1. Error Analysis

Definition:

let x is a value, $\tilde{x}$ is a estimated value

(1) absolute error, $E_a = |x - \tilde{x}|$

(2) relation error, $E_r = |\frac{x - \tilde{x}}{x}|$

(3) percentage error, $E_p = 100 \times |\frac{x - \tilde{x}}{x}|$

$\exists \epsilon > 0, |x - \tilde{x}| < \epsilon$, Then ϵ is upper limit of the absolute error measures the absolute accuracy.

1.1. Error in Implementation of Numerical Methods.

(1) Round-off Error

(2) Overflow & Underflow

(3) Floating Point Arithmetic and Error Propagation

(4) Truncation Error

(5) Machine eps (Epsilon)

(3)Floating Point Arithmetic and Error Propagation.

Let x_1, x_2 are values, E_1, E_2 are error of x_1, x_2,We want to check the change of error in $"+", "-", "*", "/"$
"+"

Let $x = x_1 + x_2$, error of x is E
Then $x + E = x_1 + x_2 + E_1 + E_2 \implies E = E_1 + E_2$
by triangle inequality
Absolute Error $= |E| \leq |E_1| + |E_2|$
Relative Error $= \frac{|E|}{|x|} \leq \frac{|E_1|}{|x|} + \frac{|E_2|}{|x|}$

"-"(Similar "+")

"*"

Let $x = x_1 * x_2$
Then $x + E = (x_1 + E_1)(x_2 + E_2) = x_1x_2 + E_2x_1 + E_1x_2 + E_1E_2$
Absolute Error $= |E| \leq |x_2E_1| + |x_1E_2|$
Relative Error $= \dfrac{|E_1|}{|x|} \leq \dfrac{|E_1|}{|x_1|} + \dfrac{|E_2|}{|x_2|}$

"/"

Let $x = x_1/x_2$
$x + E_x = \dfrac{x_1 + E_1}{x_2 + E_2}\left(\dfrac{x_2 - E_2}{x_2 - E_2}\right) = \dfrac{x_1x_2 + E_1x_2 - x_1E_2}{x_2^2 - E_2^2} + E_1E_2$
Absolute Error $= |E_x| = |\dfrac{E_1x_2 - x_1E_2}{x_2^2}| \leq \dfrac{|E_1|}{|x_2|} + \dfrac{|x_1E_2|}{x_2^2}$
Relative Error $= \dfrac{|E_x|}{|x|} \leq \dfrac{|E_1|}{|x_1|} + \dfrac{|E_2|}{|x_2|}$

(4)Truncation Error. Cause by approximation infinite with its finite terms.
Use Taylor series ($f(x) \in P(C)$) as example
Let $x = a, f(x) = f(a) + f'(a)(x-a) + f''(a)\dfrac{(x-a)^2}{2!} + \cdots + \dfrac{(x-a)^n}{n!}f^n(a) + \cdots + Rn$
$Rn = \displaystyle\int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$

Thm 1(First Mean Value Theorem)

If g is continuous on $[a, x]$, then $\exists\ \xi$ between a and x s.t.

$$\int_a^x g(t)\ dt = g(\xi)(x-a)$$

Thm 2(Second Mean Value Theorem)

If g, h is differentiable and integrable on $[a, x]$, h does not change sign on $[a, x]$
then $\exists\ \xi$ that $a \leq \xi \leq x$ s.t.

$$\int_a^x g(t)h(t)\ dt = g(\xi)\int_a^x h(t)\ dt$$

since $t \in [a, x], h(t) = (x-t)^n \dfrac{1}{n!},\ f^{(n+1)}(t)$ is continuous
$\exists\ \xi \in [a, x],\ R_n = \dfrac{f^{(n+1)}(\xi)}{(n+1)!} f^{(x+1)}(\xi),\ \xi \in [a, a+h]$
(Ref. Violin page:799)
since power series convergent, $R_n(x) \to 0,\ as_n \to \infty$

Definition

Given $\{\, a_n \,\}\{\, b_n \,\}\,,\ b_n \geq 0,\ \forall n \geq 1$
$a_n = O(b_n)$ if $\exists\ M > 0 \rightarrow |a_w| \leq Mb_n\ \forall\ n \geq 1$
$R_n(x) = O(h^{n+1})$

1.2. Condition & Stability.

Condition number is sensitivit of the function
Stability is used to describe the sensitity of the process
Condition number of the $f(n)$

$$\text{CN} = \frac{|\dfrac{f(x) - f(\tilde{x})}{x - \tilde{x}}|}{|\dfrac{x - \tilde{x}}{x}|} = |\frac{f(x) - f(\tilde{x})}{x - \tilde{x}}| \cdot |\frac{x}{f(x)}| = |\frac{x}{f(x)} \cdot f'(x)|$$

by Mean Value Theorem,

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

when CN ≤ 1 is **well condition**, other is **ill condition**
when the function is more sensitive to change, the condition number will be more big.

2. Methods for $f(x) = 0$

we have four way to deal this problem

(1) Direct analytical Method

(2) Graphical

(3) Trial and Error Method

(4) Iterative Method

Thm. 3(Mean Value Theorem)

Let f be a continuous function on $[a, b] = I$(connected),
if $f(a) \leq c \leq f(b)$ that $\exists\ \xi \in [a, b] \to f(\xi) = c$

Corollary

Let f be a continuous function on $[a, b] = I$(connected)
i.e. $f(a) \cdot f(b) < 0\ \ni\ \exists c \in (a, b)\ \ni\ f(c) = 0$
c is a root of f(t)

Iterative Method

2.1. Bisection Method.

Let a, b be fixed satisfying Thm.3
$\therefore f(a) \cdot f(b) < 0, f$ is continuous on $[a, b]$. The first approximation is $x_0 = \dfrac{a+b}{2}$
if $f(a) \cdot f(x_0) \leq 0$, then By Thm. 3 the root will lie on (a, x_0) and $x_1 = \dfrac{a + x_0}{2}$
continue the process, let $x_{n-3}, x_{n-2}, x_{n-1}$ be same step, then nth approximation
if $f(x_n - 1) \cdot f(x_{n-3}) \leq 0$, then $x_n = \dfrac{x_{n-1} + x_{n-2}}{2}$
else $f(x_n - 1) \cdot f(x_{n-3}) \geq 0$, then $x_n = \dfrac{x_{n-1} + x_{n-3}}{2}$
we shall label the interval by algorithm

$$[a, b] = [a_0, b_0], [a_1, b_1][a_2, b_2], \cdots$$

by construction $b_n a_n = \dfrac{1}{2}(b_{n-1} - a_{n-1})$, Hence $b_n - a_n = \dfrac{1}{2^n}[b_0 - a_0],\ \forall n \geq 1$
Clearly $a_0 \leq a_1 \leq \cdots \leq b, b_0 \geq b_1 \geq \cdots \geq a, \{\, a_n \,\}, \{\, b_n \,\}$ is bdd and monotonic

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = f(r)$$

by assumption $f(a_n)f(b_n) < 0, \lim\limits_{n\to\infty} f(a_n) = f(\lim\limits_{n\to\infty} a_n) = f(r)$
$\therefore f(b_n) = f(r), 0 \leq [f(r)]^2 \leq 0 \implies f(r) = 0$
The process is called **nested internal property**

Let $\{ C_k \}_{k=1}^{\infty}$ is a $\downarrow$ sequence of nonempty closed compact subset of X, then $\cap k \subset k \neq \varnothing$ if $c_k \to 0$, then $\cap_k c_k = \{ r \}$

Let ξ be the solution $f(x) = 0$, then $\{ x_0 - \xi \} \leq \frac{b-a}{2}, \cdots, \{ x_n - \xi \} \leq \frac{b-a}{2^{n+1}}$

Definition(p-order-convergence)

$\{ x_n \}$: seq, $x_n \to z, s_n \to \infty$, define $\epsilon_n = z - x_n$, if $\exists c > 0, p \geq 1$

$$\lim_{n\to\infty} \frac{|\epsilon_{n+1}|}{|\epsilon_n|^p} = c$$

we call $\{ x_n \}$ is p order convergence
if $c \leq 1$, then it's good(only check this when it's a first order convergence)

Let ϵ_n be the error i.e. $\epsilon_n = |x_n - \xi|$, $\epsilon_n \leq \frac{b-a}{2^{n+1}} \leq \epsilon$, i.e. $h \geq \frac{ln(b-a) - ln\epsilon}{ln2} - 1$

$\epsilon_n = |x_n - \xi| \leq \frac{1}{2}(\frac{b-a}{2^n}) \approx \frac{1}{2}\epsilon_n - 1 \implies \lim_{n\to\infty} |\frac{\epsilon_n}{\epsilon_n - 1}| = \frac{1}{2}$

Then Bisection Method is first order convergence

2.2. Newton-Taphson Method.

observation:

Let x_0 be an initial approximate to the root of $f(x) = 0$, then $x_0 + h$ is the exact root of $f(x) = 0$, i.e. $f(x_0 + h) = 0$, from Taylor series, $f(x_0 + h) = f(x_0) + h \cdot f(x_0) + \cdots$
i.e. $x_0 \approx x_0 + h$

the first order approximation, $f(x_0 + h) = f(x_0) + h \cdot f'(x_0) = 0 \implies h = \frac{-f(x_0)}{f'(x_0)}$

Let $x_1 = x_0 + h$ be the next approximation to the root, $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

Ingeneral $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ $\forall n \geq 1$

Example

Consider the $f(x) = x^2 - M = 0 (M > 0)$

$$x_{n+1} = x_n - \frac{x_n^2 - M}{2x_n} = \frac{1}{2}(x_n + \frac{M}{x_n})(\star)$$

Ingeneral, also can obtain for the kth root of M, i.e. $\sqrt[k]{M}$ with $f(x) = x^k - M = 0$
if $x_1 > \sqrt{M}$, and define $x_2, \cdots$ by the interaction formula $(\star)$, then
(1)$\{ x_n \}$ is $\downarrow$(trivial) (2)$\{ x_n \}$ is bounded above($x_{n+1} = \frac{1}{2}(x_n + \frac{M}{x_n}) \geq \sqrt{x_n(\frac{M}{x_n})} = \sqrt{M}$)
By (1)(2), $\lim_{n\to\infty} x_n = \sqrt{M}$ exists.

observation
let $(x_0, f(x_0))$ be any point on the curve
$y = f(x)$, then $y - f(x_0) = f'(x_0)(x - x_0)$
Thm. 4(The NR method is 2 order convergence)

Let x denote the exact value of the root of $f(x) = 0$
x_n, x_{n+1} be two approximation S to the exact root $a, (f(a) = 0)$
if $\epsilon_n, \epsilon_{n+1}$ corresponding error S, then $x_n = a + \epsilon_n, x_{n+1} = a + \epsilon_{n+1}$
by(NR)

$$
\begin{aligned}
a + \epsilon_{n+1} &= a + \epsilon_n - \frac{f(a - \epsilon)}{f'(a + \epsilon_n)} \\
\epsilon_{n+1} &= S_n - \frac{f(a) + \epsilon_n f'(a) + \frac{\epsilon_n^2}{2!} f''(a) + \cdots}{f'(a) + \epsilon f''(a) + \frac{\epsilon_n^2}{2!} f'''(a) + \cdots} \\
&= \epsilon_n - \frac{\epsilon_n \left(f'(a) + \epsilon f'(a) + \frac{\epsilon_n^2}{2!} f''(a) + \cdots \right)}{f'(a) + \epsilon_n f''(a) + \frac{\epsilon^2}{2!} f''(a) + \cdots} \\
&= \frac{\epsilon[f'(a) + \epsilon_n f''(a) + \frac{\epsilon_n^2}{2!} f(a) + \cdots - [f'(a) + \frac{\epsilon_n}{2!} f''(a) + \cdots]]}{f'(a) + \epsilon_n f''(a) + \frac{\epsilon_n^2}{2!} f'''(a) + \cdots} \\
&= \frac{\epsilon_n [\frac{\epsilon_n}{2} f'(a) + \frac{\epsilon_n^2}{3} f''(a) + \cdots]}{f'(a) + \epsilon_n f''(a) + \frac{\epsilon_n^2}{2!} f'''(a) + \cdots} \\
&= \frac{\epsilon_n^2 [\frac{1}{2} f'(a) + \frac{\epsilon_n}{3} f''(x) + \cdots]}{f'(a)[1 + \epsilon_n \frac{f''(a)}{f'(a)} + \frac{\epsilon_n^2}{2!} \frac{f'''(a)}{f''(a)} + \cdots]} \\
\implies \frac{e_{n+1}}{\epsilon_n^2} &= \frac{\frac{1}{2} f''(a) + \frac{\epsilon_n}{3} f'''(a) + \cdots}{f'(a)(1 + \epsilon_n \frac{f''(a)}{f'(a)} + \cdots)}
\end{aligned}
$$

$\lim_{n \to \infty} |\frac{\epsilon_{n+1}}{\epsilon^n}| > \frac{1}{2} |\frac{f''(a)}{f'(a)}| < +\infty$
Remark: if f(x) has double root S

3. Eigen Problem

3.1. Review eigenvalue & eigenvector.

$A \in M_{n\times n}(\mathrm{R}/C),\ AX = \lambda X = \lambda(IX) = \lambda IX \implies (A-\lambda I)X = 0$
it's a homogeneous system of n linear equation, it determinate is 0
$p(\lambda) = det(A-\lambda I) = 0,\ deg(p(\lambda)) = n$

Define $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, X is a eigen vector of A, λ is a eigenvalue of A

the normalized eigenvector $\hat{X} = \dfrac{1}{||X||}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ where $||X|| = (X^T X)^{\frac{1}{2}} = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}}$
if T is diagonalizable, then $\exists$ order basis β, $\beta \ni [T]_\beta = D$, which is a diagonal matrix
similarly A is diagonalizable if L_A is diagonalizable

diagonalizable

$$\begin{cases} \text{the c.p split} \begin{cases} n \text{ distinct eigenvalue} \\ \text{other} \begin{cases} \text{algebraic multiplicity} = \text{geometric multiplicity} \\ \text{algebraic multiplicity} \neq \text{geometric multiplicity}\textbf{(not diagonalizable)} \end{cases} \end{cases} \\ \text{the cp does not split (not diagonalizable)} \end{cases}$$

(c.p. is charateristic polynomial)

E_λ is subspace ,$E_\lambda = N(T-\lambda I)$, E_λ is T-invariant, i.e. $T(E_\lambda) \subseteq E_\lambda$,$1 \le \dim(E_\lambda) \le m$
if T is diagonalizable, then

$$V = E_{\lambda_1} \oplus E_{\lambda_2} + \cdots + E_{\lambda_n} \Leftrightarrow V = k\lambda_1 \oplus \cdots \oplus k\lambda_n$$

Let any eigenvalue λ be repeated r times with k linearly independent eigenvector
r is algebraic multiplicity, k is geometric multiplicity

3.2. some introduction.

we will learn ODE and PDE next time
$\frac{dX}{dt} = AX$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\frac{dx_1}{dt} = a_{11}x_1 + a_{12}$, $\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2$
$X = \chi e^{\lambda t}$ is the solution of system, χ is column vector, λ is parameter to be determind
$\frac{d\chi e^{\lambda t}}{dt} = \lambda\chi e^{\lambda t} \implies \lambda\chi e^{\lambda t} = A\chi e^{\lambda t} \implies \lambda\chi = A\chi$

Definition

The spectrum of A, radius p of the smallest circle with center at the origin and contains all the spactual radius

3.3. Power Method.

Definition

Let $A \in M_{n\times m}(\mathrm{C})$, for $1 \leq i, j \leq n$
define $p_i(A)$ to be the sum of the abs-values of the entries of row i of A and $r_i(A)$ to be the sum of the abs-values of the entries of column j of A
$p_i(A) = \sum_j^n ||A_ij||$, $r_j(A) = \sum_i^n ||A_ij||$
$e(A) = \max(p_i(A))$, $r(A) = \max(r_j(A))$, $1 \leq i, j \leq n$

Definition

an $n \times n$ matrix A, we define the ith Geisg disk c_i to be the disk in the complex plain with center A_{ii} an radius $r_i = p_i(A) - |A_{ii}|$, $c_i = \{\, z \in \mathrm{C} \mid |z - A_{ii}| < r_i \,\}$

Theorem(Geisg Disk Theorem 1)

Let $A \in M_{n\times n}(\mathrm{C})$, then every eigenvalue of A is contained in a Geisg Disk

pf: Let λ be eigenvalue of A r.t. eigenvector $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, clearly $Av = \lambda v$

Then $I_j^n = A_{ij}v_j = \lambda_{ri}$, $1 \leq i \leq n(\star)$
suppose v_k is the coordinate of V having the largest abs-solute,($v_k \neq 0$)
claim $\lambda \in C_k$, i.e. $|\lambda - A_{kk}| \leq r_k$ For $i = k$,by $(\star)$

$$\begin{aligned} |\lambda v_k - A_{kk}v_k| &= |\sum_{j=1}^{n} A_{kj}v_j - A_{kk}v_k| \\ &= |\sum_{j\neq k} A_{kj}v_j| \\ &\leq \sum_{j\neq k} |A_{kj}||v_j| \\ &\leq \sum_{j\neq k} |A_kj||v_k| = r_k|v_k| \end{aligned}$$

Corollary 1

Let λ be any eigenvalue of $A \in M_{n\times n}(\mathrm{C})$, then $|\lambda| \leq p(A) = \max(p_i(A))$
pf: by Thm. $|\lambda - A_{kk}| \leq r_k$ for some k, $1 \leq k \leq n$
$|\lambda| = |\lambda - A_{kk}| + |A_{kk}| \leq r_k + |A_{kk}| = p_k(A) \leq p(A)$

Corollary 2

$A^T \in M_{n\times n}(\mathrm{C})$, $|\lambda| \leq r(A) = \max(r_j(A))$

Corollary 3

Let λ be eigenvalue of $A \in M_{n\times n}(\mathrm{C})$, $|\lambda| \leq \min\{\, p(A), r(A)\,\}$
by corollary 1 & 2, we are done.

Theorem(Geisg Disk Theorem 2)

Let $A \in M_{n\times n}(\mathrm{C})$, k of the disks are disjoint from the others, then exactly k eigenvalue are contained in the union of these disks.
pf: the gumltprinciple
Ref:Matrix Analysis 2/e (Horn/Johnson) P.388,389

Rayleign Power Method

Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalue of matrix, $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$