
i=0: &\quad a_0\int_a^b 1dx + a_1\int_a^b xdx + a_2\int_a^b x^2dx + \cdots + a_n\int_a^b x^ndx = \int_a^b f(x). \\
i=1: &\quad a_0\int_a^b xdx + a_1\int_a^b x^2xdx + a_2\int_a^b x^3dx + \cdots + a_n\int_a^b x^{n+1}dx = \int_a^b xf(x). \\
&\quad \vdots \\
i=n: &\quad a_0\int_a^b x^ndx + a_1\int_a^b x^{n+1}xdx + a_2\int_a^b x^{n+2}dx + \cdots + a_n\int_a^b x^{2n}dx = \int_a^b x^nf(x).
\end{aligned}
$$

Denoting

$$\int_a^b x^idx = s_i, \quad i=0,1,2,\cdots,2n, \text{ and } b_i = \int_a^b x^if(x)dx, \quad i=0,1,2,\cdots n,$$

the above $(n+1)$ equations can be written as

$$
\begin{aligned}
&s_0a_0 + s_1a_1 + \cdots + s_na_n = b_0 \\
&s_1a_0 + s_2a_1 + \cdots + s_{n+1}a_n = b_0 \\
&\quad \vdots \\
&s_na_0 + s_{n+1}a_1 + \cdots + s_{2n}a_n = b_0
\end{aligned}
$$

or in matrix notation

$$\begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & & \ddots & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Hence, we have the system of normal equations

$$\mathbf{Sa} = \mathbf{b}, \tag{3.7}$$

where

$$\mathbf{S} = \begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & & \ddots & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

A Special Case:

Let the interval be $[0,1]$. Then

$$s_i = \int_0^1 x^idx = \frac{1}{i+1}, \quad i=0,1,2,\cdots,2n.$$

Thus, in this case the matrix of the normal equations

$$\mathbf{S} = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n} \end{bmatrix},$$

which is a **Hilbert Matrix**. It is well-known to be **ill-conditioned.**

Example 3.7

Find Linear and Quadratic least-squares approximations to $f(x) = e^x$ on $[-1, 1]$. ***Solution:***
Linear approximation: $n = 1$ $P_1(x) = a_0 + a_1 x$ ***Step 1:***

$$
\begin{aligned}
s_0 &= \int_{-1}^{1} 1dx = 2,\\
s_1 &= \int_{-1}^{1} xdx = \left[\frac{x^2}{2}\right]_{-1}^{1} = \frac{1}{2} - \left(\frac{1}{2}\right) = 0,\\
s_2 &= \int_{-1}^{1} x^2dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3} - \left(\frac{-1}{3}\right) = \frac{2}{3}.
\end{aligned}
$$

Step 2:

$$
\begin{aligned}
b_0 &= \int_{-1}^{1} 1de^x = [e^x]_{-1}^{1} = e - \frac{1}{e} = 2.3504,\\
b_1 &= \int_{-1}^{1} xe^xdx = \frac{2}{e} = 0.7358,
\end{aligned}
$$

Step 3: *From the matrix* $\mathbf{S}$ *and vector* $\mathbf{b}$*:*

$$
\mathbf{S} = \begin{bmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2.3504 \\ 0.7358 \end{bmatrix}
$$

Step 4: *Solve the normal system is:*

$$
\begin{bmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 2.3504 \\ 0.7358 \end{bmatrix}
$$

This gives

$$
a_0 = 1.1752, \quad a_1 = 1.1037
$$

The linear least-squares polynomial $P_1(x) = 1.1752 + 1.1037x$.
Accuracy Check:

$$
P_1(0.5) = 1.7270, \quad e^{0.5} = 1.6487.
$$

Relative Error:

$$
\frac{|e^{0.5} - P_1(0.5)|}{|e^{0.5}|} = \frac{|1.6487 - 1.7270|}{|1.6487|} = 0.0475.
$$

Quadratic fitting *n=2;* $P_2(x) = a_0 + a_x + a_2x^2$
Step 1: *Compute* s_i*'s*

$$
\begin{aligned}
s_0 &= 2, \quad s_1 = 0, \quad s_2 = \frac{2}{3}\\
s_3 &= \int_{-1}^{1} x^3dx = \left[\frac{x^4}{4}\right]_{-1}^{1} = \frac{1}{4} - \left(\frac{1}{4}\right) = 0,\\
s_4 &= \int_{-1}^{1} x^4dx = \left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{1}{5} - \left(\frac{-1}{5}\right) = \frac{2}{5}.
\end{aligned}
$$

Example

Step 2: *Compute b_i's*

$$b_0 = 2.3504, \quad b_1 = 0.7358,$$
$$b_2 = \int_{-1}^{1} x^2 e^x dx = e - \frac{5}{e} = 0.8789.$$

Step 3: *From the matrix* $\mathbf{S}$ *and vector* $\mathbf{b}$*:*

$$\mathbf{S} = \begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2.3504 \\ 0.7358 \\ 0.8789 \end{bmatrix}$$

Step 4: *Solve the normal system is:*

$$\begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2.3504 \\ 0.7358 \\ 0.8789 \end{bmatrix}$$

This gives

$$a_0 = 0.9963, \quad a_1 = 1.1037, \quad a_2 = 0.5368.$$

The linear least-squares polynomial $P_2(x) = 0.9963 + 1.1037x + 0.5368x^2$.

Accuracy Check:

$$P_2(0.5) = 1.6889, \quad e^{0.5} = 1.6487.$$

Relative Error:

$$\frac{|e^{0.5} - P_1(0.5)|}{|e^{0.5}|} = \frac{|1.6487 - 1.6889|}{|1.6487|} = 0.0204.$$