1 Chapter 1: Diagonalization of matrices

1.1 Definitions

Let E be an $n-$dimensional space vector over a field K, where $K = \mathbb{R}$ or $\mathbb{C}$. $dimE = n$, B a basis of E. Let $f : E \longrightarrow E$ a linear application (endomorphism of E), A the square matrix $(n \times n)$ associated with f: $A = \mathbb{M}_B(f) = (a_{ij})$.

1.1.1 Definition 1. Characteristic Polynomial of a Matrix

If A is an $n \times n$ matrix, the **characteristic polynomial** $P(\lambda)$ of A is defined by:

$$P(\lambda) = det(A - \lambda I_n)$$

1.1.2 Definition 2. Eigenvalues and Eigenvectors

If A is $n \times n$ matrix, a number λ is called an eigenvalue of A if there is $V \in E$ such that:

$$AV = \lambda V$$

In this case, V is called an eigenvector of A corresponding to the eigenvalue λ.

Example. If $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ and $V = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ then $AV = \begin{bmatrix} 20 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 5 \\ 1 \end{bmatrix} = 4V$
So $\lambda = 4$ is an eigenvalue of A with corresponding eigenvector V.

Theorem. Let A be an $n \times n$ matrix.

1. The eigenvalues λ of A are the roots of the characteristic polynomial $P(\lambda)$ of A.
$$P(\lambda) = 0$$
2. The $\lambda-$ eigenvectors X are the nonzero solutions to the homogeneous system
$$(A - \lambda I)X = 0$$

1.1.3 Definition 3.

Let A be $n \times n$ matrix and λ an eigenvalue of the matrix A. The set

$$E(\lambda) = \{V \in E, AV = \lambda V\}$$

is called the **eigenspace** of A associated to the eigenvalue λ in which $E(\lambda)$ is vector sub-space of E. Its dimension $(dimE(\lambda))$ is called the the geometric multiplicity of λ.

1.1.4 Definition 4. Similarity and Diagonalization

If A, B are two $n \times n$ matrices, then they are **similar** if and only if there exists an invertible matrix P such that:

$$A = P^{-1}BP$$

1.1.5 Definition 5. Trace of a matrix

If $A = (a_{ij})$ is an $n \times n$ matrix, then the trace of A is

$$trace(A) = tr(A) = \sum_{i=1}^{n} a_{ij}$$

Lemma. Properties of a trace For $n \times n$ matrices A and B, and any $k \in \mathbb{R}$,

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(kA) = k.tr(A)$
3. $tr(AB) = tr(BA)$

Theorem. Properties of similar matrices If A and B are $n \times n$ matrices and A, B are similar, then

1. $det(A) = det(B)$
2. $rank(A) = rank(B)$
3. $tr(A) = tr(B)$
4. $P_A(\lambda) = P_B(\lambda)$
5. A and B have the same eigenvalues.

Proof. **1.** We have $B = P^{-1}AP$, then $det(B) = det(P^{-1}AP) = det(A)$

4. $P_B(\lambda) = det(B - \lambda I_n) = det(P^{-1}AP - P^{-1}\lambda P) = det[P^{-1}(A - \lambda I_n)P] = det(P^{-1}) \times det(A - \lambda I_n) \times det(P)$

□

1.1.6 Definition 6. Digonalizable

Let A be an $n \times n$ matrix. Then A is said to be **diagonalizable** if there exists an invetible matrix P such that

$$P^{-1}AP = D$$

where D is a diagonal matrix.

Proposition. Let λ_1 and λ_2 be two distinct eigenvalues ($\lambda_1 \neq \lambda_2$) of A, then

$$E(\lambda_1) \cap E(\lambda_2) = \{0\}$$

Proof. If $V \in E(\lambda_1) \cap E(\lambda_2)$, then $AV = \lambda_1 V = \lambda_2 V$ i.e. $(\lambda_1 - \lambda_2)V = 0$.
Since $\lambda_1 \neq \lambda_2$, then we have $V = 0$ □

1.1.7 Definition 7. Diagonalization

A square $n \times n$ matrix A is **diagonalizable** if A is similar to a diagonal matrix, i.e.

$$A = PDP^{-1}$$

for a diagonal matrix D and an invertible matrix P.

Proposition. Let A be an $n \times n$ matrix. We suppose that $P(\lambda)$ have k distinct roots $\lambda_1, \lambda_2, ..., \lambda_k$. If $E = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_k)$, then A is diagonalizable.

Proof. For $i = 1, 2, ..., k$, we choose the basis B_i of $E(\lambda_i)$. The basis $B' = \cup_{i=1}^{i=k} B_i$ of E consists of the eigenvectors of A associated with the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$, then the matrix $D = \mathbb{M}_{B'}(f)$ is diagonal. □

Examples Find the characteristic polynomial, eigenvalues and eigenvectors of the matrices:

1. $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$
2. $A = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{bmatrix}$

Solution.

1. $P(\lambda) = (\lambda - 4)(\lambda + 2)$
   $\lambda_1 = -2$ and $\lambda_2 = 4$
   $V_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $V_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

2. $P(\lambda) = -\lambda(\lambda - 1)(\lambda - 2)$
   $\lambda_1 = 0, \lambda_2 = 4$ and $\lambda_3 = 2$
   $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $V_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

1.2 Sufficient condition for a matrix to be diagonalizable

Proposition. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. We have $P(\lambda) = (-1^n)(\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$, where $\lambda_1, \lambda_2, ..., \lambda_n$ n distinct eigenvalues of A and $V_1, V_2, ..., V_n$ the n eigenvectors associated with λ_i.
$AV_1 = \lambda_1 V_1$
$AV_2 = \lambda_2 V_2$
.
.
.
$AV_n = \lambda_n V_n$
We can prove that $B' = (V_1, V_2, ..., V_n)$ is a basis of E by induction:
We prove that the set $(V_1, V_2, V_3, ..., V_{k+1})$ is linearly independent of E.

$$\alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_k V_k + \alpha_{k+1} V_{k+1} = 0 \quad (1)$$

We have $A(\alpha_1 V_1 + \alpha_2 V_2 + ... + \alpha_k V_k + \alpha_{k+1} V_{k+1}) = 0$, then
$\alpha_1 AV_1 + \alpha_2 AV_2 + ... + \alpha_k AV_k + \alpha_{k+1} AV_{k+1} = 0$

$$\alpha_1 \lambda_1 V_1 + \alpha_2 \lambda_2 V_2 + ... + \alpha_k \lambda_k V_k + \alpha_{k+1} \lambda_{k+1} V_{k+1} \quad (2)$$

From $(2) - \lambda_{k+1}(1)$:
$(\lambda_1 - \lambda_{k+1})\alpha_1 V_1 + (\lambda_2 - \lambda_{k+1})\alpha_2 V_2 + ... + (\lambda_k - \lambda_{k+1})\alpha_k V_k = 0$
Since the set $(V_1, V_2, ..., V_k)$ is linearly independent of E by induction hypothesis, then $(\lambda_1 - \lambda_{k+1})\alpha_1 = (\lambda_2 - \lambda_{k+1})\alpha_2 = ... = (\lambda_k - \lambda_{k+1})\alpha_k = 0$ (because λ_k are distinct).
Therefore $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$
By (1) we have $\alpha_{k+1} V_{k+1} = 0$, then $\alpha_{k+1} = 0$ □

1.3 Necessary and sufficient condition for diagonalizability

Proposition 1. Let A be an $n \times n$ matrix, then

$$dim(E(\lambda_1)) \leq m_1$$

where λ_1 is an eigenvalue of A multiplicity m_1.

Proof. Let $(e_1, e_2, ..., e_r)$ the basis of $E(\lambda_1)$, then we can find the basis $B = (e_1, e_2, ..., e_r, e_{r+1}, ..., e_n)$ of E.
The matrix A is similar of the matrix A' of the form

$$A' = \left(\begin{array}{cccc|c} \lambda_1 & & & & \\ & \lambda_1 & & & \\ & & \ddots & & A_1 \\ & & & \lambda_1 & \\ \hline & & 0 & & A_2 \end{array}\right)$$

$$P(\lambda) = det(A - \lambda I_n) = \left|\begin{array}{cccc|c} \lambda_1 - \lambda & & & & \\ & \lambda_1 - \lambda & & & \\ & & \ddots & & A_1 \\ & & & \lambda_1 - \lambda & \\ \hline & & 0 & & A_2 - \lambda I_{n-r} \end{array}\right|$$

$$= (\lambda_1 - \lambda)^r det(A_2 - \lambda I_{n-r})$$

Then $m \geq r$, where $r = dimE(\lambda_1)$ □

Proposition 2. Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if:

1. $P(\lambda)$ is factored.
2. For each eigenvalue λ_i of A, $dim(E(\lambda_i)$ is equal to the multiplicity of λ_i i.e.
$$dimE(\lambda_i) = m_i, i = 1, ..., k$$

Proof. By induction, the sub-spaces $E(\lambda_i)$, $i = 1, ..., j$, verify

$$E = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_k)$$

for $j = 1, ..., k$
Denote $S_j = E(\lambda_1) \oplus E(\lambda_2) \oplus ... \oplus E(\lambda_j)$
It is sufficient to demonstrate that $S_j \cap E(\lambda_{j+1}) = \{0\}$
Let $V \in S_j \cap E(\lambda_{j+1})$, then

$$\begin{cases} V = V_1 + V_2 + ... + V_j \\ \text{and} \\ AV = \lambda_{j+1} V \end{cases} \tag{3}$$

For (3), we have $AV = AV_1 + AV_2 + ... + AV_j$, then

$$\lambda_{j+1} V = \lambda_1 V_1 + \lambda_2 V_2 + ... + \lambda_j V_j \tag{4}$$

For $(4) - \lambda_{j+1}(3)$, we have

$$0 = (\lambda_1 - \lambda_{j+1})V_1 + (\lambda_2 - \lambda_{j+1})V_2 + ... + (\lambda_j - \lambda_{j+1})V_j$$

Using induction hypothesis, we get $V_1 = V_2 = ... = V_j = 0$
Since $\sum_{i=1}^{n} dimE(\lambda_i) = \sum_{i=1}^{n} m_i = n$, we see that $E = \oplus_{i=1}^{k} E(\lambda_i)$. Then A is diagonalizable and we write:

$$D = \begin{bmatrix} \lambda_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && \lambda_1 &&&&&& \\ &&& \lambda_2 &&&&& \\ &&&& \ddots &&&& \\ &&&&& \lambda_2 &&& \\ &&&&&& \ddots && \\ &&&&&&& \lambda_k && \\ &&&&&&&& \ddots & \\ &&&&&&&&& \lambda_k \end{bmatrix}$$

□

Examples.

1. $A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}$
   $P(\lambda) = -\lambda(\lambda - 1)^2$
   $P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$
   $E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$
   $E(\lambda_2) = E(1) =< V_2, V_3 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$, $V_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $dimE(\lambda_2) = 1 = m_2 = 2$.
   Then the matrix A is diagonalizable.

2. $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{bmatrix}$
   $P(\lambda) = -\lambda(\lambda - 1)^2$
   $P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$
   $E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$
   $E(\lambda_2) = E(1) =< V_2 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $dimE(\lambda_2) = 1 \neq m_2 = 2$
   Then the matrix A isn't diagonalizable.

2 Chapter 2: Triangulability of matrices

Example 1. Consider the matrix $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ 1 & 3 & -4 \end{bmatrix}$, then

$P(\lambda) = -\lambda(\lambda - 1)^2$

$P(\lambda) = 0 \Rightarrow \begin{cases} \lambda_1 = 0, m_1 = 1 \\ \lambda_2 = 1, m_2 = 2 \end{cases}$

$E(\lambda_1) = E(0) =< V_1 >$, where $V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $dimE(\lambda_1) = 1 = m_1$

$E(\lambda_2) = E(1) =< V_2 >$, where $V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ and $dimE(\lambda_2) = 1 \neq m_2 = 2$

Then the matrix A isn't diagonalizable.

What to do if matrix A is not diagonalizable?
Therefore, we use triangulation:

2.1 Proposition

Let $f : E \to F$ a linear map and A the matrix of f, we suppose the characteristic polynomial $P(\lambda)$ of f (or A) is factored in $K[\lambda]$. Then f (or A) is triangulable.

Proof. By induction over $dimE$: the result is true for the space of dimension 1. Suppose they are true for spaces of dimension $\leq n - 1$ and let E be a space of dimension n.
Let $P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$ in $K[\lambda]$, ($K = \mathbb{R}$ or $\mathbb{C}$).
We suppose that the eigenvalues λ_i are not necessarily distinct. We denote V_1, an eigenvector associated with λ_1 (i.e $f(V_1) = \lambda_1 V_1$).
By the incomplete basis theorem, there exists a basis B' of E where
$B' = (V_1, e_2, e_3, ..., e_n)$ then the matrix A' has the form

$$A' = M_{B'}(f) = \begin{bmatrix} \lambda_1 & a_{12} & . & . & . & a_{1n} \\ 0 & a_{22} & & & & . \\ 0 & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & a_{n2} & . & . & . & a_{nn} \end{bmatrix}$$

The family $B_1 = (e_2, ..., e_n)$ is a basis of the subspace $F =< e_2, ..., e_n >$ of E. We denote $g : F \to F$, the linear map such that the associated matrix is

$$A_1 = \begin{bmatrix} a_{12} & . & . & . & a_{1n} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{n2} & . & . & . & a_{nn} \end{bmatrix} = M_{B_1}(g)$$

Then $P(\lambda) = (\lambda_1 - \lambda) \times det(A_1 - \lambda I_{n-1})$
i.e. $P(\lambda)$ is factored and since $dimF = n - 1$, by induction hypothesis, there exists a basis $B_2 = (V_2, ..., V_n)$ of F such that $M_{B_2}(g)$ is upper triangular. We get

$$M_{B'=(V_1,V_2,...,V_n)}(f) = \begin{bmatrix} \boxed{\lambda_1} & a_{12} & . & . & . & a_{1n} \\ & \lambda_2 & . & . & . & . \\ & & & . & . & . \\ & & & & . & . \\ & & & & & \lambda_n \end{bmatrix} \qquad \square$$

Remark.
1/ If A is triangulable, the diagonal of the matrix $T = M_{B'}(f)$ are the eigenvalues of A.
2/ All matrix of $A \in M_n(\mathbb{C})$ is triangulable.

Corollary.
$tr(A) = \sum_i \lambda_i$
$det(A) = \prod_i \lambda_i$

Remark.
We can triangulate the matrix A of Example 1.

We consider the basis B' of E where $\begin{cases} V_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = e_1 + e_2 + e_3 \\ V_2 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = e_1 + 3e_2 + 2e_3 \\ V_3 = e_1 \end{cases}$

Because $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 2 & 0 \end{vmatrix} = 2 - 3 = -1 \neq 0$

And $\begin{cases} e_1 = V_3 \\ e_2 = -2V_1 + V_2 + V_3 \\ e_3 = 3V_1 - V_2 - 2V_3 \end{cases}$

Then $T = M_{B'}(f) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = P^{-1}AP$

Where $\begin{cases} f(V_1) = \lambda_1 V_1 = 0 \\ f(V_2) = \lambda_2 V_2 = V_2 \\ f(V_3) = f(e_1) = e_1 + 2e_2 + e_3 = -V_1 + V_2 + V_3 \end{cases}$

Finally, $T = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ is the upper triangular matrix,

$P = (V_1 V_2 V_3) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 0 \\ 1 & 2 & 0 \end{bmatrix}$ and $P^{-1}(e_1 e_2 e_3) = \begin{bmatrix} 0 & -2 & 3 \\ 0 & 1 & -1 \\ 1 & 1 & -2 \end{bmatrix}$

2.2 Annihilating polynomials

Let E a vector space over K and $R \in K[\lambda]$
$R(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_2\lambda_2 + a_1\lambda^1 + a_0\lambda^0$
If $f \in End_K(E)$, we denote $R(f)$, the linear map of E defined by
$R(f) = a_n f^n + a_{n-1}f^{n-1} + ... + a_2 f^2 + a_1 f^1 + a_0 id$
or $R(A)$ the matrix
$R(A) = a_2 A^n + a_{n-1}A^{n-1} + ... + a_2 A^2 + a_1 A^1 + a_0 I_n$
Where $f^k = \underbrace{f \circ f \circ ... \circ f}_{\text{k times}}$

Remark.
We have $P(f) \circ Q(f) = Q(f) \circ P(f)$.

2.2.1 Definition.

Let $f \in End_K(E)$, the polynomial $R \in K[\lambda]$ is called annihilating polynomial of f if $R(f) = 0$ (or $R(A) = 0$).

2.3 Cayley-Hamilton theorem

Let $f \in End_k(E)$ and $P(\lambda)$ the characteristic polynomial of f (or A).
Then

$$P(f) = 0$$

(or $P(A) = 0$). i.e $P(\lambda)$ annihilates f (or A).

Proof. We suppose $K = \mathbb{C}$, in this case f (or A) is triangulable.
Let $B' = (V_1, V_2, .., V_n)$, a basis of E such that

$$M_{B'}(f) = \begin{pmatrix} \lambda_1 & a_{12} & . & . & . & a_{1n} \\ & \lambda_2 & a_{23} & . & . & a_{2n} \\ & & . & & & . \\ & & & & . & . \\ & & & & & \lambda_n \end{pmatrix}$$

We have $f(V_1) = \lambda_1 V_1 \Rightarrow (\lambda_1 id - f)(V_1) = 0$ and
$P(\lambda) = det(T - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda)$
Then $P(f) = (\lambda_1 id - f) \circ ... \circ (\lambda_n id - f)$ and
$P(f)(V_1) = (\lambda_2 id - f) \circ ... \circ (\lambda_n id - f) \circ (\lambda_1 id - f)(V_1) = 0$. Therefore, $P(f)(V_1) = 0$
$P(f)(V_2) = (\lambda_3 id - f) \circ ... \circ (\lambda_n id - f) \circ (\lambda_1 id - f) \circ (\lambda_2 id - f)(V_2) = (\lambda_3 id - f) \circ ... \circ (\lambda_n id - f) 0 (\lambda_1 id - f)(-a_{12} V_1) = 0$. Therefore, $P(f)(V_2) = 0$
We can similarly show that $P(f)(V_3) = 0$
By induction, we find $P(f)(V_i) = 0, \forall i = 1, ..., n$. Finally, $P(f) = 0$.

□

Example.

$$A = \begin{bmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 6 & 1 & 1 \end{bmatrix}$$

$P(\lambda) = det(A - \lambda I_3) = (2 - \lambda)(1 - \lambda)^2 = -\lambda^3 + 4\lambda^2 - 5\lambda + 2$
Since $det(A) = P(0) = 2 \neq 0$, A is invertible.
By the Cayley-Hamilton theorem, we have $P(A) = 0$
i.e $-A^3 + 4A^2 - 5A + 2I_3 = 0$. Then $-A^3 + 4A^2 - 5A = -2I_3 \Rightarrow$
$A[-A^2 + 4A - 5I_3] = -2I_3 \Rightarrow A[\frac{1}{2}A^2 - 2A + \frac{5}{2}I_3] = I_3$
Therefore,

$$A^{-1} = \frac{1}{2}A^2 - 2A + \frac{5}{2}I_3$$