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

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 $f(n)$

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

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 $\{ x_0 - \xi \} \leq \dfrac{b-a}{2}, \cdots, \{ x_n - \xi \} \leq \dfrac{b-a}{2^{n+1}}$

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

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

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, $f(x_0 + h) = f(x_0) + h \cdot f'(x_0) = 0 \implies h = \dfrac{-f(x_0)}{f'(x_0)}$

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

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

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($x_{n+1} = \dfrac{1}{2}(x_n + \dfrac{M}{x_n}) \geq \sqrt{x_n(\dfrac{M}{x_n})} = \sqrt{M}$)
By (1)(2), $\lim\limits_{n\to\infty} x_n = \sqrt{M}$ 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)

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

$\lim_{n\to\infty} |\frac{\epsilon_{n+1}}{\epsilon^n}| > \frac{1}{2} |\frac{f''(a)}{f'(a)}| < +\infty$
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 $||X|| = (X^TX)^{\frac{1}{2}} = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}}$
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. $T(E_\lambda) \subseteq E_\lambda$,$1 \leq \dim(E_\lambda) \leq m$
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, $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$
our goal is to find $|\lambda_1|$

Let $x_1, \cdots, x_n$ be eigenvectors, r.t. $\lambda_1, \cdots, \lambda_n, \implies Ax_i = \lambda_i x_i, \ \forall 1 \leq i \leq n$
if the matrix A(which is diagonalizable) has n linearly independent eigenvectors
then $x = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ for some $c_i \in$ C

$$\begin{aligned}
Ax &= A(c_1x_1 + \cdots + c_nx_n) \\
&= c_1Ax_1 + \cdots + c_nAx_n \\
&= c_1\lambda_1 x + \cdots + c_n\lambda_n x \\
&= \lambda_1\left(c_1x + c_2\left(\frac{\lambda_2}{\lambda_1}\right)x + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)\right)
\end{aligned}$$

$$\begin{aligned}
A^2x &= A\left(\lambda_1\left(c_1x + c_2\left(\frac{\lambda_2}{\lambda_1}\right)x + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)\right)\right) \\
&= \lambda\left(c_1Ax + c_2\left(\frac{\lambda_2}{\lambda_1}\right)Ax + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)Ax\right) \\
&= \lambda_1^2\left(ax + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^2 x + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^2 x\right)
\end{aligned}$$

Continue process

$$\begin{aligned}
A^kx &= \lambda_1^k\left(c_1x_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k x_n\right) \\
A^{k+1}x &= \lambda_1^{k+1}\left(c_1x_1 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} x_n\right) \\
&\lim_{k\to\infty}\frac{A^{k+1}x}{A^kx} = \lambda_1
\end{aligned}$$

A stepwise procedure

(i) $X^{(0)}$ is initial vector

(ii) $Y^{(0)} = AX^0$

(iii) $\lambda^{(1)}$ is the absolutely largest element, common from the vector $Y^{(0)}$
Let the remainly vector be X^1, $Y^(0) = \lambda^{(1)}X^{(1)}$

(iv) reapeating (ii) and (iii), $Y^{(k)} = \lambda^{(k+1)}X^{k+1}$

(v) $|\lambda^{(k+1)}|, x^{(k+1)}$ is goal

Example

$$A = \begin{pmatrix} 0 & 2 & 4 \\ 1 & 1 & -2 \\ -2 & 0 & 5 \end{pmatrix}, \quad X^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad Y^0 = AX^0 = \begin{pmatrix} 6 \\ 0 \\ 3 \end{pmatrix},$$

$$\lambda^{(1)} = 6,\ Y^{(0)} = 6\begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} = \lambda^{(1)}X^{(1)} \implies Y^{(1)} = AX^{(1)} = A^{(1)} = \begin{pmatrix} 2 \\ 0 \\ 0.5 \end{pmatrix}, \lambda^{(2)} = 2$$

Inverse Power Method

Let λ_i be an eigenvalue of matrix A, then $\dfrac{i}{\lambda_i}$ is eigenvalue of the matrix A^{-1},The eigenvector of A^{-1} is X_i
pf: $Ax_i = \lambda_i x_i \implies \dfrac{1}{\lambda_i}(Ax_i) = x_i \implies \dfrac{1}{\lambda_i}x_i = A^{-1}x_i$

Shifted Power Method

Let λ_i be an eigenvalue of matrix A, then $(\lambda_i - k)$ is an eigenvalue of the matrix $A - kI$ with the same eigenvector as that matrix A
pf: $Ax_i = \lambda_i x_i \implies (A - kI)x_i = AX_i - kX_i = \lambda_i x_i - kx_i = (\lambda_i - k)x_i$

4. Review Linear Algebra

4.1. Lagrange polynomials. Let $T : P_n(\mathrm{F} \to \mathrm{F}^{n_1}$ be linear transform defined by $T(f) = (f(c_0), \cdots, f(c_n))$, which $c_0, c_1, \cdots, c_n$ are distinct scalars in an infinite field F,β be the stander order basis for $P_n(\mathrm{F})$, γ be the stander order basis for F^{n+1}

Claim 1:

$$[T]_\beta^\gamma = M = \begin{bmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 1 & c_0 & c_1^2 & \cdots & c_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{bmatrix}, \beta = \{1, x, \cdots, x^n\}, \gamma = \{(1, \cdots, 0), \cdots, (0, \cdots, 1)\}$$

$T(1) = (1, \cdots, 1), T(x) = (c_0, \cdots, c_n), \cdots, T(x^n) = (c_0^n, \cdots, c_n^n)$

M is called a Vandemonde Matrix

Claim2: $\det(M) \neq 0$

$\because \dim(P_n(F)) = \dim(\mathrm{F}^{\mathrm{n}+1}) = n+1$, T is linear **check**,T is one-to-one **check**,

$\therefore T$ is invertible, $\therefore [T]_\beta^\gamma$ is invertible $\Leftrightarrow \det([T]_\beta^\gamma \neq 0 \implies \det(M) \neq 0$

Claim3: $\det(M) = \displaystyle\prod_{0 \le i < j \le n} (c_j - c_i)$

Proof. we use the induction on $n = \deg(P_n(\mathrm{F}))$

$n = 1, \det \begin{bmatrix} 1 & c_0 \\ 1 & c_1 \end{bmatrix} = c_1 - c_0$

Suppose the statement holds for n

$$\det \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{pmatrix} = \det \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 0 & c_1 - c_0 & c_1^2 - c_0^2 & \cdots & c_1^n - c_0^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & c_n - c_0 & c_n^2 - c_0^2 & \cdots & c_n^n - c_0^n \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & c_1 - c_0 & c_1^2 - c_1 c_0 & \cdots & c_1^n - c_0 c_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & c_n - c_0 & c_n^2 - c_0 c_n & \cdots & c_n^n - c_0 c_n^{n-1} \end{pmatrix} = \det \begin{pmatrix} c_1 - c_0 & c_1(c_1 - c_0) & \cdots & c_1^{n-1}(c_1 - c_0) \\ \vdots & \vdots & \vdots & \vdots \\ c_n - c_0 & c_n(c_n - c_0) & \cdots & c_n^{n-1}(c_n - c_0) \end{pmatrix}$$

$$= (c_1 - c_0) \cdots (c_n - c_0) \cdot \det \begin{pmatrix} 1 & c_1 & \cdots & c_n^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & c_n & \cdots & c_n^{n-1} \end{pmatrix} = (c_1 - c_0) \cdots (c_n - c_0) \prod_{1 \le i < j \le n} (c_j - c_i)$$

$$= \prod_{0 \le i < j \le n} (c_j - c_i)$$ ■

Let $P_n(X) = a_0 + a_1 x + \cdots + a_n x^n$, where $a_0, \cdots, a_n \in F$

$P_n(X)$ is a polynomial s.t. it interpolated the $n+1$ points

$P_n(x_0) = a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n = y_0$

$\vdots$

$P_n(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n = y_n$

In matrix form
$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \cdots & x_n^n & \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$$
Now we define the Lagrange polynomials of degree $l_0(x), \cdots, l_n(x)$ as

$$l_i(x_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

The $P_n(x) = y_0 l_0(x) + \cdots + y_n l_n(x) = \sum_{i=0}^{n} y_i l_i(x)$

$l_i(x)$ is an n-degree polynomial with roots says

$l_i(x) = c_i(x - x_0) \cdots (x - x_n) = c_i \prod_{j \neq i} (x - x_j)$

$l_i(x_i) = 1 = 1c_i \prod_{j \neq i} (x_i - x_j),\ c_i = \frac{1}{\prod_{j \neq i}(x_i - x_j)},\ l_i(x) = \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}$

4.2. special matrix.

Theorem (Shor's Lemma). *Let T is a linear operator on V which is a finite dimension inner product space, Suppose the characteristic polynomial splits, Then $\exists$ order normal basis $\beta \implies [T]_\beta$ is uppertriangle.*

Note

normal : $AA^* = A^*A(TT^* = T * T)$
self-adjoint : $A* = A(T^* = T)$

Theorem (Spectral Theorem). *Let T be a linear operator on V which is a finite dimensional inner product space*

$\mathbb{C}$: *T is normal $\Leftrightarrow \exists$ order normal basis β containing eigenvectors $\Leftrightarrow T$ is diagonal over $\mathbb{C}$*

$\mathbb{R}$: *T is self-adjoint $\Leftrightarrow \exists$ order normal basis β containing eigenvector $\Leftrightarrow T$ is diagonal over $\mathbb{R}$*

T is diagonal $\implies \exists$ order normal basis $\implies [T]\beta$ is diagonal,T is normal(over C)
$\implies [T]_\beta^*[T]_\beta = [T]_\beta[T]^* \implies [T^*]_\beta[T]_\beta = [T]_\beta[T^*]_\beta \implies [T^*T]_\beta = [TT^*]_\beta$

T is diagonalizable $\begin{cases} (C) \Leftrightarrow T \text{ is normal(unitary equivalent)(by Shur's Lemma)} \\ (R) \Leftrightarrow T \text{ is self-adjoint(orthogonal equivalent)(eigenvalue is real + Shur's)} \end{cases}$

Property

T is unitory $\Leftrightarrow$ every row(column) vectors is orthonormal basis
unitary equivalent $A \sim B \Leftrightarrow \exists$ unitary matrix $Q \implies A = Q^*BQ$
orthogonally equivalent $A \sim B \Leftrightarrow \exists$ orthogonal matrix $P \implies A = P^*BP$

Define: Let V be a vector space, $W_1, W_2 \leq V \implies V = W_1 \oplus W_2$

A function $T : V \to V$ is called projection on W_1 along W_2 if for $x = x_1 + x_2$, $x_1 \in W_1, x_2 \in W_2,\ T(x) = x_1$

Property

$R(T) = W_1,\ N(T) = W_2,\ V = R(T) \oplus N(T)$

Proof. Claim: $R(T) = W_1$

$(\supseteq) x \in W_1 \implies T(x) = x \in R(T)$

$(\subseteq) x \in R(T) \implies \exists y \in V \implies T(y) = x$

$\because V = W_1 \oplus W_2 \therefore y = x_1 + x_2$ for some $x_1 \in W_1,\ x_2 \in W_2 \implies T(y) = x_1 = x \in W_1$

Claim: $N(T) = W_2$ **exercise**

Claim: $V = R(T) \oplus N(T), \because$(1),(2), it's trivial ∎

Property

T is projection $\Leftrightarrow T = T^2$

Proof.

$(\Rightarrow)$ T is projection$(V = W_1 \oplus W_2)$

given $y \in V, \because V = W_1 \oplus W_2,\ \therefore \exists x_1 \in W_1, x_2 \in W_2 \ni y = x_1 + x_2$

$\therefore T(y) = T(x_1 + x_2) = x_1 = T(x_1) = T(T(y))$

$(\Leftarrow)$ $T = T^2$(use the previous proposition to build $R(T), N(T)$

$\implies V = R(T) \oplus N(T) \implies T$ is projection)

Given $x \in V$

(1) $x = T(x) + [x - T(x)]$

(2)$T^2(x) = T(x)$(assumption)

$$\begin{cases} (i) T(x) \in R(T) \\ (ii) x - T(x) \in N(T) \\ (iii) R(T) \cap N(T) = \{ 0 \} \\ (iv) V = R(T) + N(R) \end{cases}$$

$(i) T(T(x)) = T(x) \in R(T),\ (ii) T(x - T(x)) = T(x) - T(T(x)) = T(x) - T(x) = 0,$

(iv)trivial($\because$(1))

(iii) Suppose $N(T) \cap R(T) = \{ v \}, v \neq 0, v \in N(T) \implies T(v) = 0$

$v \in R(T) \implies \exists y \in V \implies T(y) = v \implies T(T(y)) = T(v) = 0, y \in N(T), v = 0 \rightarrow\leftarrow$

∎

Property: every projections is uniquely determined by the range & kernal

Let $T, U : V \to V, R(T) = R(U) = W_1,\ N(T) = N(U) = W_2$

$\forall x \in V$, let $y' = T(x), y = U(x) \in W_1$

$T(x - y') = T(x) - T(y') = y' - y' = 0,\ x - y' \in W_2 \implies x \in y' + W_2$(coset)

$\implies \exists z' \in W_2 \ni x = y' + z' \implies y = U(x) = U(y' + z') = U(y') + U(z') = y' + 0 = y' = T(x)$

Theorem. *orthogonal projection* T, $V = R(T) \oplus N(T)$, $R(T)^{\perp} = N(T)$, $N(T)^{\perp} = R(T)$
T *is orthogonal projection* $\Leftrightarrow T = T^2 = T^*$

Theorem. *Let matrix* A *normal*($\mathbb{C}$), *self-adjoint*($\mathbb{R}$)
A *is unitary equivalent to a diagonal matrix*
$u_1, \cdots, u_n$:*eigenvectors(orthonormal)*, $\lambda_1, \cdots, \lambda_n$:*eigenvalues*

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix} = A = \sum_{i=1}^{n} \lambda_i u_i u_i^T \textit{(spectral decomposition)}$$

Check A is normal
$A = u_i u_i^*, L_A = u_i u_i^*,\ L_A L_A = L_A^2 = (u_i u_i^*)(u_i u_i^*) = u_i u_i^* = L_A$
$L_A^* = (L_A)^* = (u_i u_i^*)^* = u_I^{**} u_i^* = u_i u_i^* = L_A$
example

$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$,c.p. of $A = (1-t)(-2-t) - 4 = t^2 + t - 6 = (t+3)(t-2)$
the eigenvector are $\begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
$\because$ they are distinct eigenvalue $\implies$ orthogonal
$r_1 = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} -1 \\ 2 \end{pmatrix}, r_2 = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
$\begin{bmatrix} r_1 & r_2 \end{bmatrix} \begin{bmatrix} -3 & - \\ 0 & 2 \end{bmatrix} \begin{bmatrix} r_1^T \\ r_2^T \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$
$A = -3 r_1 r_1^T + 2 r_2 r_2^T$