Numerical Analysis 1 : (2 nd year (CP2))

Dr :Kamel BENYETTOU

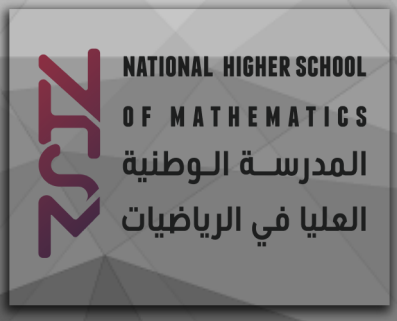
National Higher School of Mathematics,

Scientic and Technology Hub of Sidi Abdellah,

P.O. Box 75, Algier 16093, Algeria.

Email : kamel.benyettou@nhsm.edu.dz[1]

04/02/2024



1. file:///C:/Users/LENOVO/Documents/nhsm/2023/S2/formations/atelier%201%C3%A0%20reproduire%20-%20Download-20240202T163256Z-001/cour%20a%20reproduire/Cours%201%C3%A0%20reproduire%20version%20fran%C3%A7aise/Cours%201%C3%A0

Table des matières

Objectifs

The objectives of numerical analysis **((2 nd year CP 2 (Preparatory cycle Department))** are manifold and depend on the context in which it is used. However, here are some general objectives of numerical analysis:

- Solving complex mathematical problems: One of the main objectives of numerical analysis is to provide methods and tools for solving complex mathematical problems that cannot be solved analytically. This includes solving differential equations, approximating functions, finding the roots of equations, etc.
- Obtain accurate numerical solutions: Numerical analysis aims to obtain accurate numerical solutions that can be used to approximate analytical solutions or to solve practical problems in various fields such as engineering, physics, finance, etc.
- Studying the behaviour of dynamic systems: In many fields, it is necessary to study the behaviour of dynamic systems using mathematical models. Numerical analysis provides methods for simulating and studying the behaviour of these systems over time.
- Optimisation: Numerical analysis is used to solve optimisation problems, i.e. to find the optimal values of an objective function under given constraints. This is common in fields such as engineering, economics, data science, etc.
- Studying the stability and convergence of numerical methods: Another important objective of numerical analysis is to study the stability, accuracy and convergence of numerical methods used to solve mathematical problems. This ensures that the solutions obtained are reliable and accurate.
- Development of new numerical methods: Numerical analysis also involves the development and improvement of new numerical methods to solve specific problems more efficiently, accurately and quickly.

In short, the aims of numerical analysis are to provide methods and tools for solving complex mathematical problems accurately, efficiently and reliably, and to study the behaviour of dynamic systems using mathematical models.

In the other hand, the overall goal of the field of numerical analysis is the design and analysis of techniques to give approximate but accurate solutions to a wide variety of hard problems, many of which are infeasible to solve symbolically:

1. Computing the trajectory of a spacecraft requires the accurate numerical solution of a system of ordinary differential equations.
2. Advanced numerical methods are essential in making numerical weather prediction feasible.
3. Car companies can improve the crash safety of their vehicles by using computer simulations of car crashes. Such simulations essentially consist of solving partial differential equations[2] numerically.
4. Airlines use sophisticated optimization algorithms to decide ticket prices, airplane and crew assignments and fuel needs. Historically, such algorithms were developed within the overlapping field of operations research[3].

2. https://en.wikipedia.org/wiki/Partial_differential_equation
3. https://en.wikipedia.org/wiki/Operations_research

5. Car companies can improve the crash safety of their vehicles by using computer simulations of car crashes. Such simulations essentially consist of solving partial differential equations[4] numerically.
6. Insurance companies use numerical programs for actuarial[5] analysis.

4. https://en.wikipedia.org/wiki/Partial_differential_equation
5. https://en.wikipedia.org/wiki/Actuary

Introduction

Numerical Analysis refers to the process of in-depth analysis of algorithms and natural approximations. It is considered a part of both computer science and mathematical sciences. It is used in several disciplines like engineering, medicine, social science, etc. Because this branch is concerned with understanding, creating, and implementing the algorithms for common problems.

Some commonly used **numerical computation** methods include differential equations (that help predict planetary motion), linear algebra, etc.

Numerical analysis is a branch of mathematics that solves continuous problems using numeric approximation. It involves designing methods that give approximate but accurate numeric solutions, which is useful in cases where the exact solution is impossible or prohibitively expensive to calculate*Brezinski, C.; Zaglia, M.R. (2013). Extrapolation methods: theory and practice. Elsevier. ISBN 978-0-08-050622-7.*[Brezinski, C.; Zaglia, M.R. (2013). Extrapolation methods: theory and practice. Elsevier. ISBN 978-0-08-050622-7.]*. Numerical analysis also involves characterizing the convergence, accuracy, stability, and computational complexity of these methods.*Bultheel, Adhemar; Cools, Ronald, eds. (2010). The Birth of Numerical Analysis. Vol. 10. World Scientific. ISBN 978-981-283-625-0.*[Bultheel, Adhemar; Cools, Ronald, eds. (2010). The Birth of Numerical Analysis. Vol. 10. World Scientific. ISBN 978-981-283-625-0.]*

Numerical analysis can be divided into the following fields:*Brenner, S.; Scott, R. (2013). The mathematical theory of finite element methods (2nd ed.). Springer. ISBN 978-1-4757-3658-8.*[Brenner, S.; Scott, R. (2013). The mathematical theory of finite element methods (2nd ed.). Springer. ISBN 978-1-4757-3658-8.]*

(1) Numerical Solutions of Linear Algebraic Equations.

(2) Numerical Solutions of Nonlinear Algebraic Equations.

(3) Interpolation and Extrapolation.

(4) Approximation Theory and Curve Fitting.

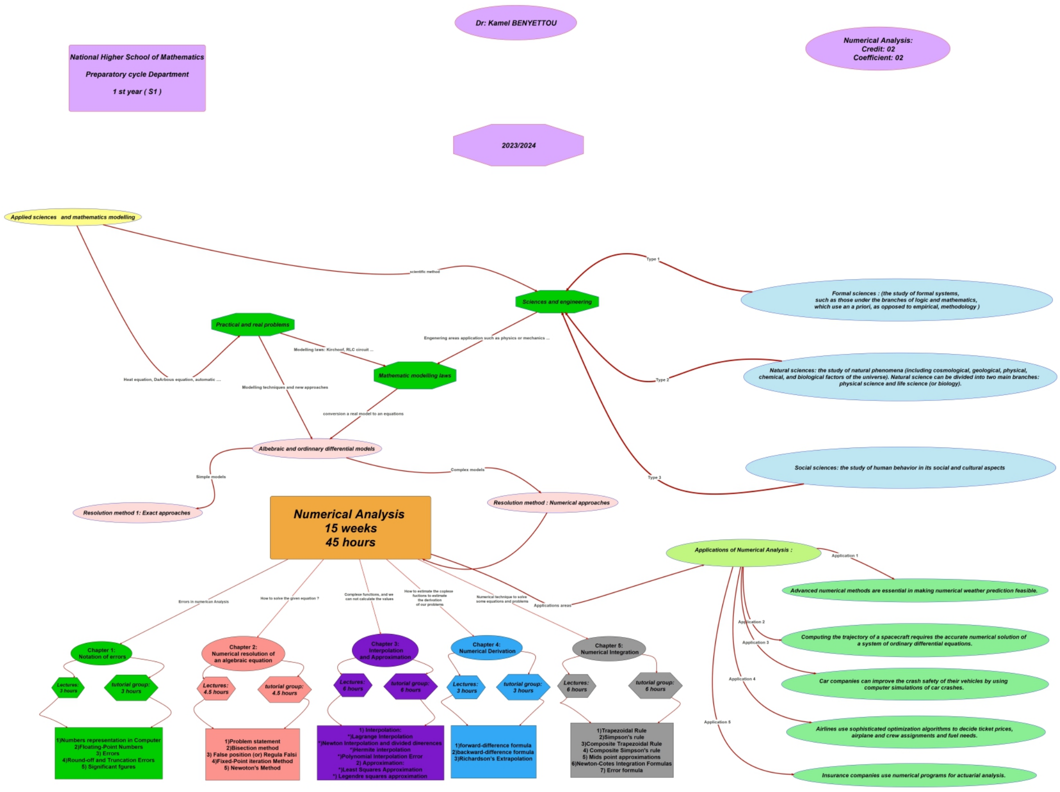
(5) Numerical Differentiation.

(6) Numerical Integration.

(7) Numerical Optimization.

(8) Numerical Solutions

One of the goals of numerical analysis is to compute answers within a specified level of accuracy.Working in double precision means that we store and operate on numbers that are kept to 52-bit accuracy, about 16 decimal digits.*Quarteroni, A.; Saleri, F.; Gervasio, P. (2014). Scientific computing with MATLAB and Octave (4th ed.). Springer. ISBN 978-3-642-45367-0.*[Quarteroni, A.; Saleri, F.; Gervasio, P. (2014). Scientific computing with MATLAB and Octave (4th ed.). Springer. ISBN 978-3-642-45367-0.]**Brezinski, C.; Zaglia, M.R. (2013). Extrapolation methods: theory and practice. Elsevier. ISBN 978-0-08-050622-7.*[Brezinski, C.; Zaglia, M.R. (2013). Extrapolation methods: theory and practice. Elsevier. ISBN 978-0-08-050622-7.]*



Dr: Kamel BENYETTOU
National Higher School of Mathematics
Preparatory cycle Department
1 st year (S1)
Numerical Analysis:
Credit: 02
Coefficient: 02
2023/2024
Applied sciences and mathematics modelling
scientific method
Sciences and engineering
Type 1
Formal sciences : (the study of formal systems, such as those under the branches of logic and mathematics, which use an a priori, as opposed to empirical, methodology)
Type 2
Natural sciences: the study of natural phenomena (including cosmological, geological, physical, chemical, and biological factors of the universe). Natural science can be divided into two main branches: physical science and life science (or biology).
Type 3
Social sciences: the study of human behavior in its social and cultural aspects
Practical and real problems
Engenering areas application such as physics or mechanics ...
Modelling laws: Kirchoof, RLC circuit ...
Mathematic modelling laws
Heat equation, DaArbous equation, automatic ...
Modelling techniques and new approaches
conversion a real model to an equations
Albebraic and ordinnary differential models
Simple models
Resolution method 1: Exact approaches
Complex models
Resolution method : Numerical approaches
Numerical Analysis
15 weeks
45 hours
Errors in numerican Analysis
How to solve the given equation ?
Complexe functions, and we can not calculate the values
How to estimate the coplexe fuctions to estimate the derivation of our problems
Numerical technique to solve some equations and problems
Applications areas
Chapter 1: Notation of errors
Lectures: 3 hours
tutorial group: 3 hours
1)Numbers representation in Computer
2)Floating-Point Numbers
3) Errors
4)Round-off and Truncation Errors
5) Significant fgures
Chapter 2: Numerical resolution of an algebraic equation
Lectures: 4.5 hours
tutorial group: 4.5 hours
1)Problem statement
2)Bisection method
3) False position (or) Regula Falsi
4)Fixed-Point iteration Method
5) Newoton's Method
Chapter 3: Interpolation and Approximation
Lectures: 6 hours
tutorial group: 6 hours
1) Interpolation:
*)Lagrange Interpolation
*)Newton Interpolation and divided dierences
*)Hermite interpolation
*)Polynomial Interpolation Error
2) Approximation:
*)Least Squares Approximation
*) Legendre squares approximation
Chapter 4: Numerical Derivation
Lectures: 3 hours
tutorial group: 3 hours
1)forward-difference formula
2)backward-difference formula
3)Richardson's Extrapolation
Chapter 5: Numerical Integration
Lectures: 6 hours
tutorial group: 6 hours
1)Trapezoidal Rule
2)Simpson's rule
3)Composite Trapezoidal Rule
4) Composite Simpson's rule
5) Mids point approximations
6)Newton-Cotes Integration Formulas
7) Error formula
Applications of Numerical Analysis :
Application 1
Advanced numerical methods are essential in making numerical weather prediction feasible.
Application 2
Computing the trajectory of a spacecraft requires the accurate numerical solution of a system of ordinary differential equations.
Application 3
Car companies can improve the crash safety of their vehicles by using computer simulations of car crashes.
Application 4
Airlines use sophisticated optimization algorithms to decide ticket prices, airplane and crew assignments and fuel needs.
Application 5
Insurance companies use numerical programs for actuarial analysis.

I Preconditions

1. Introduction

Peconditions (les prérequis)

2. Preconditions (Les prérequis)

The prerequisites for numerical analysis depend on your level of study and the specific field in which you wish to specialise. However, here are some useful general skills and prerequisite knowledge:

- Fundamental mathematics: A solid understanding of differential and integral calculus is essential. This includes knowledge of derivatives, integrals, series, ordinary and partial differential equations and convergence.
- Linear Algebra: A good understanding of linear algebra concepts such as vectors, matrices, vector spaces, linear transformations, eigenvalues and eigenvectors is required.
- Real analysis: Knowledge of real analysis concepts such as convergence of series, continuity, differentiability and integrability of functions is important.
- Basic Numerical Methods: You should have a basic understanding of common numerical methods such as solving non-linear equations, interpolation, function approximation, numerical integration and solving linear systems.
- Computer programming: Experience of programming, preferably in a language suitable for numerical analysis such as Python, MATLAB, or Julia, is very useful.
- Computer programming: Experience of programming, preferably in a language suitable for numerical analysis such as Python, MATLAB, or Julia, is very useful. You should be able to implement the numerical methods you learn and apply them to real-world problems.
- Error theory: Understanding how to assess and minimise errors in numerical calculations is crucial to obtaining accurate and reliable results.
- Application-specific knowledge: Depending on your specific field of application (engineering, physics, finance, etc.), you may need to acquire additional knowledge in this area to apply numerical analysis effectively.

II Chapter 1 : Notation of errors

1. Numbers representation in Computer

Digital computers serve as the primary tool for numerical analysis, making it crucial to comprehend their functioning. This section delves into the representation of numbers in computers and explores the implications of computerized number representation and arithmetic.

Most computers feature two modes for representing numbers: integer and floating-point. The integer mode is exclusively for integers and will not be further discussed here. The floating-point mode, however, is employed for representing real numbers. While the allowed numbers can vary greatly in size, there are constraints on both their magnitude and the number of digits they can contain. The representation of floating-point numbers closely resembles scientific notation, as found in many high school mathematics textbooks.

Decimal notation entails representing a number in the form of a fraction with the base 10, accompanied by a decimal point. It comprises digits ranging from 0 to 9, divided into two components: a whole number and a fractional part, with the decimal point serving as the separator between them.

Définition **: (Scientific Notation)**

Let k be a real number, then k can be written in the following form

$$k = m \times 10^n, 1 \leq m < 10$$

where m is any real number and the exponent n is a whole number (integer) is said to be in standard form is an integer. This notation is called the scientific notation or scientific form and sometimes referred to as standard form.

Exemple **: Scientific notation**

Express the following numbers in scientific notation :

Numbers
2
300
4321.768
-53000
6720000000
0.2
987
0.00000000751

The scientific notations :

Decimal notation	Scientific notation
2	2×10^0
300	3×10^2
4321.768	4.321768×10^3
-53000	-5.3×10^4
6720000000	6.72×10^9
0.2	2×10^{-1}
987	9.87×10^2
0.00000000751	7.51×10^{-9}

1.1. Floating-Point Numbers

In the decimal system any real number

$$a \neq 0$$

can be written in the decimal normalized oating-point form in the following way

$$a = \pm 0.\, d_1 d_2 d_3 \cdots d_k d_{k+1} d_{k+2} \cdots \times 10^n, 1 \leq d_1 \leq 9, 0 \leq d_i \leq 9,$$

for each i=2,..., and n is an integer called the exponent (ncan be positive, negative or zero). In computers we use a finite number of digits in representing the numbers and we obtain the following form

$$b = \pm 0.\, d_1 d_2 d_3 \cdots d_k \times 10^n, 1 \leq d_1 \leq 9, 0 \leq d_i \leq 9,$$

for each i=2, ..., k. These numbers are called k-digits decimal machine numbers.

2. Errors

Errors are inevitable in the realm of scientific computing. However, numerical analysts dedicate their efforts to exploring potential and optimal methods for error minimization. The examination, estimation, and mitigation of errors constitute fundamental aspects of error analysis.

2.1. Errors analysis

In numerical analysis, we approximate the exact solution of a problem using numerical methods, thereby inevitably introducing errors. The numerical error is defined as the difference between the exact solution and the approximate solution.

Définition : **(Numerical Error)**

Let x be the exact solution of the underlying problem and x^* its approximate solution, then the error (denoted by e) in solving this problem is

$$e = x - x^*$$

a) Sources of Error in Numerical Computations

Errors caused by human mistakes and oversights can be minimized through careful attention during scientific investigations. Such errors contribute to the overall error of the underlying problem and can significantly impact the accuracy of the solution.

- **Modeling Errors:** These errors emerge during the modeling process when scientists overlook influential factors in the model to simplify the problem. They are also known as formulation errors.
- **Data Uncertainty:** These errors stem from the uncertainty surrounding the physical problem data and are also referred to as data errors.
- **Discretization Errors:** Computers represent a function of continuous variables using a series of discrete values. Scientists also approximate and substitute complex continuous problems with discrete ones, resulting in discretization errors.

i) Absolute and Relative Errors

Absolute and Relative Errors

Définition **: Absolute Error**

The absolute error Δ_x of the error e is defined as the absolute value of the error e :

$$\Delta_x = |x - x^*|$$

Définition **: Relative Error**

The relative error δ_x of the error e is defined as the ratio between the absolute error Δ_x and the absolute value of the exact solution x :

$$\delta_x = \frac{\Delta_x}{|x|}, x \neq 0$$

Absolute and relative errors are two main types of measurement errors. There are some major differences between these two, which are given below.

Subject	Absolute Error	Relative Error
Definition	The difference between the actual valueand the measured value of a quantity iscalled absolute error.	The ratio of absolute error of ameasurement and the actual value ofthe quantity is known as a relativeerror.
Determination	It determines how large the error is.	It determines how good or bad the error is.

Size of the quantity	It varies depending on the size of thequantity.	It doesn't depend on the size of thequantity.

Operations	Absolute error	Relative error
$x \pm y$	$\Delta_x + \Delta_y$	$\frac{\Delta_x+\Delta_y}{\lvert x \pm y \rvert}$
$x \times y$	$\lvert x \rvert \Delta_y + \lvert y \rvert \Delta_x$	$\delta_x + \delta_y$
$\frac{x}{y}, y \neq 0$	$\frac{\lvert x \rvert \Delta_y + \lvert y \rvert \Delta_x}{y^2} = \frac{\Delta_{x \times y}}{y^2}$	$\delta_x + \delta_y$

3. Exercice : 1

Compute the absolute error and the relative error for : $a+b, a-b, a \times b, a \div b$ in the following cases :

- **Case (1)** $a = 2.89 \pm 0.2 \quad , \quad b = 3.01 \pm 0.4$
- **Case (2)** $a = 8.9 \pm 0.15 \quad , \quad b = 2.45 \pm 0.14$

Solution :

1. **Case (1) :** $a = 2.89 \pm 0.2, b = 3.01 \pm 0.4$:
   - **Addition :** $(a+b)$:

     Absolute Error: $\Delta_a + \Delta_b = 0.2 + 0.4 = 0.6$

     Relative Error: $\frac{\Delta_a+\Delta_b}{\lvert a+b \rvert} \approx \frac{0.6}{\lvert 2.89+3.01 \rvert} \approx 0.1034$
   - **Subtraction** : $(a-b)$

     Absolute Error: $\Delta_a + \Delta_b = 0.2 + 0.4 = 0.6$

     Relative Error: $\frac{\Delta_a+\Delta_b}{\lvert a-b \rvert} \approx \frac{0.6}{\lvert 2.89-3.01 \rvert} \approx 1.5$
   - **Multiplication** : $(a \times b)$

     Absolute Error: $\lvert a \rvert \Delta_b + \lvert b \rvert \Delta_a = \lvert 2.89 \rvert \cdot 0.4 + \lvert 3.01 \rvert \cdot 0.2 = 1.156 + 0.602 = 1.758$

     Relative Error: $\delta_a + \delta_b = \frac{\Delta_{a \times b}}{a \times b} \approx \frac{1.758}{2.89 \cdot 3.01} \approx 0.2031$
   - **Division** : $(a \div b)$

     Absolute Error: $\frac{\lvert a \rvert \Delta_b + \lvert b \rvert \Delta_a}{b^2} = \frac{1.156+0.602}{(3.01)^2} \approx 0.133$

     Relative Error: $\delta_a + \delta_b = \frac{\Delta_{a \times b}}{(a/b)^2} \approx \frac{1.758}{(2.89/3.01)^2} \approx 0.1987$
2. **Case (2)** : $a = 8.9 \pm 0.15 b = 2.45 \pm 0.14$
   - **Addition :** $(a+b)$:

     Absolute Error: $\Delta_{a+b} = 0.15 + 0.14 = 0.29$

     Relative Error: $\delta_{a+b} = \frac{\Delta_{a+b}}{\lvert a+b \rvert} \approx 0.02557$
   - **Subtraction** : $(a-b)$

     Absolute Error: $\Delta_{a \times b} = \lvert a \rvert \Delta_b + \lvert b \rvert \Delta_a = 1.246 + 0.3675 = 1.6135$

     Relative Error: $\delta_{a \times b} = 0.0157 + 0.0575 = 0.0732$
   - **Multiplication** : $(a \times b)$

Absolute Error: $\frac{|a|\Delta_b+|b|\Delta_a}{b^2} \approx 0.2674$

Relative Error: $\delta_{a\div b} = 0.0157 + 0.0575 = 0.0732$

- **Division** : $(a \div b)$

Absolute Error: $\frac{|a|\Delta_b+|b|\Delta_a}{b^2} = \frac{1.156+0.602}{(3.01)^2} \approx 0.133$

Relative Error: $\delta_a + \delta_b = \frac{\Delta_{a\times b}}{(a/b)^2} \approx \frac{1.758}{(2.89/3.01)^2} \approx 0.1987$

4. Round-off and Truncation Errors

4.1. Round-off error

Computers represent numbers with a finite number of digits, meaning that certain quantities cannot be represented precisely. The error that arises from substituting a number with its nearest machine number is termed the round-off error, and the procedure is referred to as correct rounding.

There are two common rounding rules, round-by-chop and round-to-nearest. The IEEE standard uses round-to-nearest.

1. **Round-by-chop**: The base- β expansion of x is truncated after the $(p-1)$-th digit.
    - This rounding rule is biased because it always moves the result toward zero.
2. **Round-to-nearest**: $fl(x$ is set to the nearest floating-point number to x. When there is a tie, the floating-point number whose last stored digit is even (also, the last digit, in binary form, is equal to 0) is used.
    - For IEEE standard where the base β is 2, this means when there is a tie it is rounded so that the last digit is equal to 0.
    - This rounding rule is more accurate but more computationally expensive.
    - Rounding so that the last stored digit is even when there is a tie ensures that it is not rounded up or down systematically. This is to try to avoid the possibility of an unwanted slow drift in long calculations due simply to a biased rounding.

Exemple **:**

- The following example illustrates the level of roundoff error under the two rounding rules. The rounding rule, round-to-nearest, leads to less roundoff error in general.

x	Round-by-chop	Roundoff Error	Round-to-nearest	Roundoff Error
1.649	1.6	0.049	1.6	0.049
1.650	1.6	0.050	1.6	0.050
1.651	1.6	0.051	1.7	−0.049
1.699	1.6	0.099	1.7	−0.001
1.749	1.7	0.049	1.7	0.049
1.750	1.7	0.050	1.8	−0.050

5. Truncation Errors

Truncation errors are the difference between the actual value of the function and the truncated value of the given function. The truncated value of the functions is the approximated value up to a given number of digits. For example, the speed of light in vacuum is 2.99792458×10^8 ms^{-1}. The truncated value up to two decimal places is 2.99×10^8. Hence the truncation error is the difference between 2.99792458×10^8 and 2.99×10^8, which is 0.00792458×10^8, or in scientific notation, it is 7.92458×10^5.

5.1. Truncation Errors

What is a truncation error?

A truncation error is the difference between an actual and a truncated, or cut-off, value. A truncated quantity is represented by a numeral with a fixed number of allowed digits, with any excess digits chopped off -- hence, the expression *truncated*.

Exemple :

Truncation error explained with Taylor series

In mathematical and computing applications, the true or analytical derivative or value of a function may be different from the value obtained by numerical approximation. The truncation error is the difference between these two values. It refers to the discrepancy that arises from executing a finite number of steps to approximate an infinite process -- a process known as *discretization*[6] -- usually for ease of calculation.

Infinite series

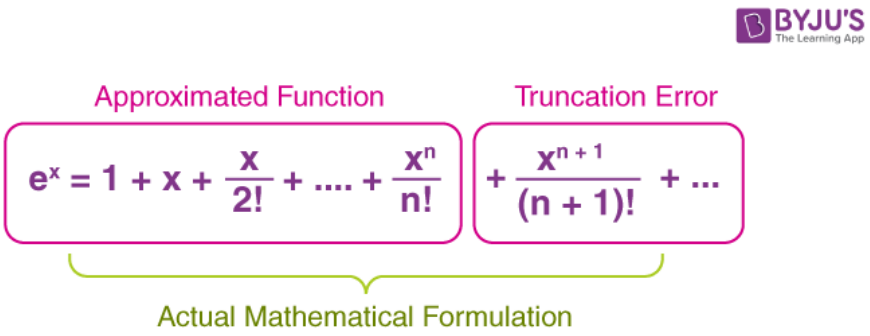
A summation series for e^x is given by an infinite series such as

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

In reality, we can only use a finite number of these terms as it would take an infinite amount of computational time to make use of all of them. So let's suppose we use only three terms of the series, then

$$e^x \approx 1 + x + \frac{x^2}{2!}$$

In this case, the truncation error is $\frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$



6. https://www.techtarget.com/searchenterpriseai/post/Wrangling-data-with-feature-discretization-standardization

Exemple **:**

Alternating Convergent Series Theorem

Maclaurin series of $\ln(1+x)$

$S = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n} \quad (-1 < x \leq 1)$

With $n = 5$,

$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} = 0.7833333340$

$\ln 2 = 0.693$

Error estimated using the althemating convergent the actual error series theorem

$|R| = |S - \ln 2| = 0.16666$

5.2. Estimation of Truncation Error for Geometric Series

Let S be an infinite geometric series if its terms are such that $|t_{j+1}| \leq k|t_j|$ where $0 \leq k \leq 1$ for every $j \geq n$, then while approximating the series up to n terms, the truncation error R_n is given by

$|R_n| = t_{n+1} + t_{n+2} + t_{n+3} + \ldots$

$\leq t_{n+1} + k|t_{n+1}| + k^2|t_{n+1}| + k^3|t_{n+1}| + \ldots$

$= |t_{n+1}|\left(1 + k^2 + k^3 + \ldots\right)$

$= |t_{n+1}|/(1-k)$

$\Rightarrow |R_n| \leq [k_n|t_n|/(1-k)$

For example, we have to calculate the truncation error $|R_6|$ of the infinite geometric series

$S = 1 + \frac{1}{\pi^2} + \frac{\sqrt{2}}{\pi^4} + \frac{\sqrt{3}}{\pi^6} + \ldots + \frac{\sqrt{j}}{\pi^{2j}} + \ldots$

Clearly,

$t_j = \frac{\sqrt{j}}{\pi^{2j}}$

Now, we to find k such that $|t_{j+1}| \leq k|t_j|$ where Now, we to find k such that $|t_{j+1}| \leq k|t_j|$ where $0 \leq k \leq 1$ for every $j \geq n(n=6) \leq k \leq 1$ for every $j \geq n(n=6)$

$|t_{j+1}| \leq k|t_j| = |t_{j+1}|/|t_j| \leq k = \sqrt{(1+1/j)}\pi^{-2}$

$\Rightarrow |t_{j+1}|/|t_j| \leq \sqrt{(1+1/6)}\pi^{-2} < 0.11 \text{ as } j \geq 6$

Therefore, by $|R_n| \leq [k|t_n|]/(1-k)\$and\$t_6 < 3 \times 10^{-6}$

$|R_6| \leq [k|t_n|]/(1-k) < \left[0.11 \times 3 \times 10^{-6}\right]/(1-0.11)$

6. Significant figures

Significant figures, also referred to as **significant digits** or **sig figs**, are specific digits[7] within a number written in positional notation[8] that carry both reliability and necessity in conveying a particular quantity. When presenting the outcome of a measurement (such as length, pressure, volume, or mass), if the number of digits exceeds what the measurement instrument can resolve, only the number of digits within the resolution[9]'s capability are dependable and therefore considered significant.

7. https://en.wikipedia.org/wiki/Numerical_digit
8. https://en.wikipedia.org/wiki/Positional_notation
9. https://en.wikipedia.org/wiki/Measurement_resolution

6.1. Rules to identify significant figures in a number

Note that identifying the significant figures in a number requires knowing which digits are reliable (e.g., by knowing the measurement or reporting resolution with which the number is obtained or processed) since only reliable digits can be significant; e.g., 3 and 4 in 0.00234 g are not significant if the measurable smallest weight is 0.001 g.

- **Non-zero digits within the given measurement or reporting resolution** are **significant**.
    - 91 has two significant figures (9 and 1) if they are measurement-allowed digits.
    - 123.45 has five significant digits (1, 2, 3, 4 and 5) if they are within the measurement resolution. If the resolution is 0.1, then the last digit 5 is not significant.
- **Zeros between two significant non-zero digits** are **significant (*significant trapped zeros)***.
    - 101.12003 consists of eight significant figures if the resolution is to 0.00001.
    - 125.340006 has seven significant figures if the resolution is to 0.0001: 1, 2, 5, 3, 4, 0, and 0.
- **Zeros to the left of the first non-zero digit** (leading zeros[10]) are **not significant**.
    - If a length measurement gives 0.052 km, then 0.052 km = 52 m so 5 and 2 are only significant; the leading zeros appear or disappear, depending on which unit is used, so they are not necessary to indicate the measurement scale.
    - 0.00034 has 2 significant figures (3 and 4) if the resolution is 0.00001.
- **Zeros to the right of the last non-zero digit (trailing zeros[11]) in a number with the decimal point** are **significant** if they are within the measurement or reporting resolution.
    - 1.200 has four significant figures (1, 2, 0, and 0) if they are allowed by the measurement resolution.
    - 0.0980 has three significant digits (9, 8, and the last zero) if they are within the measurement resolution.
    - 120.000 consists of six significant figures (1, 2, and the four subsequent zeroes) if, as before, they are within the measurement resolution.
- **Trailing zeros in an integer may or may not be significant**, depending on the measurement or reporting resolution.
    - 45,600 has 3, 4 or 5 significant figures depending on how the last zeros are used. For example, if the length of a road is reported as 45600 m without information about the reporting or measurement resolution, then it is not clear if the road length is precisely measured as 45600 m or if it is a rough estimate. If it is the rough estimation, then only the first three non-zero digits are significant since the trailing zeros are neither reliable nor necessary; 45600 m can be expressed as 45.6 km or as 4.56×10^4 m in scientific notation[12], and neither expression requires the trailing zeros.
- **An exact number has an infinite number of significant figures.**

10. https://en.wikipedia.org/wiki/Leading_zero
11. https://en.wikipedia.org/wiki/Trailing_zero
12. https://en.wikipedia.org/wiki/Scientific_notation

- If the number of apples in a bag is 4 (exact number), then this number is 4.0000... (with infinite trailing zeros to the right of the decimal point). As a result, 4 does not impact the number of significant figures or digits in the result of calculations with it.

- **A mathematical or physical constant has significant figures to its known digits.**
  - π is a specific real number[13] with several equivalent definitions. All of the digits in its exact decimal expansion 3.14159265358979323... are significant. Although many properties of these digits are known — for example, they do not repeat, because π is irrational — not all of the digits are known. As of 19 August 2021, more than 62 trillion digits[4] have been calculated. A 62 trillion-digit approximation has 62 trillion significant digits. In practical applications, far fewer digits are used. The everyday approximation 3.14 has three significant figures and 7 correct binary[14] digits. The approximation 22/7 has the same three correct decimal digits but has 10 correct binary digits. Most calculators and computer programs can handle the 16-digit expansion 3.141592653589793, which is sufficient for interplanetary navigation calculations.[5]
  - The Planck constant[15] is $h = 6.62607015 \times 10 - 34J \cdot s$ and is defined as an exact value so that it is more properly defined as $h = 6.62607015(0) \times 10 - 34J \cdot s$.

13. https://en.wikipedia.org/wiki/Real_number
14. https://en.wikipedia.org/wiki/Binary_number
15. https://en.wikipedia.org/wiki/Planck_constant

Références

Brenner, S.; Scott, R. (2013). The mathematical theory of finite element methods (2nd ed.). Springer. ISBN 978-1-4757-3658-8. Book

Brezinski, C.; Zaglia, M.R. (2013). Extrapolation methods: theory and practice. Elsevier. ISBN 978-0-08-050622-7. Book

Bultheel, Adhemar; Cools, Ronald, eds. (2010). The Birth of Numerical Analysis. Vol. 10. World Scientific. ISBN 978-981-283-625-0. Book

Quarteroni, A.; Saleri, F.; Gervasio, P. (2014). Scientific computing with MATLAB and Octave (4th ed.). Springer. ISBN 978-3-642-45367-0. Book