1 Important Background

[*Fundamental Theorem of Algebra*] Every non-zero, single-variable, degree n polynomial with complex coefficients has, counted with multiplicity, exactly n complex roots.

Application in matrix and eigenvalues: Every $n \times n$ matrix characteristic polynomial $C(\lambda)$ can be factor into linear factors as $C(\lambda) = (\lambda -_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$. Therefor, every $n \times n$ matrix has exactly n eigenvalues, counted with multiplicity.

2 Problem Statement

Every square complex matrix $A : \mathbb{C}^n \to \mathbb{C}^n$ can be reduced into Jordan Normal Form.

i.e. Every A can be expressed as $A = CA_JC^{-1}$, where A_J composed by Jordan Blocks, C changing basis.

3 Proof Structure

Method used here is to simply the matrix by choosing suitable new basis. When there are eigenvalues' multiplicity > 1 and eigenvectors not enough to make a basis, pull root vectors of repeated eigenvalues, basis of root subspaces to form a linearly independent set as basis.

The proof is consisted by three steps.

STEP 1. Every square complex matrix can be reduced into matrix A in upper triangular form.

(Schur decomposition, skip in this proof)

STEP 2. Upper triangular A can be expressed as

$$A = \begin{bmatrix} A|V^{\lambda_1} & & & 0 \\ & A|V^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & A|V^{\lambda_k} \end{bmatrix}$$

, where $A|V^{\lambda_i}$ is restricted A on root subspace V^{λ_i} of each distinct eigenvalues λ_i.

$$A|V^{\lambda_i} = \begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}$$

Further, $\mathbb{C}^n = V^{\lambda_1} \bigoplus V^{\lambda_2} \bigoplus ... \bigoplus V^{\lambda_k}$.

STEP 3. Every upper triangular $A|V^{\lambda_i}$ can be reduced into Jordan Block.

4 STEP 2: Understand Root Subspace V^{λ_i}

- Find $A|V^{\lambda_i}$.

  By definition of root vector, v is a root vector if for some m, $(A-\lambda Id)^m v = 0$. Root subspace $V^{\lambda}(A) = ker(A-\lambda Id)^m$, is the space of all root vectors of λ.

  Also, $ker(A-\lambda Id) \subset ker(A-\lambda Id)^2 \subset ... \subset ker(A-\lambda Id)^m = ker(A-\lambda Id)^{m+1}$.

  Suppose e_1, e_2 basis of $ker(A-\lambda Id)$, then e_1, e_2 eigenvectors of A. $Ae_1 = \lambda e_1$, $Ae_2 = 0e_1 + \lambda e_2$.

  Further suppose e_1, e_2, e_3, e_4 basis of $ker(A-\lambda Id)^2$. Then $(A-\lambda Id)(A-\lambda Id)e_3 = 0$. $(A-\lambda Id)e_3 \in ker(A-\lambda Id)$, is linearly combination of e_1, e_2. $A(e_3) - \lambda e_3 = ae_1 + be_2$, then $A(e_3) = ae_1 + be_2 + \lambda e_3$.

  Similarly, we can get basis of V^{λ_i} such that $A(e_i) = l.c.\{e_1, e_2, ..e_{i-1}\} + \lambda e_i$.

  So that, $A|V^{\lambda_i}$ **is upper triangular matrix with λ on all diagonal entries** .

$$A|V^{\lambda_i} = \begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}$$

  Assume $dimV^{\lambda}(A) = k$, then subordinate basis $\{e_1, .., e_k\}$, $A|V^{\lambda_i}$ is $k \times k$ matrix.

  Further, characteristic polynomial of $A|V^{\lambda_i}$ will be $det(A|V^{\lambda_i} - tId|V^{\lambda_i}) = (\lambda - t)^k$.

Lemma 1. **For $\lambda_i \neq \lambda_j$, $(A-\lambda_j Id)|V^{\lambda_i}(A)$ is invertible.**

Proof: Diagonal entries: $\lambda_i - \lambda_j \neq 0$. Thus, $det = (\lambda_i - \lambda_j)^k \neq 0$. Invertible.

(will be used in proof of Proposition 2)

- Proposition 1

  Let $k = dimV^{\lambda}(A)$, then multiplicity of λ in characteristic polynomial of A is k.

  Proof: Expend $e_1, ..., e_k$ to basis of $\mathbb{C}^n$, then A can be written as:

$$\begin{array}{c} \quad\quad e_1 \quad .... \quad e_k, e_{k+1}... \\ A = \begin{bmatrix} \lambda_i & & * & & \\ & \ddots & & * & \\ 0 & & \lambda_i & & \\ & & & & \\ & 0 & & B & \\ & & & & \end{bmatrix} \end{array}$$

  The characteristic polynomial of A: $det(A-tId) = (\lambda - t)^k \times det(B-tId)$.

  If not a root of $det(B-tId)$, λ only exists in $(\lambda-t)^k$, then multiplicity is k.

  Assume λ is a root $det((B-tId)$, then λ is an eigenvalue of B. There is $w \neq 0$, $Bw = \lambda w$.

  Then extend w into $\mathbb{C}^n$ with 0 on $e_1, .., e_k$, we can get w'. $Aw' = \lambda w' + v$, where $v \in V^{\lambda}(A)$.

Then $(A - \lambda Id)w' = v \in V^{\lambda}(A)$, which means$w'$ also in $V^{\lambda}(A)$.

But $w' = (0, ..0, w_1, .., w_{n-k})$, can not be expressed in $e_1, ...e_k$ described above, not a root vector of λ. Contradiction.

Thus, there is no λ in $det(B - tId)$, the multiplicity of λ is k.

Then, for each distinct eigenvalues $\lambda_1, \lambda_2, .., \lambda_k$ of A, there are root subspaces $V^{\lambda_1}, V^{\lambda_2}, ..., V^{\lambda_k}$ of each eigenvalues. λ_i has multiplicity $m_i = dimV^{\lambda_i}$.

According to *Fundamental Theorem of Algebra*, $m_1 + m_2 + ... + m_k = n$. Then $dimV^{\lambda_1} + dimV^{\lambda_2} + ... + dimV^{\lambda_k} = n$.

To prove $\mathbb{C}^n = V^{\lambda_1} \bigoplus V^{\lambda_2} \bigoplus ... \bigoplus V^{\lambda_k}$, we still need to prove that set of all vectors in bases of $V^{\lambda_1}, V^{\lambda_2}, ..., V^{\lambda_k}$ is linearly independent.

Lemma 2. **Let $m = dimV^{\lambda}(A)$, $v \in V^{\lambda}(A)$. Then $v \in ker(A - \lambda Id)^m$.**

Proof: By definition of root vector, $(A - \lambda Id)^k v = 0$, for some k. $V^{\lambda}(A) = ker(A - \lambda Id)^k$.

Assume $k > m$.

Then $ker(A - \lambda Id) \subset ... \subset ker(A - \lambda Id)^m \subset ... \subset ker(A - \lambda Id)^k = V^{\lambda}(A)$.

Each time degree of $(A - \lambda Id$ increase by 1, dimension of set increase at least by 1.

Then $dimV^{\lambda}(A) > m$. Contradiction. Only $k \leqslant m$ possible.

Thus, $(A - \lambda Id)^m v = (A - \lambda Id)^{m-k} v \times (A - \lambda Id)^k v = 0$. $v \in ker(A - \lambda Id)^m$.

- Proposition 2.

Let $e_1 \in V^{\lambda_1}, ..., e_k \in V^{\lambda_k}$, $e_1, ...e_k \neq 0$. Then $\{e_1, ..., e_k\}$ linearly independent.

Proof: To prove linearly independence, suppose $\mu_1 e_1 + \mu_2 e_2 + ... + \mu_k e_k = 0$.

When $k = 2$:

To kill e_1, let $dimV^{\lambda_1} = m_1$. Then by Lemma 2, $(A - \lambda_1 Id)^{m_1} e_1 = 0$.

Multiply by $(A - \lambda_1 Id)^{m_1}$, get $\mu_1 0 + \mu_2 (A - \lambda_1 Id)^{m_1} e_2 = 0$.

Since $\lambda_1 \neq \lambda_2$, diagonal entries of $(A - \lambda_1 Id)|V^{\lambda_2} = \lambda_2 - \lambda_1 \neq 0$.

Then $(A - \lambda_1 Id)^m$ can not send e_2 gradually to 0. $(A - \lambda_1 Id)^{m_1} e_2 \neq 0$.

$\mu_2 (A - \lambda_1 Id)^{m_1} e_2 = 0$, thus μ_2 must be 0. $e_1 \neq 0$, then μ_1 also must be 0.

Therefore, $\{e_1, e_2\}$ linearly independent.

When k = 3:

Similarly, $\mu_1 0 + \mu_2 (A - \lambda_1 Id)^{m_1} e_2 + \mu_3 (A - \lambda_1 Id)^{m_1} e_3 = 0$

To kill e_2, let $dimV^{\lambda_2} = m_2$. Then by Lemma 2, $(A - \lambda_2 Id)^{m_2} e_2 = 0$.

Multiply, get $\mu_2 (A - \lambda_1 Id)^{m_1} (A - \lambda_2 Id)^{m_2} e_2 + \mu_3 (A - \lambda_1 Id)^{m_1} (A - \lambda_2 Id)^{m_2} e_3 = 0$.

Similarly, by Lemma 1, $(A - \lambda_1 Id)^{m_1} (A - \lambda_2 Id)^{m_2} | V^{\lambda_3}$ $det \neq 0$, invertible.

$(A - \lambda_1 Id)^{m_1} (A - \lambda_2 Id)^{m_2} e_3 \neq 0$, then μ_3 must be 0.

Then $\mu_2 (A - \lambda_1 Id)^{m_1} (A - \lambda_2 Id)^{m_2} e_2 = 0$, μ_2 must 0. Similarly, $\mu_1 = 0$.

Therefore, $\{e_1, e_2, e_3\}$ linearly independent.

For k root spaces, repeated this process applying $(A-\lambda_i Id)^{m_i}$ to kill terms $e_1, .., e_{k-1}$.

Then we can get $\mu_k, \mu_{k-1}, .., \mu_1 = 0$, $\{e_1, ..., e_k\}$ linearly independent.

$V^{\lambda_2}, ..., V^{\lambda_k}$ each has adapted basis. Let basis of V^{λ_j} be $E_j = \{e_1^{\lambda_j}, e_2^{\lambda_j}, .., e_{m_k}^{\lambda_j}\}$

Then we can let $\sum_{i=1}^{m_1} \mu_i^{\lambda_1} e_i^{\lambda_1} + \sum_{i=1}^{m_2} \mu_i^{\lambda_2} e_i^{\lambda_2} + ... + \sum_{i=1}^{m_k} \mu_i^{\lambda_k} e_i^{\lambda_k} = 0$

By Proposition 2, each $\sum_{i=1}^{m_j} \mu_i^{\lambda_j} e_i^{\lambda_j}$ is 0. Then every μ is 0. Set of all vectors $E_1, E_2, ..., E_k$ is linearly independent.

Consider $E = \cup E_j$ for $j = 1, .., k$. By Proposition 1, $dim E = n$.

Thus E is a basis of $\mathbb{C}^n$. $\mathbb{C}^n = V^{\lambda_1} \bigoplus V^{\lambda_2} \bigoplus ... \bigoplus V^{\lambda_k}$

$$A = \begin{bmatrix} A|V^{\lambda_1} & & & 0 \\ & A|V^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & A|V^{\lambda_k} \end{bmatrix}$$

5 STEP 3: Jordan Block

Instead $A|V^{\lambda_i}$, we will consider $N = A - \lambda Id$ in this step, without loss of generality.

By definition of root space, $(A - \lambda Id)^m = N^m = 0$, for some m.

Then $N : V \to V$ is nilpotent. There exists $k = height(v)$ such that, $N^{k-1}v \neq 0$, $N^k v = 0$.

Lemma 3. **$\forall v \in V$, let $k = height(v)$, then $\{v, Nv, .., N^{k-1}v\}$ is linearly independent.**

Proof: To prove linearly independence, suppose $\mu_1 v_1 + \mu_2 N v_2 + ... + \mu_k N^{k-1} v = 0$.

$\{v, Nv, .., N^{k-1}v\}$ none is zero, $N^k v = 0$.

Multiply N^{k-1}, get $\mu_1 K^{k-1} v_1 + \mu_2 N^k v_2 + ... + \mu_k N^{2k-1} v = 0$.

Then $\mu_1 K^{k-1} v_1 = 0$ and $K^{k-1} v_1 \neq 0$, thus μ_1 must be 0.

We get $\mu_1 v_2 + \mu_2 N v_3 + ... + \mu_k N^{k-1} v = 0$

Multiply N^{k-2}, similarly μ_2 must be 0.

Therefore all μ_i are 0, $\{v, Nv, .., N^{k-1}v\}$ is linearly independent.

Easy to see that cyclic subspace $span\{v, Nv, .., N^{k-1}v\} \subset V$, is invariant under N.

$$\begin{array}{cc} & v, Nv, \quad ..., \quad N^{k-1}v \\ N|span\{v, Nv, .., N^{k-1}v\} = & \begin{bmatrix} 0 & 0 & & \cdots & 0 \\ 1 & 0 & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & & \\ 0 & & & & 1 \end{bmatrix} \end{array}$$

Rearrange the basis, we can make N upper triangular.

$$\begin{array}{cc} & N^{k-1}v, N^{k-2}v \quad ... \quad Nv, v \\ N|span\{N^{k-1}v, N^{k-2}v, Nv, v\} = & \begin{bmatrix} 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & 0 \\ & & & & & 1 \\ 0 & & & & & 0 \end{bmatrix} \end{array}$$

- Proposition 3.

  Let $N : V \to V$ **nilpotent, then** $\exists v_1, ..., v_q$**,** $ht(v_i) = m_i$**. Then** $\{v_i, Nv_i, ..., N^{m_i-1}v_i, \quad i = 1, 2, .., q\}$ **is a basis of V.** Let $U_i = span\{v_i, NV_i, ..., N^{m_i-1}v_i\}$, then $V = U_1 \bigoplus U_2 \bigoplus ... \bigoplus U_k$.

  Proof: Induction on $dimV = n$.

  $dimN(V) < n$, since $kerN(V) > 0$, exists $v \neq 0$ such that $N^m v = 0$.

  Let U any $n - 1$ dimension subspace contains $N(v)$, $N(v) \subset U$.

  Then $N(U) \subset N(V) \subset U$.

  Apply inductive hypothesis, we have $U = U_1 \bigoplus U_2 \bigoplus ... \bigoplus U_k$, each U_i cyclic.

  Then consider $N|V$, take $v \in V$, $u \in U$, $Nv = u_1 + u_2 + .. + u_k$.

  Then there are two situations: 1) $\forall u_i = 0$, 2)$\exists u_i \neq= 0$

  1). $V =< v > \bigoplus U_1 \bigoplus ... \bigoplus U_k$, $q = k + 1$. Done

  2). Then $height(u_i) = m_i$. We can get $e_2 = Ne_1$, $e_3 = N^2 e_1$,...

  Here, we can write $V = span\{v, Nv, .., N^{m_1}v\} \bigoplus U_2 \bigoplus ... \bigoplus U_k$.

  Further, it can be proved that $span\{v, Nv, .., N^{m_1}v\} \cap U_2 \bigoplus ... \bigoplus U_k = \{0\}$

By Proposition 3, we get the 1s above diagonal part of Jordan Norm Formal.

$$N = A - \lambda Id = \begin{bmatrix} 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & 0 \\ & & & & & 1 \\ 0 & & & & & 0 \end{bmatrix}$$

Add back eigenvalues λs on diagonal, $A = N + Id$, we can get the matrix in Jordan Normal Form finally.

6 Special Case: Diagonalization of Symmetirc Matrix

Attached as last page.

7 Note

" Camille Jordan (1838 – 1922) was a French mathematician, know nfor his foundational work in group theory and for his influential Cours d'analyse. He is remembered now by name in a number of results:

1. The Jordan curve theorem, a topological result required in complex analysis

2. The Jordan normal form and the Jordan matrix in linear algebra

3. The Jordan measure, an area measure that predates measure theory in math analysis

... " by Wiki. Respect.

Jordan Normal Form really opens the door towards modern linear algebra for me. I believe that the proof described here is just a basic story-telling version of this problem. From more higher or abstract level, there should be more elegant and intuitive solutions.

This problem practice my math induction techniques further (again, just like every time go further in math study), emphasizing on understanding conditions of applying inductive hypothesis. Also, from this problem, I realize how linear algebra and abstract algebra join together, showing different perspectives of same object (and make things much more complex, not surprisingly), e.x. stabilizing, cyclic, quotient group,... which will encourage me to think abstract algebra concept more linearly and linear algebra more abstractly.

Proof of Proposition 3 still not completed. Will go back and edit again.

5520H AU20

✡ Spectral Theorem

Symmetric matrix diagonalizable.

A $n\times n$, $\exists$ O orthognal, D diagonal $A = ODO^{-1}$

proof: by induction on $n\times n$

Base: A 1×1: $A = [a] = [1][a][1]$

Suppose: $n-1 \times n-1$ $B = QD'Q^{-1}$

$n\times n$: A has real eigenvalues, eigenvectors

let $Av_1 = \lambda_1 v_1$ $v_1 \neq 0$, normalize $\to |u_1| = 1$

add vectors from $\mathbb{R}^n$ not necessary eigenvectors

$\to$ G-S process: orthonormal basis $\{u_1, \cdots, u_n\}$

let matrix $O = [u_1 \cdots u_n]$ express new basis in old

orthognal ✓

want D — consider $O^{-1}AO$:

O^{-1}: $\begin{array}{c} \\ \begin{bmatrix} e_1 \cdots e_n \\ \\ \end{bmatrix} \end{array} \begin{array}{c} u_1 \\ \vdots \\ u_n \end{array}$ A: $\begin{bmatrix} \\ \\ \end{bmatrix} \begin{array}{c} e_1 \\ \vdots \\ e_n \end{array}$ O: $\begin{bmatrix} u_1 \cdots u_n \\ \\ \end{bmatrix} \begin{array}{c} e_1 \\ \vdots \\ e_n \end{array}$

send e_i to u_i trans in e_i send u_i to e_i

- symmetric: $(O^{-1}AO)^T = O^T A^T (O^{-1})^T = O^{-1}AO$
- first column: $(O^{-1}AO)e_1 = O^{-1}A \cdot O(e_1)$

$= O^{-1}A(u_1)$ eigenvector

$= O^{-1}\lambda_1 u_1 = \lambda_1 \cdot O^{-1}(u_1) = \lambda_1 \cdot e_1 = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

then $O^{-1}AO$: $\begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix}$ $B = QD'Q^{-1}$

$n-1\times n-1$ symmetric

cont.

let $R = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$ $\quad Q$: orthognal

$$\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I \end{bmatrix} \qquad R^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix}$$

let $U = \underbrace{\underline{O}\,\underline{R}}_{\text{orthogonal}} \to$ orthogonal

$$\begin{aligned} U^{-1} A U &= (R^{-1} O^{-1}) A (O R) \\ &= R^{-1} (O^{-1} A O) R \\ &= \begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & Q^{-1}B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & Q^{-1}BQ \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & D' \end{bmatrix} \end{aligned}$$

$\to A = UDU^{-1} \quad U = OR \quad U^{-1} = R^{-1} O^{-1}$

$$A = \overset{O}{\begin{bmatrix} u_1 \cdots u_n \\ \ \end{bmatrix}}\begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix} \overset{R}{\begin{bmatrix} 1 & \\ & Q \end{bmatrix}} \overset{D}{\begin{bmatrix} \lambda_1 & \\ & \underbrace{D'}_{\to B} \end{bmatrix}} \overset{R^{-1}}{\begin{bmatrix} 1 & \\ & Q^{-1} \end{bmatrix}} \overset{O^{-1}}{\begin{bmatrix} e_1 \cdots e_n \\ \ \end{bmatrix}}\begin{matrix} u_1 \\ \vdots \\ u_n \end{matrix}$$