**NUMERICAL Methods**

# Nature of numerical computation

* **Number representation and numerical errors**
* Binary vs decimal representation
* Convert decimal to binary
* Floating point representation
* Only a finite number of bit is used to store the mantissa hence we will have **round off error**
* This is mitigated by the increase in the # bits used to store numbers
* Occurs because computers cannot represent every number on the real line perfectly
* It is good practice to avoid subtracting nearly equal numbers
  + - * Other type of numerical error **is truncation error**
* Occurs whenever we replace an infinite sum with a finite sum
* Conditioning vs. stability
* error may accumulate within the steps of an algorithm ie **error propagation**
* Hence the need to analyze algorithm in term of their stability as well
* Stability is a property of a specific algorithm
* **Conditioning** related to the difficulty of solving a problem per say (see polynomial example)

Let :

* An algorithm is stable if is of the same order of magnitude as machine precision
* Condition number K of a problem

The n

* **Order of convergence**
* Iterative algorithms uses a sequence of approximations
* Given an approximation some transformation is applied to it to get to get a better approximation
* Two things to consider
* The sequence of approximation converge to the real solution
* The convergence is hopefully fast
* The speed of convergence can be quantified by a rate
* Linear convergence :
* Quadratic convergence:
* Some points about iterative methods
* Some iterative methods may not converge
* Sometimes convergence depends on the initial guess
* Complexity issues
* When using iterative methods sometimes we have no idea of the number iterations needed to get an satisfactory answer
* Sometimes a **direct method** can be used to know the exact number of steps needed which can help measure the complexity of the algorithm
* Polynomial complexity
* **Solving system of linear equations**
* Solving by matrix inversion:
* This is quite inefficient in practice other methods may work better
* Methods for solving system of linear equations can be classified as
* Direct method:

1. Have a clearly defined complexity
2. Yield results within a given number of steps

* Iterative methods :

1. Build a sequence of solution that converge toward the solution
2. Number of steps is known a priory as it depends on convergence speed

* Reason why numerical difficulty may arise
* The matrix is singular or close to singular
* The problem conditioning is bad : need to consider the **condition number of the matrix**
* **Vector and matrix norm**
* The norm is a generalization of vector length in the Euclidean sense
* The norm can work with matrices, functions and is very useful to analyze
* Convergence
* Stability
* Conditioning issues
* Vector Norm: is a function mapping vector

1. ;
2. ;

* Matrix Norm: is a function mapping matrix such that

1. ;
2. ;
3. We say that a vector and matrix are compatible if