Numerical Analysis

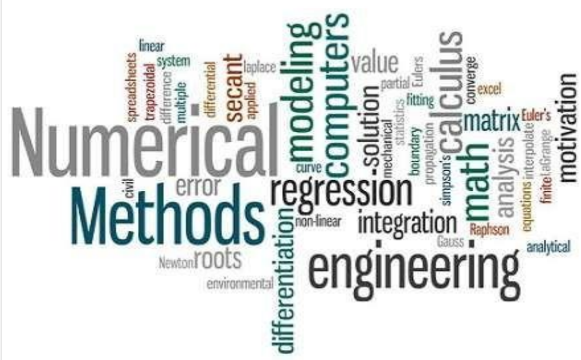
Dr. Farman Ali

Assistant Professor

DEPARTMENT OF SOFTWARE

SEJONG UNIVERSITY

Week 2, Lecture I-II



Course Syllabus

- **Introduction:** Numerical analysis and background, definitions of computer number systems, floating point representation, representation of numbers in different bases, and round-off errors.
- **Root Finding:** Bisection method, fixed-point iteration, Newton's method, the secant method and their error analysis, and order of convergence (Newton's method and Secant method).
- **Direct Methods for Solving Linear Systems:** Gaussian elimination, LU decomposition, pivoting strategies, and PA=LU-factorization,….. .
- **Polynomial:** Polynomial interpolation, piecewise linear interpolation, divided differences interpolation, cubic spline interpolation, and curve fitting in interpolation (Application: Regression).
- **Integration:** Numerical differentiation, numerical integration, and composite numerical integration.
- **Ordinary Differential Equations:** Euler's Method, higher-order Taylor method, and Runge-Kutta methods….

Representation of Real Number

Numbers Systems

Numbers Systems

- **Number systems** are the technique to represent numbers in the computer system architecture, every value that you are saving into or getting from computer memory has a defined number system.

Number system can be categorized

Decimal number system (10 digits)

Binary number system (2 digits)

Octal number system (8 digits)

Hexadecimal Number System (16 digits)

Numbering System	Base	Digits Set
Binary	2	1 0
Octal	8	7 6 5 4 3 2 1 0
Decimal	10	9 8 7 6 5 4 3 2 1 0
Hexadecimal	16	F E D C B A 9 8 7 6 5 4 3 2 1 0

Numbers Systems

Conversion of Number System

Decimal	Binary	Octal	Hexa-Decimal
0	0	0	0
1	1	1	1
2	10	2	2
3	11	3	3
4	100	4	4
5	101	5	5
6	110	6	6
7	111	7	7
8	1000	10	8
9	1001	11	9

Decimal	Binary	Octal	Hexa-Decimal
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F

Numbers Systems

Binary Number System

- Binary to Decimal Conversion Techniques:
  - Multiply each bit by 2^n, where n is the “weight” of the bit.
  - The weight is the position of the bit, starting from 0 on the right.
  - Add the results.

$$101011_2 = 43_{10}$$

Binary to Decimal Conversion

$$101011_2 = 43_{10}$$

$$= 1X2^5 + 0X2^4 + 1X2^3 + 0X2^2 + 1X2^1 + 1X2^0$$

$$= 32 + 0 + 8 + 0 + 2 + 1$$

$$= 43_{10}$$

Numbers Systems

Binary Number System

- Binary to Octal Conversion Techniques:
  - Group binary digits in a 3 bits, starting on right side
  - Convert to octal digits.
  - Example $1011010111_2 = 1327_8$

Binary to Octal Conversion

$1011010111 = 1327_8$

$$= 1 \quad 011 \quad 010 \quad 111$$

$$= 1327_8$$

Numbers Systems

Binary Number System

- Binary to Hexa-decimal Conversion Techniques:
  - Group binary digits in a 4 bits, starting on right side.
  - Convert to hexa-decimal digits.
  - Example $1010111011_2 = 2BB_{16}$

Binary to Hexa-decimal Conversion

$1010111011 = 2BB_{16}$

$$= 10 \quad 1011 \quad 1011$$

$$= 2BB_{16}$$

Numbers Systems

Octal Number System

- Octal to Decimal Conversion Techniques:
- Multiply each bit by 8^n, where n is the “weight” of the bit.
- The weight is the position of the bit, starting from 0 on the right.
- Add the results.

$$724_8 = 468_{10}$$

Octal to Decimal Conversion

$724 = 468_{10}$

$$= 7X8^2 + 2X8^1 + 4X8^0$$
$$= 448 + 16 + 4$$
$$= 468_{10}$$

Numbers Systems

Octal Number System

- Octal to Binary Conversion Techniques:
  - Convert octal digit in a 3 bits, starting on the right side.
  - Example $705_8 = 111000101_2$

Octal to Binary Conversion

$705 = 111000101_2$

$$= 7 \quad 0 \quad 5$$
$$= 111 \quad 000 \quad 101$$
$$= 111000101_2$$

Numbers Systems

Octal Number System

- Octal to Hexa-decimal Conversion Techniques:
  - Use binary as an intermediary
  - Example

$$1076_8 = 23E_{16}$$

Octal to Hexa-decimal Conversion

$1076 = 23E_{16}$

$$\begin{aligned}
&= 1 \quad 0 \quad 7 \quad 6 \\
&= 001 \quad 000 \quad 111 \quad 110 \\
&= 0010 \quad 0011 \quad 1110 \\
&= 23E_{16}
\end{aligned}$$

Numbers Systems

Decimal Number System

- Decimal to binary Conversion Techniques:
  - Divide it by 2, keep track of the remainder
  - First remainder is bit 0 (LSB, least-significant bit)

$$125_{10} = 1111101_2$$

2	125 --- 1
2	62---0
2	31---1
2	15---1
2	7---1
2	3---1
	1

Numbers Systems

Decimal Number System

- Decimal to Octal Conversion Techniques:
  - Divide it by 8, keep track of the remainder
  - First remainder is bit 0 (LSB, least-significant bit)

Example:

$$1234_{10} = 2322_8$$

8	1234--- 2
8	154--- 2
8	19--- 3
8	2

Numbers Systems

Decimal Number System

- Decimal to Hexa-decimal Conversion Techniques:
  - Divide it by 16, keep track of the remainder
  - First remainder is bit 0 (LSB, least-significant bit)

Example: $1234_{10} = 4D2_{16}$

16	1234--- 2
16	77--- 13
16	4

Numbers Systems

Hexa-decimal Number System

- Hexa-Decimal to Decimal Conversion Techniques
  - Multiply each bit by 16^n, where n is the "weight" of the bit.
  - The weight is the position of the bit, starting from 0 on the right.
  - Add the results

Example: $ABC_{16} = 2748_{10}$

$$
\begin{aligned}
&= AX16^2 + BX16^1 + CX16^0 \\
&= 10X16^2 + 11X16^1 + 12X16^0 \\
&= 2560 + 176 + 12 \\
&= 2748_{10}
\end{aligned}
$$

Numbers Systems

Hexa-decimal Number System

- Hexa-decimal to binary Conversion Techniques:
  - Convert hexa-decimal digit in a 4 bits, starting on the right side.

Example: $10AF_{16} = 0001000010101111_2$

$$= 1 \quad 0 \quad A \quad F$$

$$= 0001 \quad 0000 \quad 1010 \quad 1111$$

$$= 0001000010101111_2$$

Numbers Systems

Hexa-decimal Number System

- Hexa-decimal to octal Conversion Techniques:
    - Use binary as an intermediary

Example: $1F0C_{16} = 0001111100001100_2 = 17414_8$

$$
\begin{aligned}
&= 1 \quad F \quad 0 \quad C \\
&= 0001 \quad 1111 \quad 0000 \quad 1100 \\
&= 0 \quad 001 \quad 111 \quad 100 \quad 001 \quad 100 \\
&= 0 \quad 1 \quad 7 \quad 4 \quad 1 \quad 4 \\
&= 017414_8
\end{aligned}
$$

Round-off errors

Round-off errors

- No matter how large a computer is, it still has a limited amount of storage.
- Consider the result of dividing 2 by 3.

  0.666666 is a repeating number.
- Regardless of how many bits we use to store this number, it will get “cut off” at some point. No computer can accurately store this number

Truncation

- To truncate a number means to simply ignore the extra digits that the computer cannot store.
- Truncate the following to 3 significant digits after decimal point

0.2349

$\Downarrow$

0.234

Round-off errors

Truncation

- Truncate the following to 5 significant digits after (.)

0.666666666666 ➡ 0.66666

- Truncate the following to 8 significant binary digits after (.)

0.1010101111000101 ➡ 0.10101011

\$10.773	→	\$10.77
\$9.766	→	\$9.76
\$10.799	→	\$10.79
\$31.338		\$31.32

Rounding

- An alternative to truncation is rounding, where the last digit is "adjusted" to give a more accurate representation of the number.

Round-off errors

Rounding

- Round the following to 1 significant digits after (.)

2.53

2.53

2.5

Since 3 is less than 5, this will be like truncation

- Round the following to 2 significant digits after (.)

17.948

17.95

Since 8 is greater than 4, this will be like truncation

$10.773 → $10.77

$9.766 → $9.77

$10.799 → $10.8

$31.338 $31.34

Root Finding

1. Bisection method

2. Fixed-point iteration method

3. Newton's method

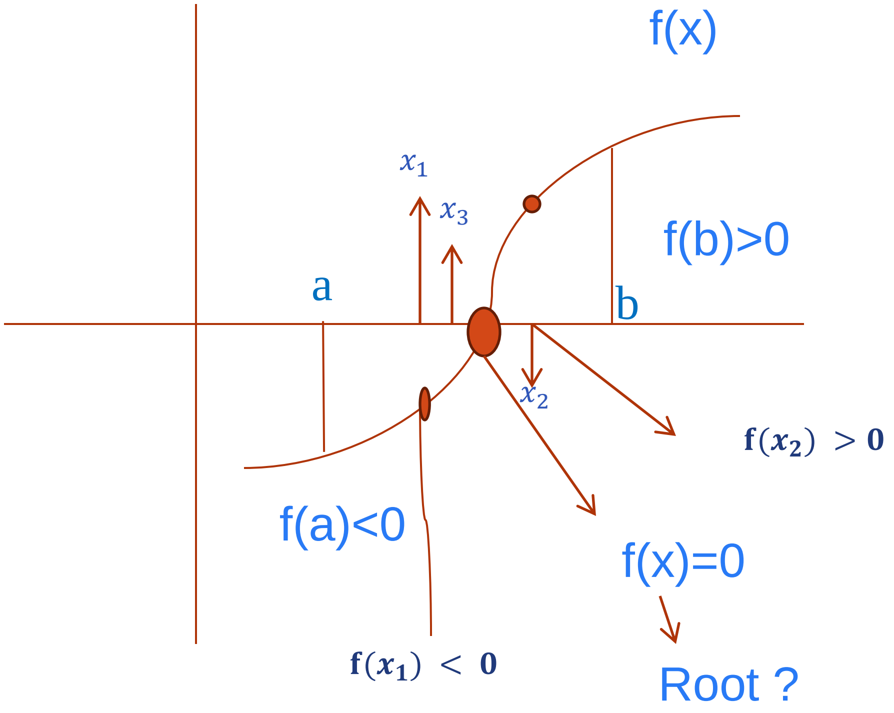
Bisection Method

- A root or solution of equation f(x)=0 are the values of x for which the equation holds true. Sometimes roots of equations are called the zeros of the equation.

- The bisection method in mathematics is a root-finding method that repeatedly bisects an interval and then selects a subinterval in which a root must lie for further processing.
- Also known as method: binary Court, Partition, and Bolzano method

Bisection method

Concept

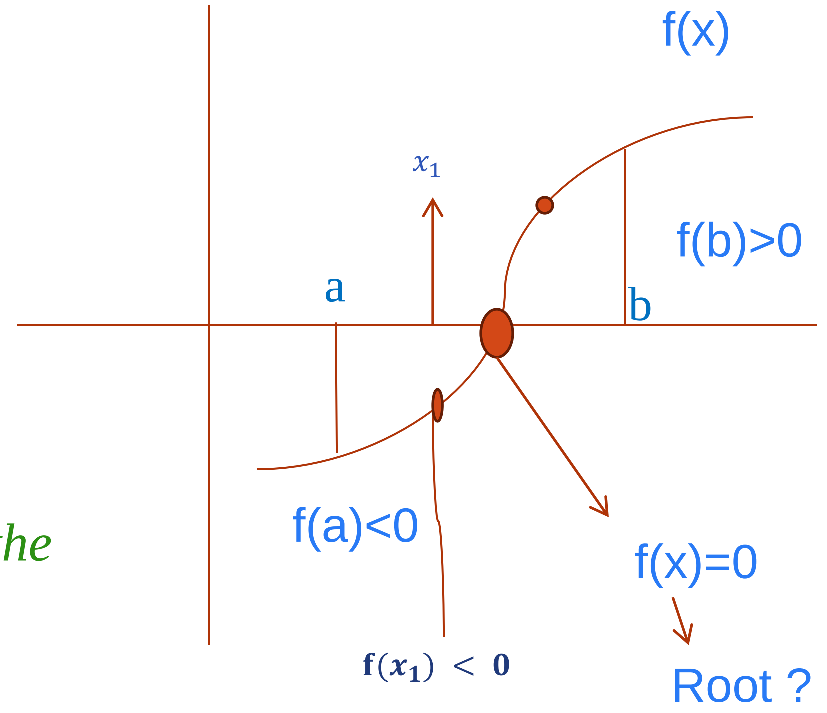


$$x_1 = \frac{a+b}{2}$$

$$x_2 = \frac{x_1+b}{2}$$

$$x_3 = \frac{x_1+x_2}{2}$$

Bisection Method



Rules:

Let f(x)=0 be given equation

Step 1: Find a and b (f(a)<0 & F(b)>0)

Step 2: Identify the first approximation of the root using bisection method

$$x_1 = \frac{a+b}{2}$$

Compute $f(x_1)$ *and examine its sign*

Step 2.1: if $f(x_1) < 0$, *this means root lies between* x_1 *and B, and the root of 2nd approximation can be computed by*

$$x_2 = \frac{x_1+b}{2}$$

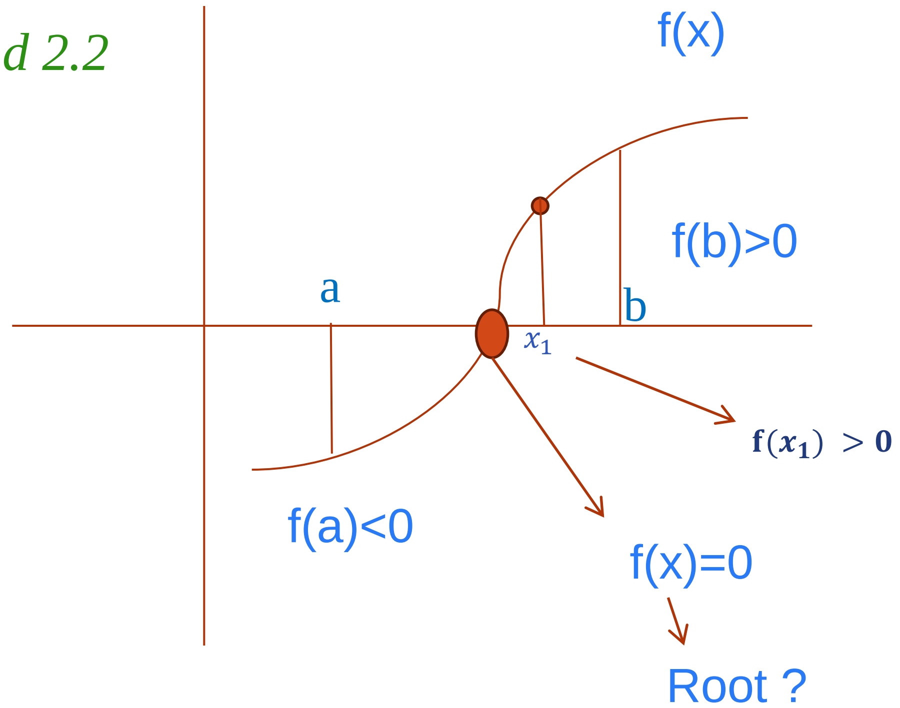
Bisection Method

Rules:

Step 2.2: if $\mathbf{f}(x_1) > 0$ *, this means root lies between* a *and* x_1*, and the root of 2nd approximation can be computed by*

$$x_2 = \frac{a + x_1}{2}$$

Compute $\mathbf{f}(x_2)$ *and repeat step 2.1 and 2.2*



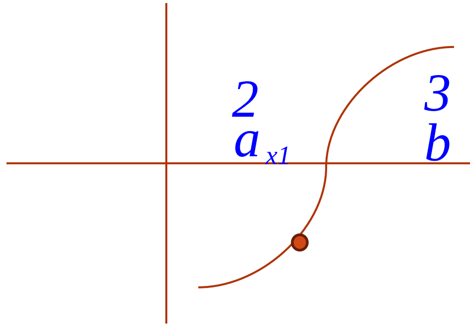
Bisection Method

Problem: Find the approximate root $\boldsymbol{x^3 - 4x - 9}$ by using bisection method, (1) In four stages and (2) Correct upto four decimal places.

Solution:

Step-by-step explanation:

Given: **Let** $\boldsymbol{F(x) = x^3 - 4x - 9}$



Then

$$F(0) = (0)^3 - 4(0) - 9 = -9$$

$$F(1) = (1)^3 - 4(1) - 9 = -12$$

$$F(2) = (2)^3 - 4(2) - 9 = -9 \rightarrow a$$

$$F(3) = (3)^3 - 4(3) - 9 = 6 \rightarrow b$$

Here f(2) is -ve and f(3) is positive. Therefore root lies between 2 and 3

$$x_1 = \frac{a+b}{2} = \frac{2+3}{2} = 2.5$$

Bisection Method

Hence the first approximation to the root is $x_1 = 2.5$

$$f(x_1) = (2.5)^3 - 4 * (2.5) - 9$$

$$f(x_1) = 15.625 - 10 - 9$$

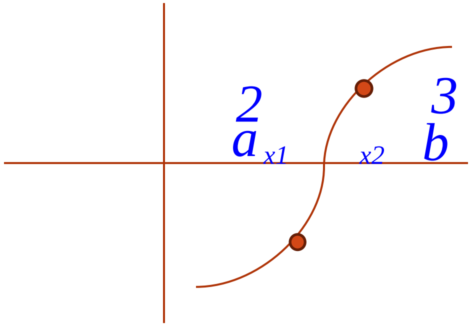
$\mathbf{f(x_1) < 0}$ *, this means root lies between* x_1 *and b,*

= -3.375

Root lies between 2.5 and 3

$$x_2 = \frac{x_1 + b}{2} = \frac{2.5 + 3}{2} = 2.75$$

$$x_2 = 2.75$$



$$f(x_2) = (2.75)^3 - 4 * (2.75) - 9$$

$$f(x_2) = 20.79 - 11 - 9$$

$$f(x_2) = 0.79$$

$\mathbf{f(x_2) > 0}$ *, this means root lies between* $\mathbf{x_1}$ *and* $\mathbf{x_2}$

Root lies between 2.5 and 2.75.

$$x_3 = \frac{x_1 + x_2}{2} = \frac{2.5 + 2.75}{2} = 2.625$$

$$x_3 = 2.625$$

Bisection Method

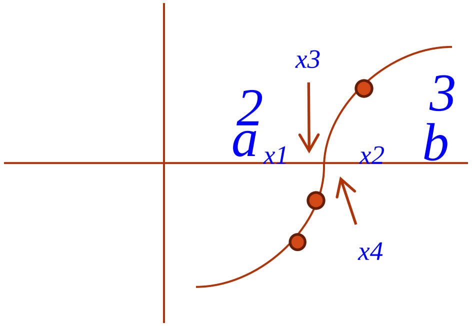
$$f(x_3) = (2.625)^3 - 4 * (2.625) - 9$$

$$f(x_3) = 18.09 - 10.5 - 9$$

$$f(x_3) = -1.41$$

$\mathbf{f}(x_3) < \mathbf{0}$, *this means root lies between* x_3 *and,* x_2

Root lies between 2.625 and 2.75.



$$x_4 = \frac{x_3 + x_2}{2} = \frac{2.625 + 2.75}{2} = 2.6875$$

$$x_4 = 2.6875$$

Hence the required root is $x_4 = 2.6875$ *Answer of first part*

$x_5 = 2.71875$
$x_6 = 2.7031$
$x_7 = 2.7109$
$x_8 = 2.707$
$x_9 = 2.7051$
$x_{10} = 2.7061$
$x_{11} = 2.7066$
$x_{12} = 2.7064$
$x_{13} = 2.7065$
$x_{14} = 2.7065$

Since $x_{13} = x_{14} = 2.7065$

Hence the root is 2.7065 and correct upto four decimal places

Bisection Method

Advantages:

- This method is always convergent. The convergence is guaranteed if a & b taken are such that

$$f(a) * f(b) < 0$$

- Error can be controlled. As iterations are conducted, the interval gets halved, so one can guarantee the error in the solution of the equation.

- It is also considered to be the safest method and simplest method.

Bisection Method

Disadvantages:

- Although convergence is guaranteed it is generally slow. The reason is every time the interval is getting half as we go from one iteration to another.

- Cannot find roots of some equations.

- Choosing a guess close to the root may result in need of many iterations to converge.

Thank You !