LECTURE 33

Diagonal Form of a Matrix

Hamilton's Theorem

The Internal Input-Output Processing Structure of a Matrix

Diagonalizable Matrices.

I. DIAGONOL FORM OF A MATRIX

Read Strang 5.2

The essential properties of a $n\times n$ matrix with a set of n linearly independent eigen vectors are expressed in terms of its eigenvalues and the transition matrix formed from these eigen vectors. This is expressed by the following

Proposition 33.1

Let $A: R^n \to R^n$ be an $n\times n$ matrix
Let A have a set of n lin. indep. eigen vectors $\{\vec{e}_i\}$
Let S be the transition matrix formed from these eigenvalues.

Then

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & O \\ O & & \ddots \lambda_n \end{bmatrix} \equiv \Lambda$$

One says that A can be diagonalized

proof:

Consider the equations

$$A\vec{e}_i = \lambda_i \vec{e}_i \qquad i = 1, \ldots, n$$

satisfied by the given eigenvectors.

One has

$$A\begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_n \end{bmatrix} = \begin{bmatrix} \vec{e}_1\lambda_1 & \cdots & \vec{e}_n\lambda_n \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_n \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} \lambda_1 & & O \\ O & & \lambda_n \end{bmatrix}}_{\Lambda} \leftarrow \Lambda$$

or

$$A\,S = S\,\Lambda \qquad \text{where } S \text{ is invertible (why?)}$$

Thus $\boxed{S^{-1}AS = \Lambda}$ or $\boxed{A = S\Lambda S^{-1}}$

A is diagonalizable by means of a similarity transformation.

Comments:

1. It is easy to compute powers of a diagonalizable matrix:

$$A^m = (S\Lambda S^{-1})\cdots(S\Lambda S^{-1})$$
$$= S\,\Lambda\cdots\Lambda\,S^{-1}$$
$$\boxed{A^m = S\,\Lambda^m S^{-1}}$$

2. One can compute matrix power series: Consider a function defined by a convergent power series

$$g(x) = a_0 + a_1 x + a_2 x^2 + \cdots \text{ ; e.g. } e^x = 1 + x + \frac{x^2}{2} + \cdots$$

Then $g(A)$ is defined and convergent in an appropriate sense:

$$e^{At} = I + At + A^2\frac{t^2}{2} + \cdots$$

$$= SIS^{-1} + S\Lambda S^{-1}t + S\Lambda^2 S^{-1}\frac{t^2}{2!} + \cdots$$

$$= S\left(I + \Lambda t + \frac{\Lambda^2 t^2}{2!} + \cdots\right)S^{-1}$$

$$= S\begin{bmatrix} e^{\lambda_1 t} & & O \\ & e^{\lambda_2 t} & \\ O & \ddots & e^{\lambda_n t}\end{bmatrix}S^{-1}$$

$$= S\, e^{\Lambda t} S^{-1}$$

Thus $\boxed{g(A) = S\, g(\Lambda) S^{-1}}$

3. Consider the secular (a.k.a. characteristic) polynomial

$$\det|A - \lambda I| = a_0 + a_1(-\lambda) + \cdots + a_n(-\lambda)^n$$

$$= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) \equiv f(\lambda)$$

$$= \lambda_1\cdots\lambda_n + (\cdots)(-\lambda) + (\cdots)(-\lambda)^2 + \cdots + (\lambda_1 + \cdots \lambda_n)(-\lambda)^{n-1} + (-\lambda)^n = f(\lambda)$$

The secular pol'l in the powers of A vanishes:

$$f(A) = S^{-1} f(\Lambda) S =$$

$$= S^{-1}\begin{bmatrix} (\lambda_1-\lambda_1)(\lambda_2-\lambda_1)\cdots(\lambda_n-\lambda_1) & & O \\ & \ddots & \\ O & & (\lambda_1-\lambda_n)(\lambda_2-\lambda_n)(\lambda_n-\lambda_n)\end{bmatrix} = S^{-1}\begin{bmatrix} 0 & & O \\ & \ddots & \\ O & & 0\end{bmatrix}S$$

$$= [O]_{\text{matrix}}$$

We have

$$f(\lambda) = \lambda_1\cdots\lambda_n + (\cdots)(-\lambda) + (\cdots)(-\lambda)^2 + \cdots + (\lambda_1 + \lambda_2 + \cdots + \lambda_n)(-\lambda)^{n-1} + (-\lambda)^n$$

$$\boxed{f(A) = a_0 + a_1(A) + a_2(-A)^2 + \cdots a_{n-1}(-A)^{n-1} + (-A)^n = \text{zero matrix}}$$

This illustrates the Cayley-Hamilton theorem:

Theorem 33.1

$\boxed{\text{A matrix satisfies its characteristic equation}}$

Here the coefficients

$$a_0 = f(0) = \lambda_1\lambda_2\cdots\lambda_n = \det A$$

$$a_{n-1} = \lambda_1 + \cdots + \lambda_n = \sum_{i=1}^{n} A_{ii} \equiv \text{trace} A \qquad (*)$$

depend only on the eigenvalues of A.

4. Question: True or False: $tr\, AB = tr\, BA$?

Answer: True. Indeed one has

$$tr\, AB = A_{ij}B_{ji} = B_{ji}A_{ij} = tr\, BA$$

Application: For $S^{-1}AS = \Lambda$ one has

$$\lambda_1 + \cdots + \lambda_n = tr\,\Lambda$$

$$= tr\, S^{-1}AS$$

$$= tr\, ASS^{-1}$$

$$= tr\, A$$

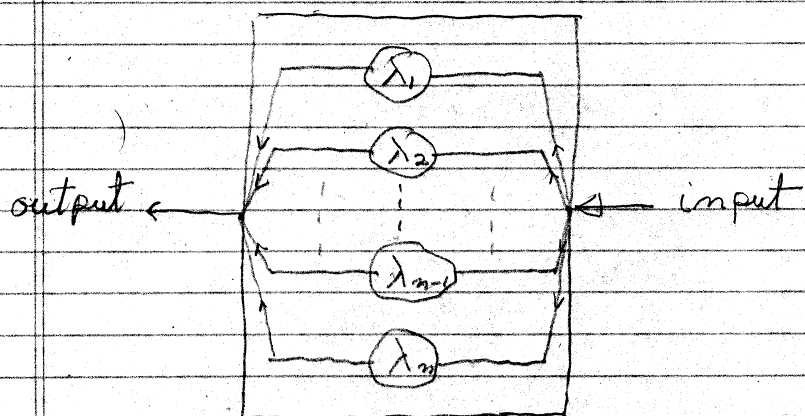
$\boxed{tr\,\Lambda = tr A}$

This validates Eq. (*).

4. A diagonalizable matrix A is a linear system whose internal structure is characterized by its eigenvalues

$$\lambda_1, \lambda_2, \ldots, \lambda_n$$

Contact with its exterior is provided by its eigenvector basis and the corresponding dual basis. Diagramatically one has



$$\boxed{A = \sum_{i=1}^{n} \vec{e}_i \lambda_i \underset{\sim}{\omega}^i}$$

eigen vectors $\{\vec{e}_j\}$

eigen covectors $\{\underset{\sim}{\omega}^i\}$

(a) This diagram conceptualizes the following physical process:

A matrix, say A, is a linear input-output device. Its output exhibits a causal relationship to the input. As depicted in the diagram, this relationship is the result of a three-stage process

1. The input is decomposed (by the eigen covector basis $\{\underset{\sim}{\omega}^i\}$; see page 33.8) into its eigenvector components,
2. Each eigenvector component gets multiplied ("amplified", "phase-shifted", "filtered") by its respective eigenvalue λ_i.
3. The results of this modification get recombined by the eigenvector basis $\{\vec{e}_i\}$. This recombination is the output from A.

(b) The mathematical (here algebraic) formulation of the diagram on page 33.6 is achieved by the following

4-step process:

STEP I.

Let $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ be the eigenvector basis corresponding to $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

Let $\{\underset{\sim}{\omega}^1, \underset{\sim}{\omega}^2, \ldots, \underset{\sim}{\omega}^n\}$ be the corresponding cobasis for the dual space V^* (See Lecture 8).

The two bases are related by the duality relation

$$\langle \underset{\sim}{\omega}^i \,|\, \vec{e}_j \rangle = \delta^i_{\ j} .$$

Explicitly one has

$$\vec{e}_j = \begin{bmatrix} e^1_j \\ e^2_j \\ \vdots \\ e^n_j \end{bmatrix}, \text{ the } j^{th} \text{ basis vector:}$$

$$A\vec{e}_j = \lambda_j \vec{e}_j$$

STEP II

Construct the cobasis elements

$$\underset{\sim}{\omega}^i = [\omega^i_1, \omega^i_2, \ldots, \omega^i_n]$$

From the duality relation

$$\delta^i_{\ j} = \langle \underset{\sim}{\omega}^i \,|\, \vec{e}_j \rangle \equiv [\omega^i_1, \ldots, \omega^i_n] \begin{bmatrix} e^1_j \\ \vdots \\ e^n_j \end{bmatrix}$$

one solves for ω^i_ℓ, the components of each $\underset{\sim}{\omega}^i$:

$$\sum_{\ell=1}^{n} \omega^i_{\ \ell} e^\ell_j = \delta^i_{\ j}$$

Note that if the $\vec{e}_j$'s are the columns of

$$S = \begin{bmatrix} e^1_1 & e^1_2 & & e^1_n \\ e^2_1 & e^2_2 & & e^2_n \\ \vdots & \vdots & & \vdots \\ e^n_1 & e^n_2 & & e^n_n \end{bmatrix}$$

Then the $\underset{\sim}{\omega}^i$ are the rows of

$$\begin{array}{cl} S^{-1} = \begin{bmatrix} \omega^1_1 & \omega^1_2 & \cdots & \omega^1_n \\ \vdots & \vdots & & \vdots \\ \omega^n_1 & \omega^n_2 & \cdots & \omega^n_n \end{bmatrix} & \begin{matrix} \leftarrow \underset{\sim}{\omega}^1 \\ \\ \leftarrow \underset{\sim}{\omega}^n \end{matrix} \end{array}$$

STEP III

Exhibit the matrix

$$A = \sum_{i=1}^{n} \vec{e}_i \, \lambda_i \, \underset{\sim}{\omega}^i$$

STEP IV

Validate the correctness of this representation by applying A to the vector

$$\vec{x} = \sum_{j=1}^{n} \vec{e}_j \langle \underset{\sim}{\omega}^j \,|\, \vec{x} \rangle = \sum_{j=1}^{n} \vec{e}_j \, x^j$$

we have

$$A(\vec{x}) = \sum_i \vec{e}_i \lambda_i \underset{\sim}{\omega}^i(\vec{x})$$

$$= \sum_i e_i \lambda_i \langle \omega^i | \vec{e}_j x^j \rangle$$

$$= \sum_i e_i \lambda_i \underbrace{\langle \omega^i | e_j \rangle}_{\delta^i_j} x^j$$

$$= \sum_{i=1}^{n} e_i \lambda_i x^i$$

If $\vec{x}$ is one of the eigenvectors, say

$$\vec{x} = \vec{e}_k,$$

then

$$A \vec{e}_k = \sum_i \vec{e}_i \lambda_i \underbrace{\langle \underset{\sim}{\omega}^i | \vec{e}_k \rangle}_{\delta^i_k}$$

$$\boxed{A\vec{e}_k = \vec{e}_k \lambda_k} \quad (\text{no sum over } k)$$

Thus $\boxed{A = \sum_{i=1}^{n} \vec{e}_i \lambda_i \underset{\sim}{\omega}^i}$

mathematizes the diagram on page 33.5 indeed.

5. The eigenvectors $\{e_j\}$,

$$A\vec{e}_j = \lambda_j \vec{e}_j \qquad i = 1, \dots, n$$

are <u>right</u> eigenvectors. By contrast $\{\underset{\sim}{\omega}^k\}$ are left eigencovectors. Indeed,

$$\omega^k A = \langle \underset{\sim}{\omega}^k | \sum_i e_i \lambda_i \underset{\sim}{\omega}^i$$

$$= \sum_i \langle \underset{\sim}{\omega}^k | e_i \rangle \lambda_i \underset{\sim}{\omega}^i$$

$$\therefore \boxed{\underset{\sim}{\omega}^k A = \lambda_k \underset{\sim}{\omega}^k}$$

<u>Comment</u>

Notice that the decomposition of A,

$$A = \sum_{i=1}^{n} \vec{e}_i \lambda_i \underset{\sim}{\omega}^i$$

and hence the eigenvalue eq'n

$$A\vec{x} = \lambda \vec{x}$$

did not in any way depend on the existence of an inner product. The only necessary incredient was the construction of the dual eigen bases $\{\underset{\sim}{\omega}^i\}$ from the already known eigenbasis $\{\vec{e}_j\}$ of A.

Notabene: The existence of each such eigenvector is guaranteed by THEOREM 23.1 on page 23.4.

II. DIAGONALIZABLE MATRICES

It is not obvious that all matrices are diagonalizable. In fact it is not even true as exemplified by the ("defective") matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, whose characteristic polynomial $(1-\lambda)^2 = 0$ has degenerate roots. However if the roots are distinct, then there exists a corresponding set of eigenvectors. This fact is made precise by the following

<u>Theorem 33.2</u>

A matrix with n distinct eigenvalues has a set of n eigenvectors which is linearly independent.

Because of this, one has the following

<u>Corollary</u>

A matrix with distinct eigenvalues can be diagonalized.

<u>Comment</u>

1. <u>The converse is <u>not</u> true:</u>
   $\exists$ many many diagonalizable matrices with degenerate eigenvalues.
   For example, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
2. The Theorem does not apply to "defective" matrices

Proof of the linear independence of the set of n eigenvectors having n distinct eigenvalues.

<u>Step I</u>

Consider

$$c_1 \vec{e}_1 + c_2 \vec{e}_2 = \vec{0} \qquad (1)$$

Apply A to both sides and obtain

$$c_1 \lambda_1 \vec{e}_1 + c_2 \lambda_2 \vec{e}_2 = 0$$

Multiply Eq.(1) by λ_2 and subtract. One obtains

$$c_1 (\lambda_2 - \lambda_1) \vec{e}_1 = 0$$

By hypothesis $\lambda_2 \neq \lambda_1$. Thus

$$c_1 = 0.$$

Introduce this result into Eq.(1), one finds

$$c_2 \vec{e}_2 = \vec{0}$$

Thus $c_2 = 0$. It follows that the set of eigenvectors in Eq.(1),

$$\boxed{\{\vec{e}_1, \vec{e}_2\} \text{ is a linearly independent set.}}$$

<u>Step II</u>

One now repeats this line of reasoning by considering the linear combination

$$c_1 \vec{e}_1 + c_2 \vec{e}_2 + c_3 \vec{e}_3 = \vec{0}$$

to conclude that

$\{e_1, e_2, e_3\}$ is a l.i. set.

<u>Step III.</u> More generally one assumes that

$$\{e_1, e_2, \dots, e_k\} \text{ is a lin. indep. set} \qquad (2)$$

and considers

$$c_1\vec{e}_1 + \dots + c_k\vec{e}_k + c_{k+1}\vec{e}_{k+1} = \vec{0} \qquad (3)$$

in order to show that

$\{e_1, e_2, \dots, e_k, e_{k+1}\}$ is a lin. indep. set.

This is certainly true for $k=1$ and $k=2$.

<u>Step IV.</u>

Apply A to Eq.(3) and obtain

$$c_1\lambda_1\vec{e}_1 + \dots + c_k\lambda_k\vec{e}_k + c_{k+1}\lambda_{k+1}\vec{e}_{k+1} = \vec{0}. \qquad (4)$$

Mply Eq.(3) by λ_{k+1}, subtract the result

from Eq.(4) and obtain

$$c_1(\lambda_1-\lambda_{k+1})\vec{e}_1 + \dots + c_k(\lambda_k-\lambda_{k+1})\vec{e}_k = \vec{0}$$

By hypotheses $\lambda_1 \neq \lambda_{k+1}, \dots, \lambda_k \neq \lambda_{k+1}$. Thus

Eq.(2) implies

$$c_1 = c_2 = \dots = c_k = 0.$$

Introduce this result into Eq.(3); thus

$$c_{k+1}\vec{e}_{k+1} = 0$$

Consequently, <u>all</u> coefficients vanish:

$$c_1 = \dots = c_k = c_{k+1} = 0$$

This means that

$\{e_1, e_2, \dots, e_k, e_{k+1}\}$ is also a lin. indep. set.

Step V

Repeat Steps <u>III</u> and <u>IV</u> until

one has exhausted all n eigenvalues.

The result is that

$\{\vec{e}_1, \dots, \vec{e}_n\}$ is a lin. indep. set.