Chapter 1:

Introduction to Numerical Methods

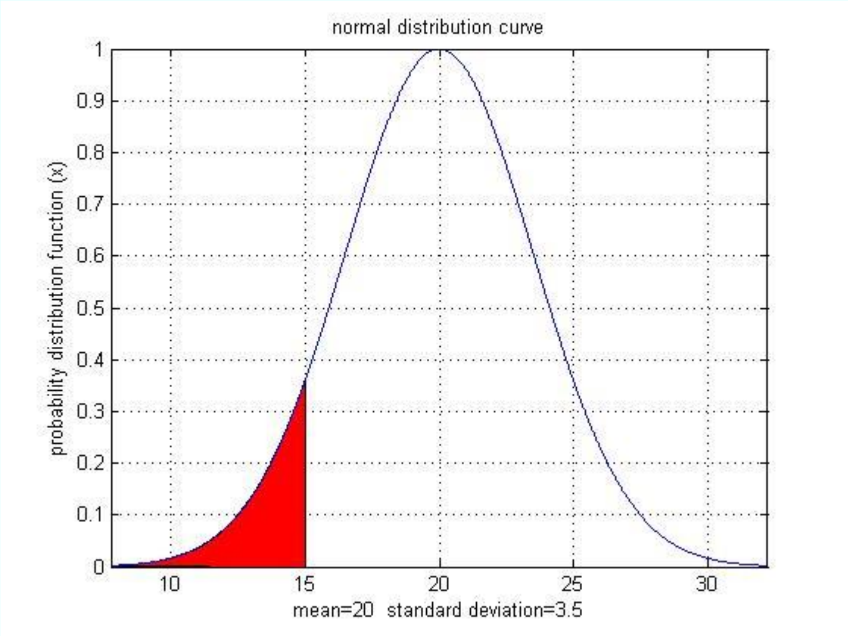
http://numericalmethods.eng.usf.edu

INTRODUCTION

Why use Numerical Methods?

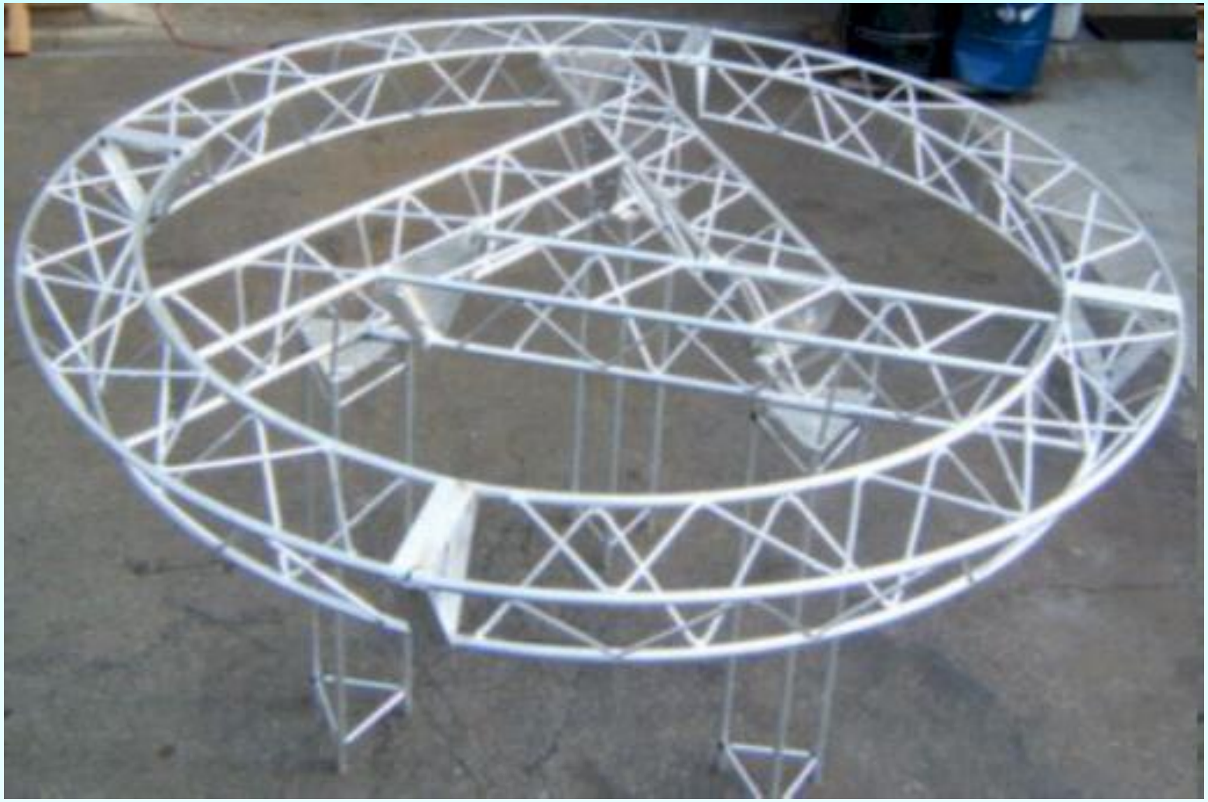
- To solve problems that cannot be solved exactly

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}\,du$$



Why use Numerical Methods?

- To solve problems that are intractable!



How do we solve an engineering problem?

Problem Description

↓

Mathematical Model

↓

Solution of Mathematical Model

↓

Using the Solution

1. Introduction to Numerical Methods

Mathematical Procedures

http://numericalmethods.eng.usf.edu

Introduction to Numerical Methods

Mathematical Procedures

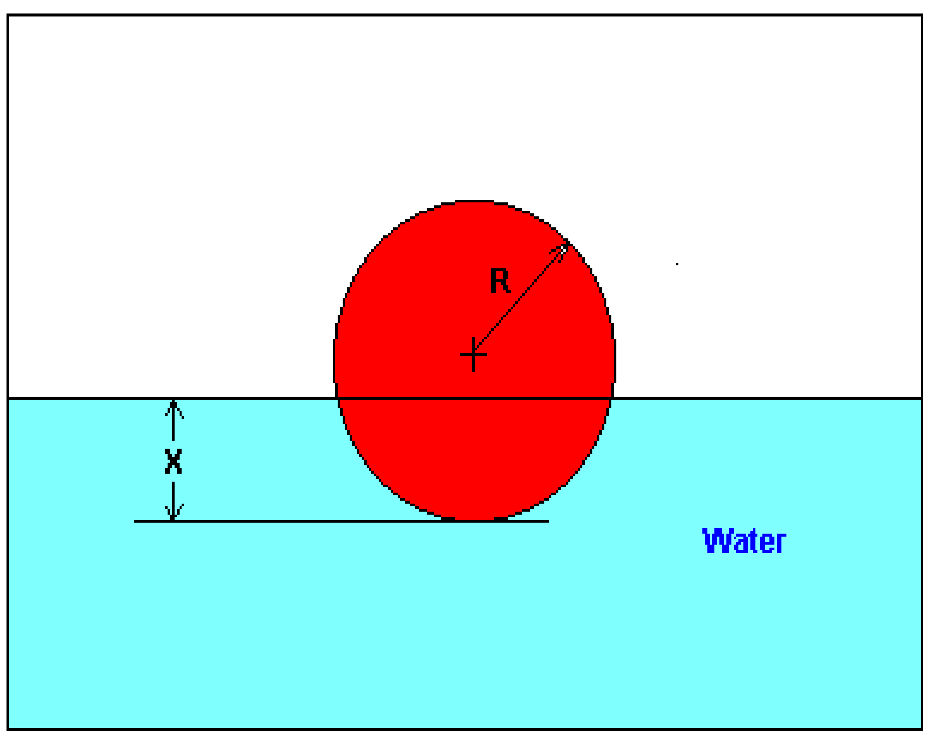
- Nonlinear Equations
- Differentiation
- Simultaneous Linear Equations
- Curve Fitting
  - Interpolation
  - Regression
- Integration
- Ordinary Differential Equations
- Other Advanced Mathematical Procedures:
  - Partial Differential Equations
  - Optimization
  - Fast Fourier Transforms

Nonlinear Equations

How much of the floating ball is under water?

Diameter=0.11m

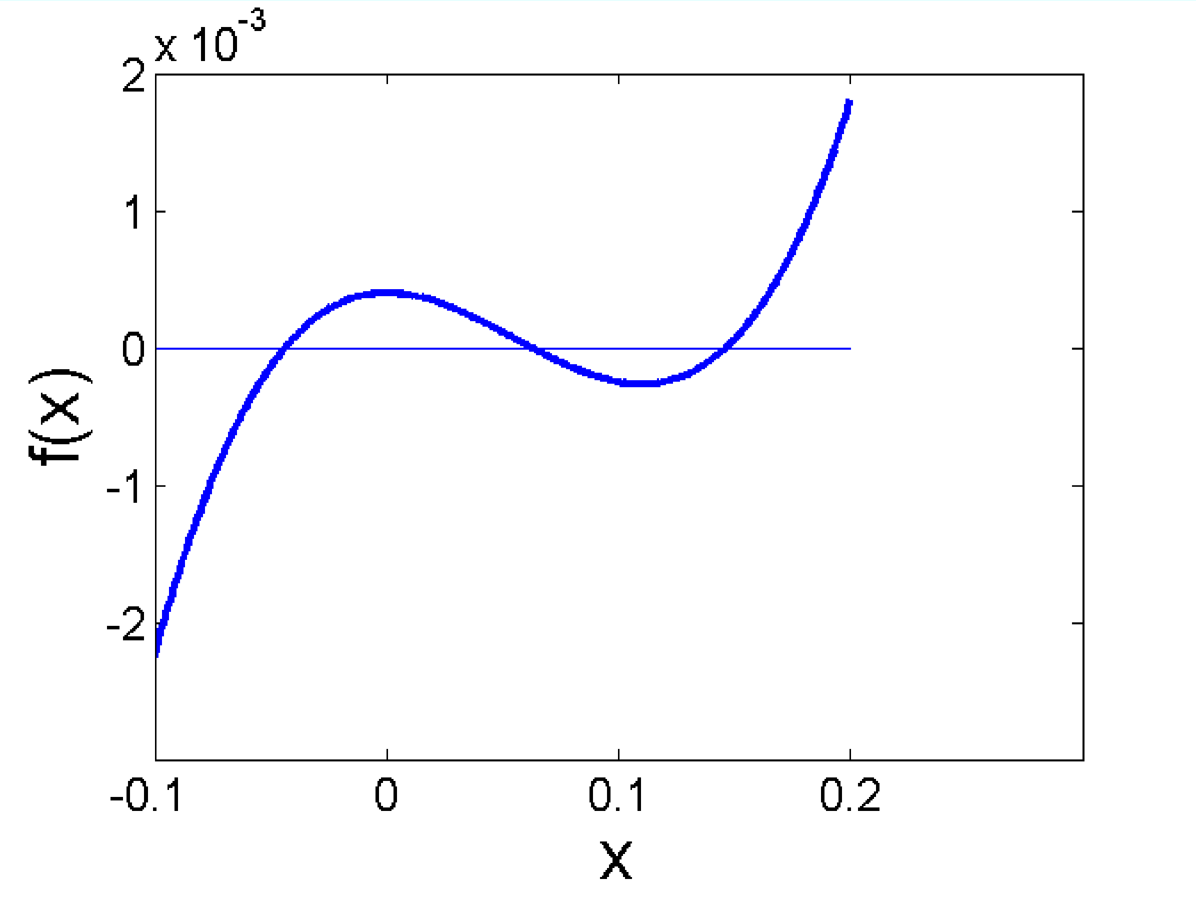
Specific Gravity=0.6



$$x^3 - 0.165x^2 + 3.993\times10^{-4} = 0$$

Nonlinear Equations

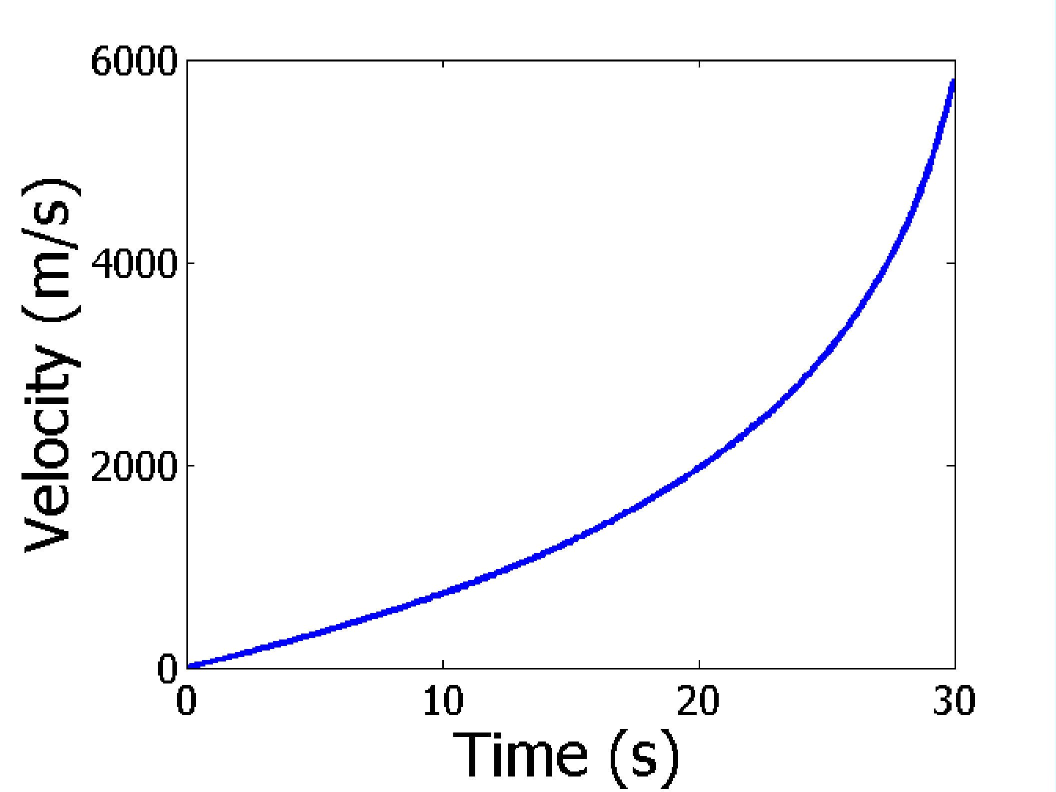
How much of the floating ball is under the water?



$$f(x) = x^3 - 0.165x^2 + 3.993\times10^{-4} = 0$$

Differentiation

What is the acceleration at t=7 seconds?



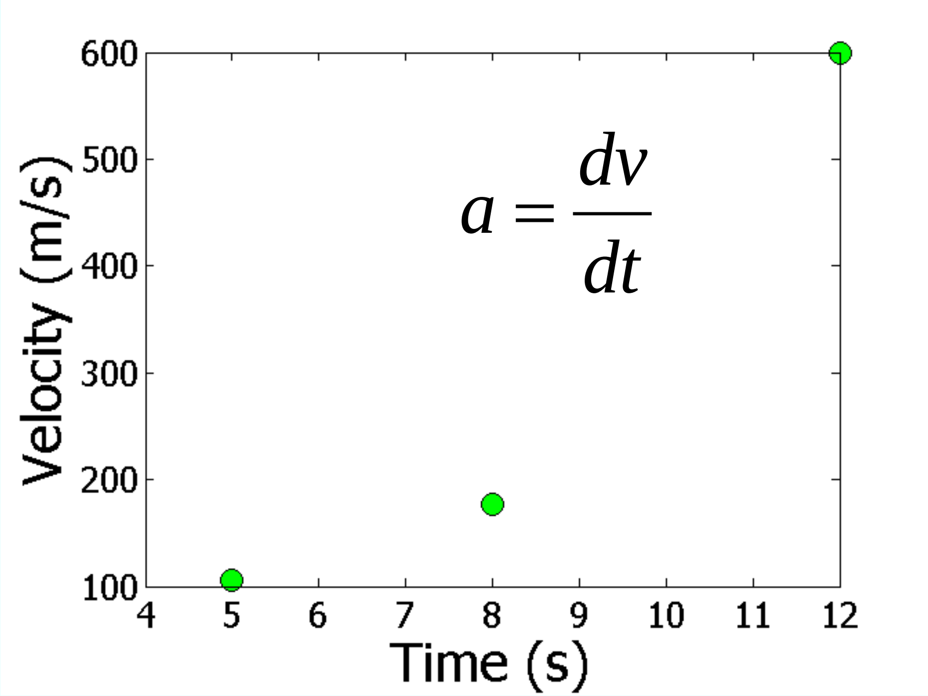
$$v(t) = 2200 \ln\left(\frac{16\times10^4}{16\times10^4 - 5000t}\right) - 9.8t$$

$$a = \frac{dv}{dt}$$

Differentiation

What is the acceleration at t=7 seconds?

Time (s)	5	8	12
Vel (m/s)	106	177	600



Simultaneous Linear Equations

Find the velocity profile, given

Time (s)	5	8	12
Vel (m/s)	106	177	600

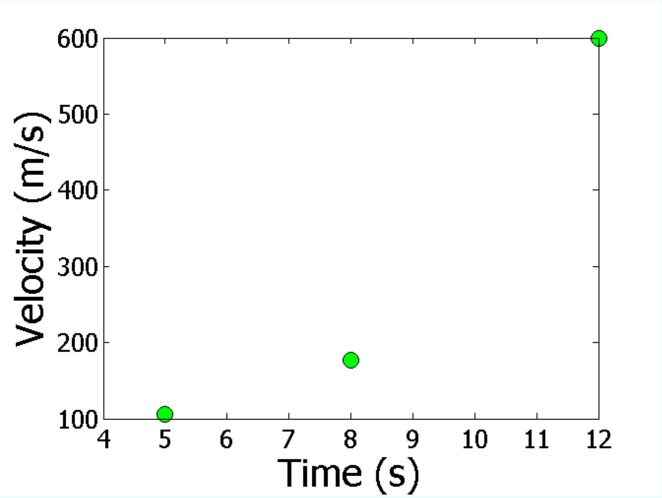
$$v(t) = at^2 + bt + c,\ 5 \le t \le 12$$

Three simultaneous linear equations

$$25a + 5b + c = 106$$

$$64a + 8b + c = 177$$

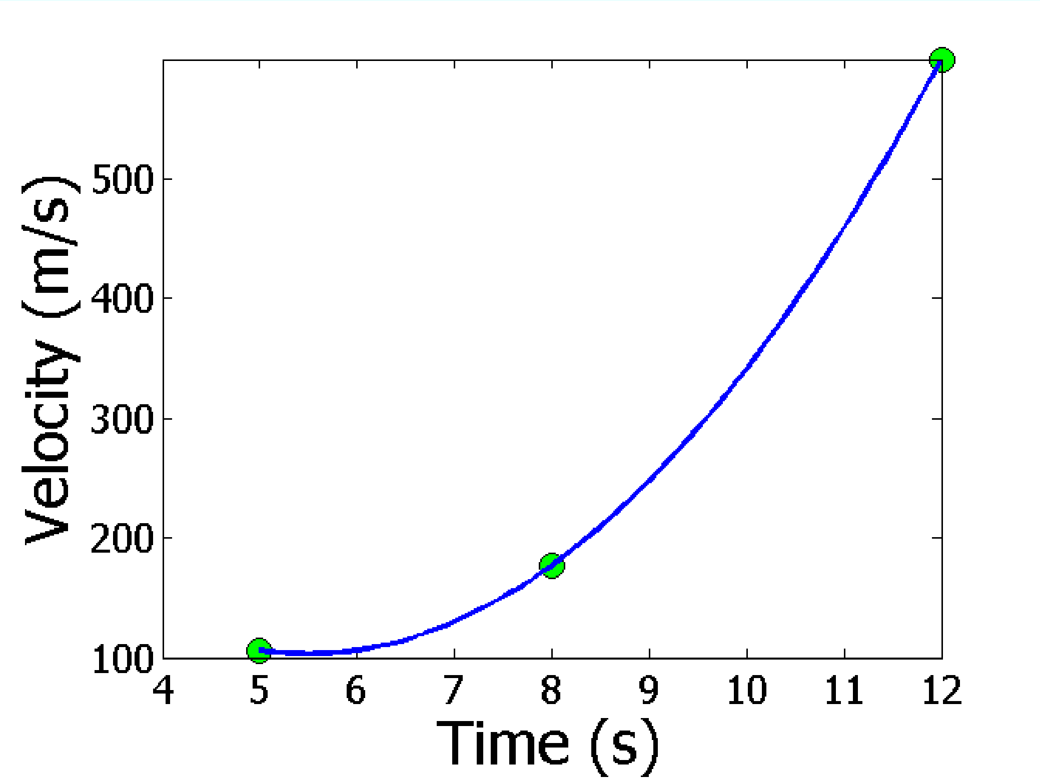
$$144a + 12b + c = 600$$



Interpolation

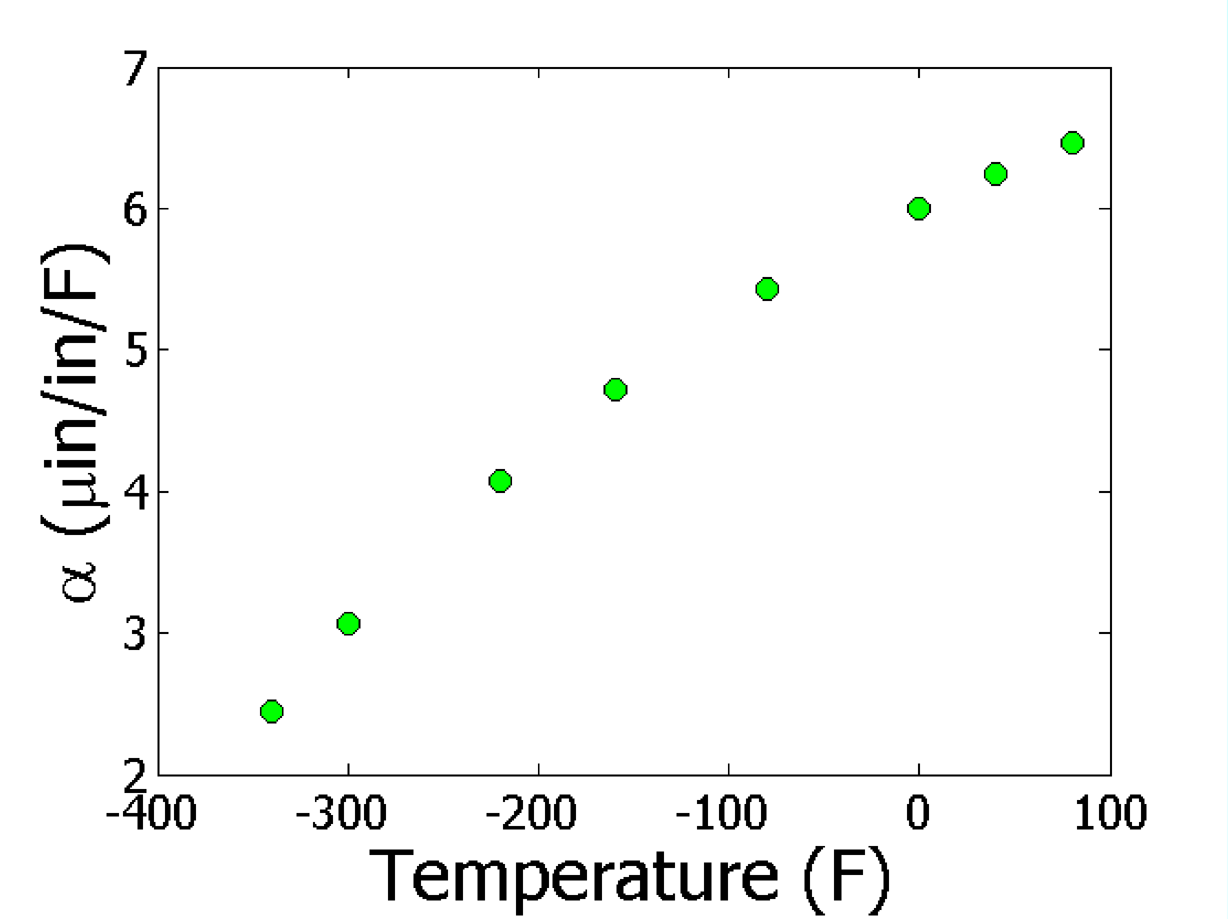
What is the velocity of the rocket at t=7 seconds?

Time (s)	5	8	12
Vel (m/s)	106	177	600

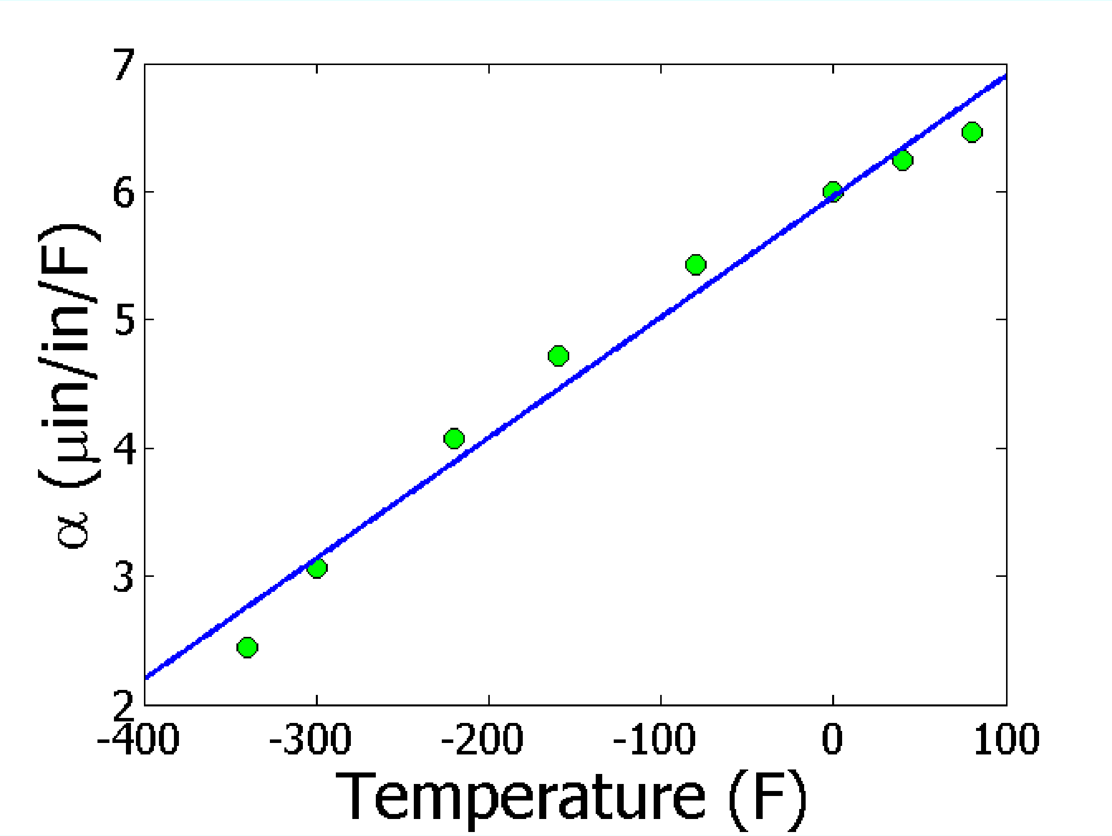


Regression

Thermal expansion coefficient data for cast steel



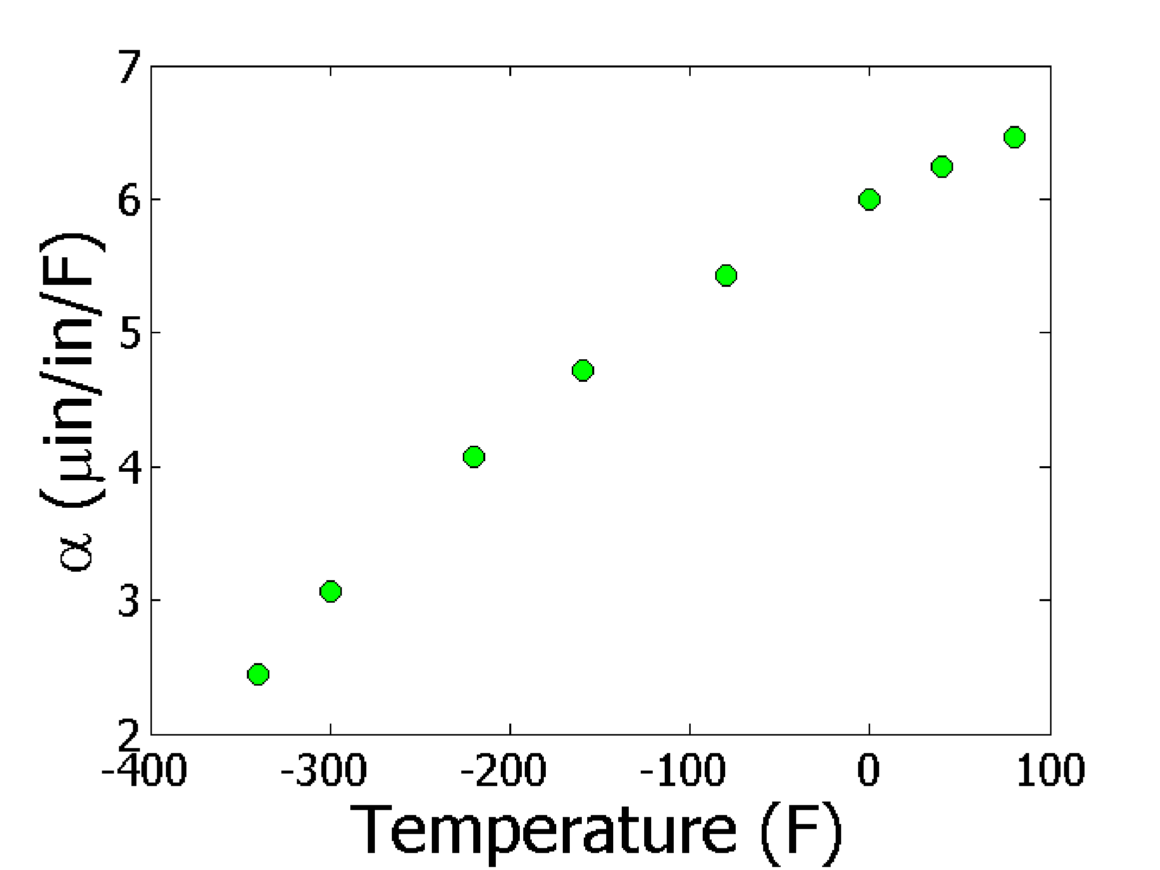
Regression (cont)



Integration

Finding the diametric contraction in a steel shaft when dipped in liquid nitrogen.

$$\Delta D = D \int_{T_{room}}^{T_{fluid}} \alpha \, dT$$



Ordinary Differential Equations

How long does it take a trunnion to cool down?

$$mc\frac{d\theta}{dt} = -hA(\theta - \theta_a), \quad \theta(0) = \theta_{room}$$

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/introduction_numerical.html

2. Measuring Errors

http://numericalmethods.eng.usf.edu

Why measure errors?

1) To determine the accuracy of numerical results.
2) To develop stopping criteria for iterative algorithms.

True Error

- Defined as the difference between the true value in a calculation and the approximate value found using a numerical method etc.

True Error = True Value – Approximate Value

Example—True Error

The derivative, $f'(x)$ of a function $f(x)$ can be approximated by the equation,

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

If $f(x) = 7e^{0.5x}$ and $h = 0.3$

a) Find the approximate value of $f'(2)$

b) True value of $f'(2)$

c) True error for part (a)

Example (cont.)

Solution:

a) For $x = 2$ and $h = 0.3$

$$f'(2) \approx \frac{f(2+0.3) - f(2)}{0.3}$$

$$= \frac{f(2.3) - f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107 - 19.028}{0.3} = 10.263$$

Example (cont.)

Solution:

b) The exact value of $f'(2)$ can be found by using our knowledge of differential calculus.

$$f(x) = 7e^{0.5x}$$

$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$

$$= 3.5e^{0.5x}$$

So the true value of $f'(2)$ is

$$f'(2) = 3.5e^{0.5(2)}$$

$$= 9.5140$$

True error is calculated as

$E_t =$ True Value – Approximate Value

$$= 9.5140 - 10.263 = -0.722$$

Relative True Error

- Defined as the ratio between the true error, and the true value.

$$\text{Relative True Error } (\in_t) = \frac{\text{True Error}}{\text{True Value}}$$

Example—Relative True Error

Following from the previous example for true error, find the relative true error for $f(x) = 7e^{0.5x}$ at $f'(2)$ with $h = 0.3$

From the previous example,

$$E_t = -0.722$$

Relative True Error is defined as

$$\in_t = \frac{\text{True Error}}{\text{True Value}}$$

$$= \frac{-0.722}{9.5140} = -0.075888$$

as a percentage,

$$\in_t = -0.075888 \times 100\% = -7.5888\%$$

Approximate Error

- What can be done if true values are not known or are very difficult to obtain?
- Approximate error is defined as the difference between the present approximation and the previous approximation.

Approximate Error (E_a) = Present Approximation – Previous Approximation

Example—Approximate Error

For $f(x) = 7e^{0.5x}$ at $x = 2$ find the following,

a) $f'(2)$ using $h = 0.3$

b) $f'(2)$ using $h = 0.15$

c) approximate error for the value of $f'(2)$ for part b)

Solution:

a) For $x = 2$ and $h = 0.3$

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$f'(2) \approx \frac{f(2+0.3) - f(2)}{0.3}$$

Example (cont.)

Solution: (cont.)

$$= \frac{f(2.3) - f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107 - 19.028}{0.3} = 10.263$$

b) For $x = 2$ and $h = 0.15$

$$f^{'}(2) \approx \frac{f(2+0.15) - f(2)}{0.15}$$

$$= \frac{f(2.15) - f(2)}{0.15}$$

Example (cont.)

Solution: (cont.)

$$= \frac{7e^{0.5(2.15)} - 7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50 - 19.028}{0.15} = 9.8800$$

c) So the approximate error, E_a is

$E_a =$ Present Approximation – Previous Approximation

$$= 9.8800 - 10.263$$

$$= -0.38300$$

Relative Approximate Error

- Defined as the ratio between the approximate error and the present approximation.

$$\text{Relative Approximate Error } (\in_a) = \frac{\text{Approximate Error}}{\text{Present Approximation}}$$

Example—Relative Approximate Error

For $f(x)=7e^{0.5x}$ at $x=2$, find the relative approximate error using values from $h=0.3$ and $h=0.15$

Solution:

From Example 3, the approximate value of $f'(2)=10.263$ using $h=0.3$ and $f'(2)=9.8800$ using $h=0.15$

$E_a =$ Present Approximation – Previous Approximation

$$=9.8800-10.263$$

$$=-0.38300$$

Example (cont.)

Solution: (cont.)

$$\in_a = \frac{\text{Approximate Error}}{\text{Present Approximation}}$$

$$= \frac{-0.38300}{9.8800} = -0.038765$$

as a percentage,

$$\in_a = -0.038765 \times 100\% = -3.8765\%$$

Absolute relative approximate errors may also need to be calculated,

$$\left|\in_a\right| = |-0.038765| = 0.038765 \text{ or } 3.8765\%$$

How is Absolute Relative Error used as a stopping criterion?

If $|\in_a| \leq \in_s$ where $\in_s$ is a pre-specified tolerance, then no further iterations are necessary and the process is stopped.

If at least m significant digits are required to be correct in the final answer, then

$$|\in_a| \leq 0.5 \times 10^{2-m}\%$$

Table of Values

For $f(x) = 7e^{0.5x}$ at $x = 2$ with varying step size, h

h	$f'(2)$	$\lvert\in_a\rvert$	m
0.3	10.263	N/A	0
0.15	9.8800	3.877%	1
0.10	9.7558	1.273%	1
0.01	9.5378	2.285%	1
0.001	9.5164	0.2249%	2

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/measuring_errors.html

3. Sources of Error

http://numericalmethods.eng.usf.edu

Two sources of numerical error

1) Round off error
2) Truncation error

Round-off Error

http://numericalmethods.eng.usf.edu

Round off Error

- Caused by representing a number approximately

$$\frac{1}{3} \cong 0.333333$$

$$\sqrt{2} \cong 1.4142...$$

Truncation Error

http://numericalmethods.eng.usf.edu

Truncation error

- Error caused by truncating or approximating a mathematical procedure.

Example of Truncation Error

Taking only a few terms of a Maclaurin series to approximate e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\dots\dots\dots\dots\dots$$

If only 3 terms are used,

$$Truncation \;\; Error = e^x - \left(1 + x + \frac{x^2}{2!}\right)$$

Another Example of Truncation Error

Using a finite Δx to approximate $f'(x)$

$$f'(x) \approx \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

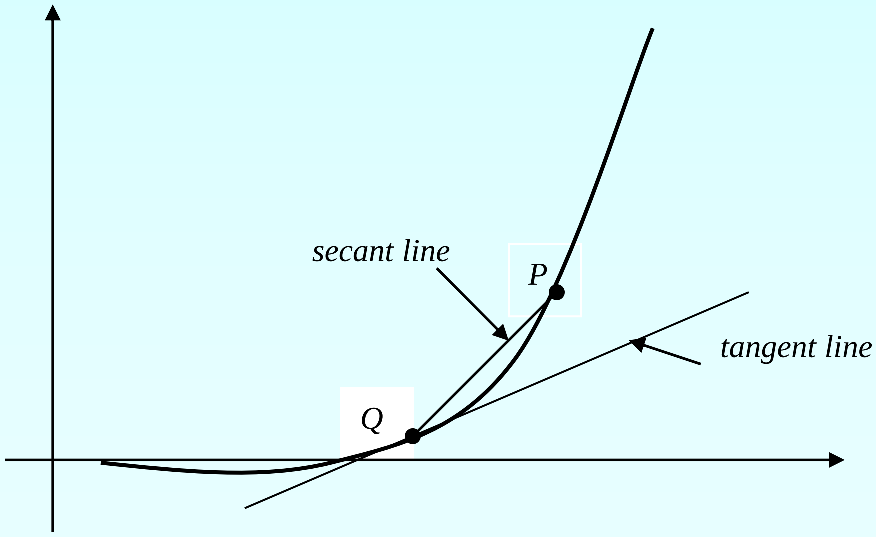
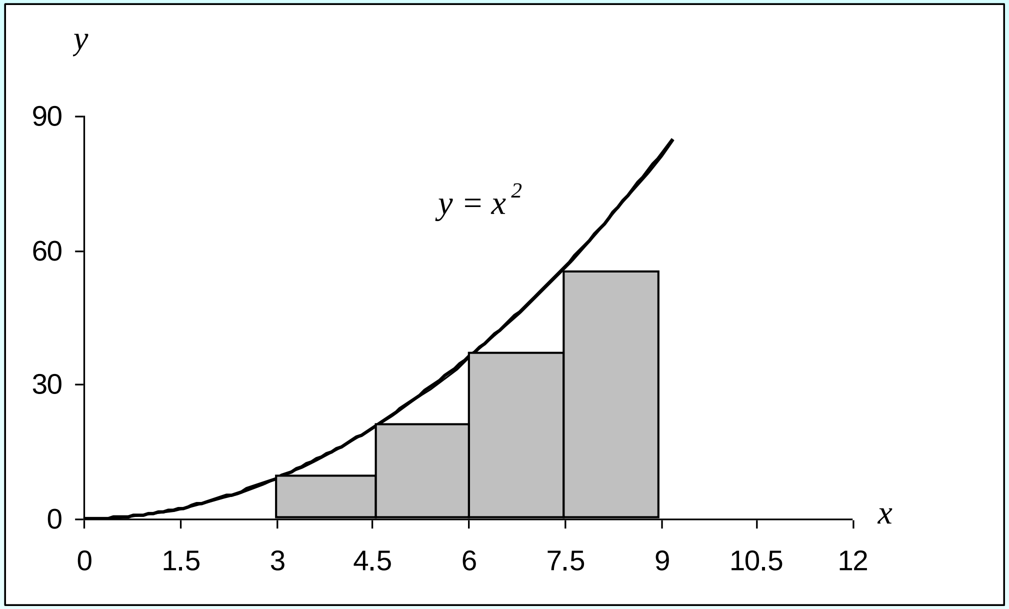


Figure 1. Approximate derivative using finite Δx

Another Example of Truncation Error

Using finite rectangles to approximate an integral.



Example 1 —Maclaurin series

Calculate the value of $e^{1.2}$ with an absolute relative approximate error of less than 1%.

$$e^{1.2} = 1 + 1.2 + \frac{1.2^2}{2!} + \frac{1.2^3}{3!} + \ldots\ldots\ldots\ldots\ldots\ldots$$

n	$e^{1.2}$	E_a	$\left\|\in_a\right\|\%$
1	1	__	___
2	2.2	1.2	54.545
3	2.92	0.72	24.658
4	3.208	0.288	8.9776
5	3.2944	0.0864	2.6226
6	3.3151	0.020736	0.62550

6 terms are required. How many are required to get at least 1 significant digit correct in your answer?

Example 2 —Differentiation

Find $f'(3)$ for $f(x) = x^2$ using $f'(x) \approx \dfrac{f(x+\Delta x) - f(x)}{\Delta x}$
and $\Delta x = 0.2$

$$f^{'}(3) = \frac{f(3+0.2) - f(3)}{0.2}$$

$$= \frac{f(3.2) - f(3)}{0.2} = \frac{3.2^2 - 3^2}{0.2} = \frac{10.24 - 9}{0.2} = \frac{1.24}{0.2} = 6.2$$

The actual value is

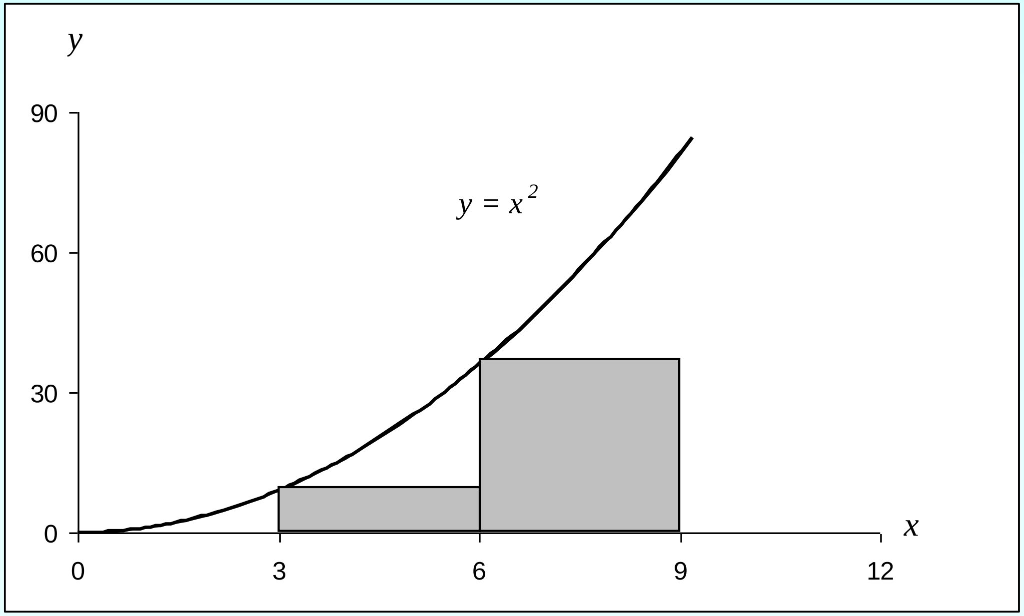
$$f^{'}(x) = 2x, \quad f^{'}(3) = 2 \times 3 = 6$$

Truncation error is then, $6 - 6.2 = -0.2$

Can you find the truncation error with $\Delta x = 0.1$

Example 3 — Integration

Use two rectangles of equal width to approximate the area under the curve for $f(x) = x^2$ over the interval $[3,9]$



$$\int_3^9 x^2 dx$$

Integration example (cont.)

Choosing a width of 3, we have

$$\int_3^9 x^2 dx = (x^2)\Big|_{x=3}(6-3) + (x^2)\Big|_{x=6}(9-6)$$

$$= (3^2)3 + (6^2)3$$

$$= 27 + 108 = 135$$

Actual value is given by

$$\int_3^9 x^2 dx = \left[\frac{x^3}{3}\right]_3^9 = \left[\frac{9^3 - 3^3}{3}\right] = 234$$

Truncation error is then

$$234 - 135 = 99$$

Can you find the truncation error with 4 rectangles?

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/sources_of_error.html

4. Binary Representation

http://numericalmethods.eng.usf.edu

How a Decimal Number is Represented

$$257.76 = 2\times10^{2} + 5\times10^{1} + 7\times10^{0} + 7\times10^{-1} + 6\times10^{-2}$$

Base 2

$$(1011.0011)_2 = \begin{pmatrix} (1\times 2^3 + 0\times 2^2 + 1\times 2^1 + 1\times 2^0) \\ + (0\times 2^{-1} + 0\times 2^{-2} + 1\times 2^{-3} + 1\times 2^{-4}) \end{pmatrix}_{10}$$

$$= 11.1875$$

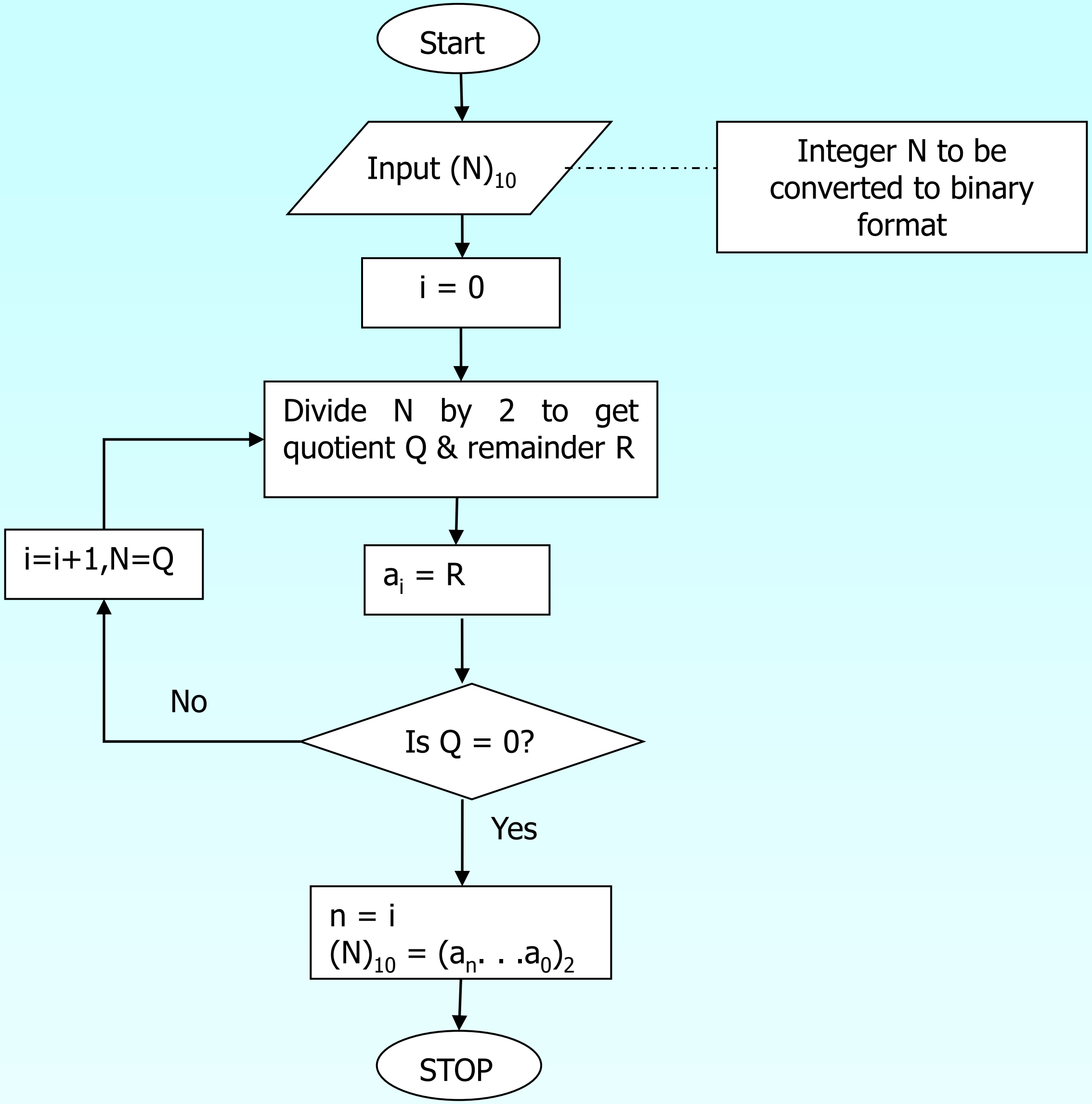
Convert Base 10 Integer to binary representation

Table 1 Converting a base-10 integer to binary representation.

	Quotient	**Remainder**
11/2	5	$1 = a_0$
5/2	2	$1 = a_1$
2/2	1	$0 = a_2$
1/2	0	$1 = a_3$

Hence

$$(11)_{10} = (a_3 a_2 a_1 a_0)_2$$
$$= (1011)_2$$


Start
Input $(N)_{10}$
Integer N to be converted to binary format
i = 0
Divide N by 2 to get quotient Q & remainder R
i=i+1,N=Q
$a_i = R$
No
Is Q = 0?
Yes
n = i
$(N)_{10} = (a_n. . .a_0)_2$
STOP

Fractional Decimal Number to Binary

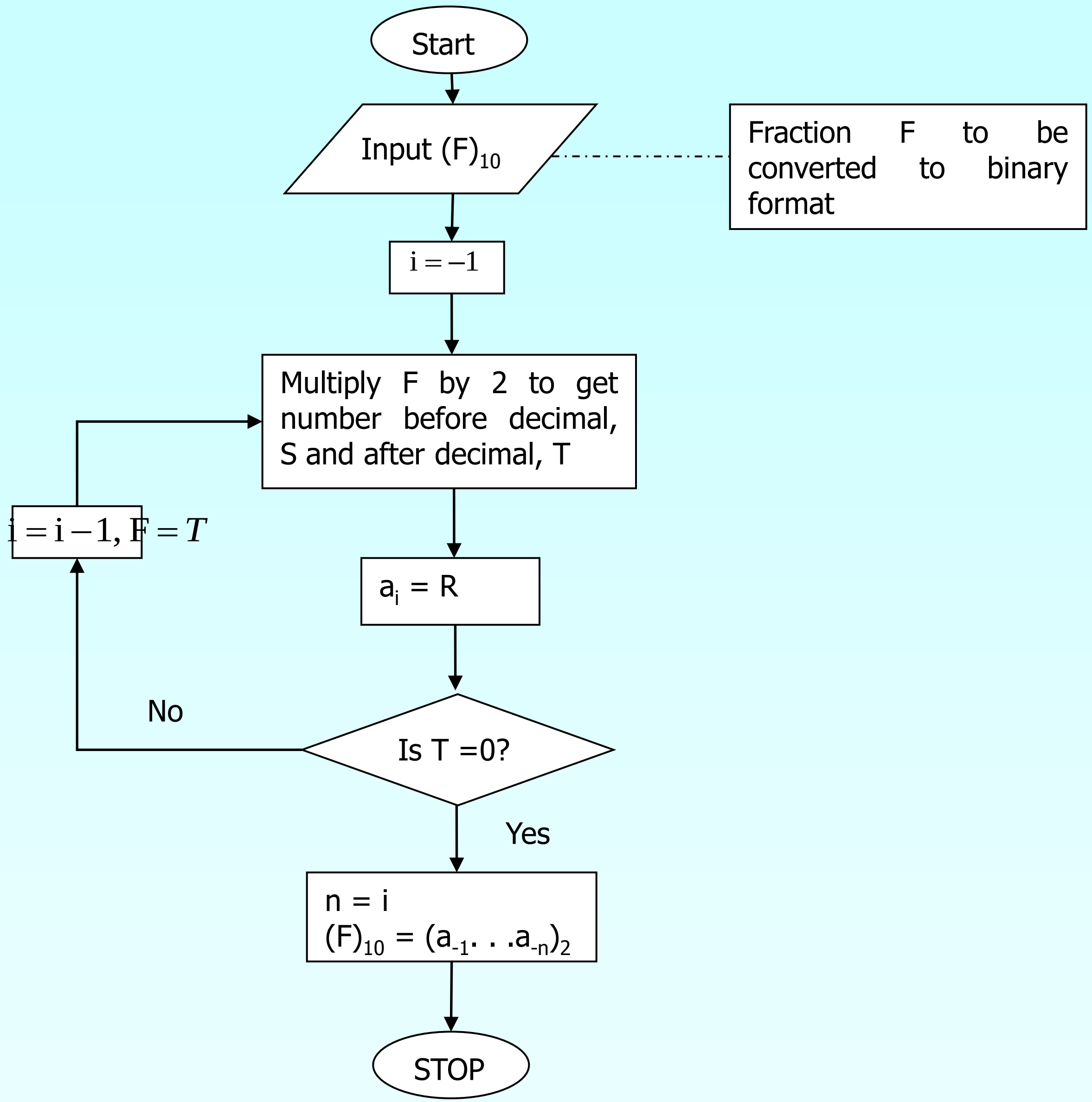
Table 2. Converting a base-10 fraction to binary representation.

	Number	Number after decimal	Number before decimal
0.1875×2	0.375	0.375	$0 = a_{-1}$
0.375×2	0.75	0.75	$0 = a_{-2}$
0.75×2	1.5	0.5	$1 = a_{-3}$
0.5×2	1.0	0.0	$1 = a_{-4}$

Hence

$$(0.1875)_{10} = (a_{-1}a_{-2}a_{-3}a_{-4})_2$$

$$= (0.0011)_2$$



Start
Input $(F)_{10}$
Fraction F to be converted to binary format
$i = -1$
Multiply F by 2 to get number before decimal, S and after decimal, T
$i = i - 1, F = T$
$a_i = R$
No
Is T =0?
Yes
n = i
$(F)_{10} = (a_{-1}. . .a_{-n})_2$
STOP

Decimal Number to Binary

$$(11.1875)_{10} = (\quad ?.? \quad)_2$$

Since

$$(11)_{10} = (1011)_2$$

and

$$(0.1875)_{10} = (0.0011)_2$$

we have

$$(11.1875)_{10} = (1011.0011)_2$$

All Fractional Decimal Numbers Cannot be Represented Exactly

Table 3. Converting a base-10 fraction to approximate binary representation.

	Number	**Number after decimal**	**Number before Decimal**
0.3×2	0.6	0.6	$0=a_{-1}$
0.6×2	1.2	0.2	$1=a_{-2}$
0.2×2	0.4	0.4	$0=a_{-3}$
0.4×2	0.8	0.8	$0=a_{-4}$
0.8×2	1.6	0.6	$1=a_{-5}$

$$(0.3)_{10} \approx (a_{-1}a_{-2}a_{-3}a_{-4}a_{-5})_2 = (0.01001)_2 = 0.28125$$

Another Way to Look at Conversion

Convert $(11.1875)_{10}$ to base 2

$$
\begin{aligned}
(11)_{10} &= 2^3 + 3 \\
&= 2^3 + 2^1 + 1 \\
&= 2^3 + 2^1 + 2^0 \\
&= 1\times 2^3 + 0\times 2^2 + 1\times 2^1 + 1\times 2^0 \\
&= (1011)_2
\end{aligned}
$$

$$
\begin{aligned}
(0.1875)_{10} &= 2^{-3} + 0.0625 \\
&= 2^{-3} + 2^{-4} \\
&= 0\times 2^{-1} + 0\times 2^{-2} + 1\times 2^{-3} + 1\times 2^{-4} \\
&= (.0011)_2
\end{aligned}
$$

$$
(11.1875)_{10} = (1011.0011)_2
$$

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/binary_representation.html

5. Floating Point Representation

http://numericalmethods.eng.usf.edu

Floating Decimal Point : Scientific Form

$$256.78 \text{ is written as } +2.5678\times10^{2}$$

$$0.003678 \text{ is written as } +3.678\times10^{-3}$$

$$-256.78 \text{ is written as } -2.5678\times10^{2}$$

Example

The form is

$$\text{sign} \times \text{mantissa} \times 10^{\text{exponent}}$$

or

$$\sigma \times m \times 10^{e}$$

Example: For

$$-2.5678 \times 10^{2}$$

$$\sigma = -1$$

$$m = 2.5678$$

$$e = 2$$

Floating Point Format for Binary Numbers

$$y = \sigma \times m \times 2^{e}$$

$$\sigma = \text{sign of number } \left(0 \text{ for } + \text{ve}, 1 \text{ for - ve}\right)$$

$$m = \text{mantissa} \left[(1)_2 < m < (10)_2\right]$$

1 is not stored as it is always given to be 1.

$$e = \text{integer exponent}$$

Example

9 bit-hypothetical word

- the first bit is used for the sign of the number,
- the second bit for the sign of the exponent,
- the next four bits for the mantissa, and
- the next three bits for the exponent

$$(54.75)_{10} = (110110.11)_2 = (1.1011011)_2 \times 2^5$$
$$\cong (1.1011)_2 \times (101)_2$$

We have the representation as

0	0	1	0	1	1	1	0	1

$$\underset{\text{Sign of the number}}{0}\;\;\underset{\text{Sign of the exponent}}{0}\;\;\underbrace{1\;0\;1\;1}_{\text{mantissa}}\;\;\underbrace{1\;0\;1}_{\text{exponent}}$$

Machine Epsilon

Defined as the measure of accuracy and found by difference between 1 and the next number that can be represented

Example

Ten bit word

- Sign of number
- Sign of exponent
- Next four bits for exponent
- Next four bits for mantissa

0	0	0	0	0	0	0	0	0	0

$= (1)_{10}$

Next number →

0	0	0	0	0	0	0	0	0	1

$= (1.0001)_2 = (1.0625)_{10}$

$$\in_{mach} = 1.0625 - 1 = 2^{-4}$$

Relative Error and Machine Epsilon

The absolute relative true error in representing a number will be less then the machine epsilon

Example

$$(0.02832)_{10} \cong (1.1100)_2 \times 2^{-5}$$

$$= (1.1100)_2 \times 2^{-(0110)_2}$$

10 bit word (sign, sign of exponent, 4 for exponent, 4 for mantissa)

0	1	0	1	1	0	1	1	0	0

Sign of the number

Sign of the exponent

exponent

mantissa

$$(1.1100)_2 \times 2^{-(0110)_2} = 0.0274375$$

$$\in_a = \left| \frac{0.02832 - 0.0274375}{0.02832} \right|$$

$$= 0.034472 < 2^{-4} = 0.0625$$

IEEE 754 Standards for Single Precision Representation

http://numericalmethods.eng.usf.edu

IEEE-754 Floating Point Standard

- Standardizes representation of floating point numbers on different computers in single and double precision.

- Standardizes representation of floating point operations on different computers.

One Great Reference

What every computer scientist (and even if you are not) should know about floating point arithmetic!

http://www.validlab.com/goldberg/paper.pdf

IEEE-754 Format Single Precision

32 bits for single precision

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Sign (s)

Biased Exponent (e’)

Mantissa (m)

$$\text{Value} = (-1)^{s} \times (1 \bullet m)_2 \times 2^{e'-127}$$

Example#1

$$\underbrace{1}_{\substack{\text{Sign}\\ \text{(s)}}}\ \underbrace{1\,0\,1\,0\,0\,0\,1\,0}_{\substack{\text{Biased}\\ \text{Exponent (e')}}}\ \underbrace{1\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0}_{\text{Mantissa (m)}}$$

$$\text{Value} = (-1)^s \times (1.m)_2 \times 2^{e'-127}$$

$$= (-1)^1 \times (1.10100000)_2 \times 2^{(10100010)_2 - 127}$$

$$= (-1) \times (1.625) \times 2^{162-127}$$

$$= (-1) \times (1.625) \times 2^{35} = -5.5834 \times 10^{10}$$

Example#2

Represent -5.5834x10^{10} as a single precision floating point number.

?	? ? ? ? ? ? ? ?	? ?
Sign (s)	Biased Exponent (e′)	Mantissa (m)

$$-5.5834\times10^{10}=(-1)^{1}\times(1.?)\times2^{\pm?}$$

Exponent for 32 Bit IEEE-754

8 bits would represent

$$0 \le e' \le 255$$

Bias is 127; so subtract 127 from representation

$$-127 \le e \le 128$$

Exponent for Special Cases

Actual range of e'

$$1 \le e' \le 254$$

$e' = 0$ and $e' = 255$ are reserved for special numbers

Actual range of e

$$-126 \le e \le 127$$

Special Exponents and Numbers

$e' = 0$ — all zeros

$e' = 255$ — all ones

s	e'	m	Represents
0	all zeros	all zeros	0
1	all zeros	all zeros	-0
0	all ones	all zeros	∞
1	all ones	all zeros	$-\infty$
0 or 1	all ones	non-zero	NaN

IEEE-754 Format

The largest number by magnitude

$$\left(1.1........1\right)_2 \times 2^{127} = 3.40 \times 10^{38}$$

The smallest number by magnitude

$$\left(1.00......0\right)_2 \times 2^{-126} = 2.18 \times 10^{-38}$$

Machine epsilon

$$\varepsilon_{mach} = 2^{-23} = 1.19 \times 10^{-7}$$

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/floatingpoint_representation.html

6. Propagation of Errors

http://numericalmethods.eng.usf.edu

Propagation of Errors

In numerical methods, the calculations are not made with exact numbers. How do these inaccuracies propagate through the calculations?

Example 1:

Find the bounds for the propagation in adding two numbers. For example if one is calculating $X+Y$ where

$$X = 1.5 \pm 0.05$$

$$Y = 3.4 \pm 0.04$$

Solution

Maximum possible value of $X = 1.55$ and $Y = 3.44$

Maximum possible value of $X + Y = 1.55 + 3.44 = 4.99$

Minimum possible value of $X = 1.45$ and $Y = 3.36$.

Minimum possible value of $X + Y = 1.45 + 3.36 = 4.81$

Hence

$$4.81 \le X + Y \le 4.99.$$

Propagation of Errors In Formulas

If f is a function of several variables $X_1, X_2, X_3, \ldots\ldots\ldots, X_{n-1}, X_n$ then the maximum possible value of the error in f is

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots\ldots + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Example 2:

The strain in an axial member of a square cross-section is given by

$$\in = \frac{F}{h^2 E}$$

Given

$$F = 72 \pm 0.9 \text{ N}$$
$$h = 4 \pm 0.1 \text{ mm}$$
$$E = 70 \pm 1.5 \text{ GPa}$$

Find the maximum possible error in the measured strain.

Example 2:

Solution

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)}$$

$$= 64.286\times10^{-6}$$

$$= 64.286\mu$$

$$\Delta \in = \left|\frac{\partial \in}{\partial F}\Delta F\right| + \left|\frac{\partial \in}{\partial h}\Delta h\right| + \left|\frac{\partial \in}{\partial E}\Delta E\right|$$

Example 2:

$$\frac{\partial \in}{\partial F}=\frac{1}{h^2E} \qquad \frac{\partial \in}{\partial h}=-\frac{2F}{h^3E} \qquad \frac{\partial \in}{\partial E}=-\frac{F}{h^2E^2}$$

Thus

$$\Delta E=\left|\frac{1}{h^2E}\Delta F\right|+\left|\frac{2F}{h^3E}\Delta h\right|+\left|\frac{F}{h^2E^2}\Delta E\right|$$

$$=\left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times0.9\right|+\left|\frac{2\times72}{(4\times10^{-3})^3(70\times10^9)}\times0.0001\right|$$

$$+\left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times1.5\times10^9\right|$$

$$=5.3955\mu$$

Hence

$$\in=(64.286\mu\pm5.3955\mu)$$

Example 3:

Subtraction of numbers that are nearly equal can create unwanted inaccuracies. Using the formula for error propagation, show that this is true.

Solution

Let

$$z = x - y$$

Then

$$|\Delta z| = \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right|$$

$$= |(1)\Delta x| + |(-1)\Delta y|$$

$$= |\Delta x| + |\Delta y|$$

So the relative change is

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x - y|}$$

Example 3:

For example if

$$x = 2 \pm 0.001$$

$$y = 2.003 \pm 0.001$$

$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|\,2 - 2.003\,|}$$

$$= 0.6667$$

$$= 66.67\%$$

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/propagation_of_errors.html

7. Taylor Series Revisited

http://numericalmethods.eng.usf.edu

What is a Taylor series?

Some examples of Taylor series which you must have seen

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \Lambda$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \Lambda$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \Lambda$$

General Taylor Series

The general form of the Taylor series is given by

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \Lambda$$

provided that all derivatives of f(x) are continuous and exist in the interval [x,x+h]

What does this mean in plain English?

As Archimedes would have said, "*Give me the value of the function at a single point, and the value of all (first, second, and so on) its derivatives at that single point, and I can give you the value of the function at any other point*" (*fine print excluded*)

Example—Taylor Series

Find the value of $f(6)$ given that $f(4)=125$, $f'(4)=74$, $f''(4)=30$, $f'''(4)=6$ and all other higher order derivatives of $f(x)$ at $x=4$ are zero.

Solution:

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2!}+f'''(x)\frac{h^3}{3!}+\Lambda$$

$$x=4$$

$$h=6-4=2$$

Example (cont.)

Solution: (cont.)

Since the higher order derivatives are zero,

$$f(4+2) = f(4) + f'(4)2 + f''(4)\frac{2^2}{2!} + f'''(4)\frac{2^3}{3!}$$

$$f(6) = 125 + 74(2) + 30\left(\frac{2^2}{2!}\right) + 6\left(\frac{2^3}{3!}\right)$$

$$= 125 + 148 + 60 + 8$$

$$= 341$$

Note that to find $f(6)$ exactly, we only need the value of the function and all its derivatives at some other point, in this case $x = 4$

Derivation for Maclaurin Series for e^x

Derive the Maclaurin series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \Lambda$$

The Maclaurin series is simply the Taylor series about the point x=0

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + f''''(x)\frac{h^4}{4} + f'''''(x)\frac{h^5}{5} + \Lambda$$

$$f(0+h) = f(0) + f'(0)h + f''(0)\frac{h^2}{2!} + f'''(0)\frac{h^3}{3!} + f''''(0)\frac{h^4}{4} + f'''''(0)\frac{h^5}{5} + \Lambda$$

Derivation (cont.)

Since $f(x) = e^x$, $f'(x) = e^x$, $f''(x) = e^x$, ..., $f^n(x) = e^x$ and $f^n(0) = e^0 = 1$

the Maclaurin series is then

$$f(h) = (e^0) + (e^0)h + \frac{(e^0)}{2!}h^2 + \frac{(e^0)}{3!}h^3 \ldots$$

$$= 1 + h + \frac{1}{2!}h^2 + \frac{1}{3!}h^3 \ldots$$

So,

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Error in Taylor Series

The Taylor polynomial of order *n* of a function *f(x)* with *(n+1)* continuous derivatives in the domain [*x,x+h]* is given by

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + \Lambda + f^{(n)}(x)\frac{h^n}{n!} + R_n(x)$$

where the remainder is given by

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

where

$$x < c < x+h$$

that is, c is some point in the domain [x,x+h]

Example—error in Taylor series

The Taylor series for e^x at point $x = 0$ is given by

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \Lambda$$

It can be seen that as the number of terms used increases, the error bound decreases and hence a better estimate of the function can be found.

How many terms would it require to get an approximation of e^1 within a magnitude of true error of less than 10^{-6}.

Example—(cont.)

Solution:

Using $(n+1)$ terms of Taylor series gives error bound of

$$R_n(x)=\frac{(x-h)^{n+1}}{(n+1)!}f^{(n+1)}(c) \qquad x=0, h=1, f(x)=e^x$$

$$R_n(0)=\frac{(0-1)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

$$=\frac{(-1)^{n+1}}{(n+1)!}e^c$$

Since

$$x<c<x+h$$
$$0<c<0+1$$
$$0<c<1$$

$$\frac{1}{(n+1)!}<\left|R_n(0)\right|<\frac{e}{(n+1)!}$$

Example—(cont.)

Solution: (cont.)

So if we want to find out how many terms it would require to get an approximation of e^1 within a magnitude of true error of less than 10^{-6},

$$\frac{e}{(n+1)!} < 10^{-6}$$

$$(n+1)! > 10^6 e$$

$$(n+1)! > 10^6 \times 3$$

$$n \geq 9$$

So 9 terms or more are needed to get a true error less than 10^{-6}

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/taylor_series.html