Lecture 24

Normal Forms for Matrices over a Field F and Jordan Normal Form

Definition

Two n by n matrices A and B are similar if there exists an n by n invertible matrix P such that

$$B = PAP^{-1}$$

Two Fundamental Problems

1. Given a matrix A, find the "simplest" (e.g. most zero entries) matrix B such that A is similar to B

2. Given two matrices A and B give criteria for deciding when A and B are similar.

Lemma 1 If A and B are similar then $\det A = \det B$.

Proof If $B = PAP^{-1}$ then $\det(B) = \det(PAP^{-1})$

$$= \det(P)\det(A)\det(P^{-1})$$

But $\det(P^{-1}) = \dfrac{1}{\det(P)}$ $\square$

The Solution of Both Problems for Real Symmetric Matrices

In Lecture 20 we solved Problem 1 in case A is a real symmetric n by n matrix. Then A is similar to a diagonal matrix.

We restate the Theorem on page 16 of Lecture 20, changing P to P^{-1}.

Theorem 1 If A is a real symmetric n by n matrix then A is similar to a diagonal matrix

$$PAP^{-1} = \begin{pmatrix} \lambda_1, \lambda_2, & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \qquad (*)$$

The λ_i's, $1 \le i \le n$, are the eigenvalues of A.

The columns of P^{-1} are eigenvectors of A.

We now address Problem 2 for real symmetric matrices.

Theorem 2 Let A and B be real symmetric n by n matrices. Then A and B are similar if and only if they have the same eigenvalues with the same multiplicities possibly arranged in a different order.

Example

Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

Then $\det(A) = 2$ and $\det(B) = 4$ so A and B are not similar by Lemma 1 even though they have the same eigenvalues 1 and 2. But if we keep track of multiplicities, the eigenvalues of A are $(1,1,2)$ and the eigenvalues of B are $(1,2,2)$ so not equal.

Proof of Theorem 2 We first prove that A, B similar implies they have the same eigenvalues with the same multiplicities.

Suppose $B = PAP^{-1}$. Suppose v is an eigenvector of A corresponding to the eigenvalue λ. Then

$$BPv = (PAP^{-1})(Pv) = PA(P^{-1}P)v = PAv = P(\lambda v) = \lambda Pv$$

Hence Pv is an eigenvector of B corresponding to the same eigenvalue λ and the map $v \longrightarrow Pv$ maps the eigenvectors of A to the eigenvectors of B without changing the associated eigenvalues. But also $A = P^{-1}BP$ and $w \longrightarrow P^{-1}w$ maps the eigenvectors of B to the eigenvectors of A without changing the associated eigenvalues. The two maps are clearly inverses of each other. So we have constructed a one-to-one correspondence between a basis of $\mathbb{R}^n$ consisting of eigenvectors of A and a basis of $\mathbb{R}^n$ consisting of eigenvectors of B which does not change the corresponding eigenvalues. Hence A and B have the same eigenvalues with the same multiplicities.

Now suppose A and B have the same eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with the same multiplicities. For each λ_i, $1 \le i \le n$, let $u_i \in \mathbb{R}^n$ be an eigenvector of A associated to the eigenvalue λ_i so $Au_i = \lambda_i u_i$, $1 \le i \le n$, and $v_i \in \mathbb{R}^n$ be an eigenvector of B associated to the eigenvalue λ_i so $Bv_i = \lambda_i v_i$, $1 \le i \le n$

We may choose the u_i's and v_i's so that $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$ are (orthonormal) bases of $\mathbb{R}^n$. Then there is a unique linear transformation $P: \mathbb{R}^n \to \mathbb{R}^n$ such that $Pu_i = v_i$, $1 \le i \le n$. We claim $B = PAP^{-1}$.

In order to prove that two linear transformations are equal it suffices to prove they coincide on a basis. Now $Bv_i = \lambda_i v_i$, $1 \le i \le n$. Also,

$$(PAP^{-1})(v_i) = PA(P^{-1}v_i) = P(Au_i) = P(\lambda_i u_i) = \lambda_i P(u_i) = \lambda_i v_i.$$

Hence $PAP^{-1} = B$ and Theorem 2 follows. $\square$

Remark The same proof as that of Theorem 2 proves the following generalization to diagonalizable matrices with entries from any field F, so in particular for $F = \mathbb{C}$. So we have

Theorem 3

Two diagonalizable matrices over any field F are similar if and only if they have the same eigenvalues with the same multiplicities.

Jordan Normal Form for Complex Matrices

It is remarkable that Problems 1 and 2 from page 1 are solved for all n by n complex matrices. Not every matrix over $\mathbb{C}$ is diagonalizable so we need to generalize the notion of a diagonal matrix.

Definition An n by n matrix A is block diagonal if A is of the form

$$A = \begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & A_k \end{pmatrix}$$

We assume all blocks are smaller than n by n (otherwise any n by n matrix would be block diagonalizable with one n by n block).
A diagonalizable matrix is a very special case with all blocks 1 by 1.

Definition An n by n matrix A is said to be in Jordan normal form if it is block diagonal (even with only one block)

$$A = \left(\begin{array}{c|c|c|c|c} J_1 & 0 & 0 & 0 & 0 \\ \hline 0 & J_2 & 0 & 0 & 0 \\ \hline 0 & 0 & \cdot & 0 & 0 \\ \hline 0 & 0 & 0 & \cdot & 0 \\ \hline 0 & 0 & 0 & 0 & J_m \end{array}\right)$$

and each block J_i, $1 \leq i \leq m$, is a Jordan block with parameter λ_i, $1 \leq i \leq m$, according to the definition on the next page.

Example (see page 13)

$$\left(\begin{array}{c|c|c} \begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix} & 0 & 0 \\ \hline 0 & \begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix} & 0 \\ \hline 0 & 0 & \begin{matrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{matrix} \end{array}\right)$$

The Theory of a Jordan Block

Definition A k by k matrix is said to be a Jordan block with parameters λ and k (or a Jordan block of length k) if

$$J = \underbrace{\begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & 0 & \\ & & \lambda & 1 & & \\ & & & \ddots & & \\ & 0 & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}}_{k} \Big\} k$$

The number k is said to be the length of the block.

Note ~~that~~ we may regard J as the matrix of a linear transformation T acting on a k-dimensional vector space V with a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ such that

$$\begin{aligned}
Tv_1 &= \lambda v_1 & &\text{or} & (T-\lambda I)v_1 &= 0 \\
Tv_2 &= v_1 + \lambda v_2 & &\text{or} & (T-\lambda I)v_2 &= v_1 \\
Tv_3 &= v_2 + \lambda v_3 & &\text{or} & (T-\lambda I)v_3 &= v_2 \\
&\vdots & &\vdots & &\vdots \\
Tv_k &= v_{k-1} + \lambda v_k & &\text{or} & (T-\lambda I)v_k &= v_{k-1}
\end{aligned}$$

It is important to note that the basis vector v_1 corresponding to the upper left entry in the Jordan block is an eigenvector of T.

We may represent the second set of equations on the previous page pictorially by

$$0 \xleftarrow{T-\lambda I} v_1 \xleftarrow{T-\lambda I} v_2 \xleftarrow{T-\lambda I} v_3 \xleftarrow{T-\lambda I} \cdots \xleftarrow{T-\lambda I} v_{k-2} \xleftarrow{T-\lambda I} v_{k-1} \xleftarrow{T-\lambda I} v_k$$

Proposition

$$v_j = (T-\lambda I)^{k-j} v_k \,, \quad 1 \le j \le k \qquad (*)$$

Proof We prove the proposition by descending induction on j. It is true for $j = k-1$. Suppose it is true for $j+1$. Now we prove it is true for j.

$$v_j = (T-\lambda I)\, v_{j+1} = (T-\lambda I)\Big[\underbrace{(T-\lambda I)^{k-(j+1)} v_k}_{\text{induction hypothesis}}\Big]$$

so

$$v_j = (T-\lambda I)^{k-j} v_k$$

$\square$

Remark on descending induction

If $(*)$ is not true then there exists a largest j between 1 and k for which it is not true. So it is true for $j+1$. But we just proved $(*)$ true for $j+1$ implies $(*)$ true for j. Contradiction.

Corollary

We may rewrite the above basis as

$$\mathcal{B} = \{(T-\lambda I)^{k-1} v_k, (T-\lambda I)^{k-2} v_k, \ldots, (T-\lambda I)v_k, v_k\}$$

So you get every vector in $\mathcal{B}$ by applying (larger and larger) powers of $T-\lambda I$ to the last basis vector v_k

Hence we obtain

Proposition

$$V = \text{span}(\{(T-\lambda I)^j v_k, 1 \leq j \leq k-1\})$$

Hence v_k is a cyclic vector for $S = T - \lambda I$ acting on V according to the following definition

<u>Definition</u> Let $S \in L(V,V)$. Then $v \in V$ is a cyclic vector for S if for some m, $V = \text{span}(v, Sv, S^2v, \ldots, S^m v)$

The following proposition is very important. We have just proved (1) and (2)

Jordan Block

<u>Proposition JB</u> Suppose the matrix of $T \in L(V,V)$ relative to the basis $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ is a Jordan block. Then

(1) The first basis vector v_1 in $\mathcal{B}$ is an eigenvector for T

(2) The last basis vector v_k is a cyclic vector for $T - \lambda I$ acting on V

(3) $(T - \lambda I)^\ell \equiv 0$, $\ell \geq k$.

Proof (of (3))

We will look at the case $k=3$
Consider the picture from page 9

$$0 \xleftarrow{T-\lambda I} v_1 \xleftarrow{T-\lambda I} v_2 \xleftarrow{T-\lambda I} v_3$$

So $T-\lambda I$ moves each vector on the line to the left one step closer to zero. So $(T-\lambda I)^2$ moves each vector on the line two steps closer to zero and $(T-\lambda I)^3$ moves each vector on the line to zero.

$$(T-\lambda I)^3 v_3 = (T-\lambda I)^2 v_2 = (T-\lambda I) v_1 = 0$$

Once you get to zero you stay at zero, $(T-\lambda I)0 = 0$. $\square$

I leave the general case to you.

Corollary If v_k is a cyclic vector for $T-\lambda I$ then $v_1 = (T-\lambda I)^{k-1} v_k$ is an eigenvector for T so (2) $\Rightarrow$ (1)

Proof $$(T-\lambda I)\Big[\underbrace{(T-\lambda I)^{k-1} v_k}_{v_1}\Big] = (T-\lambda I)^k v_k = 0$$

so

$$(T-\lambda I)v_1 = 0 \quad \text{so} \quad Tv_1 = \lambda v_1$$

$\square$

Example

$$A = \left(\begin{array}{c|c|c} \begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix} & 0 & 0 \\ \hline 0 & \begin{matrix} 3 & 1 \\ 0 & 3 \end{matrix} & 0 \\ \hline 0 & 0 & \begin{matrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{matrix} \end{array}\right)$$

We assume A is the matrix of a linear transformation T relative to a basis $\mathcal{B} = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. We will say $\mathcal{B}$ is a Jordan basis for T (see the definition on page 15). Let $\mathcal{B}_1 = \{v_1, v_2\}$, $\mathcal{B}_2 = \{v_3, v_4\}$ and $\mathcal{B}_3 = \{v_5, v_6, v_7\}$. Then $\mathcal{B}_1$ is a basis belonging to the first Jordan block, $\mathcal{B}_2$ is a basis belonging to the second Jordan block and $\mathcal{B}_3$ is a basis belonging to the third Jordan block. Hence v_1, v_3 and v_5 are eigenvectors belonging to the eigenvalues $3, 3$ and 5 respectively by Proposition JB. So the multiplicity of 3 as an eigenvalue for T is 2 (equals the number of Jordan 3-blocks) and the multiplicity of 5 as an eigenvalue of T is 1 (there is one Jordan 5-block).

We represent the partitioning of $\mathcal{B}$ schematically by

$$\mathcal{B} = \{ \underbrace{v_1, v_2}_{\mathcal{B}_1}, \underbrace{v_3, v_4}_{\mathcal{B}_2}, \underbrace{v_5, v_6, v_7}_{\mathcal{B}_3} \}$$

Put $U_1 = \mathrm{span}(\mathcal{B}_1)$, $U_2 = \mathrm{span}(\mathcal{B}_2)$, $U_3 = \mathrm{span}(\mathcal{B}_3)$. Then we have a direct sum decomposition $V = U_1 \oplus U_2 \oplus U_3$. We claim that this direct sum decomposition is invariant under T. But from top of page 13 we see that A, the matrix of T relative to the basis $\mathcal{B}$ is block diagonal with blocks of sizes 2 by 2, 2 by 2 and 3 by 3. But this structure of the matrix of T is equivalent to the direct sum decomposition being invariant under T.

On the previous page we said "$\mathcal{B}$ was a Jordan basis for T". The next definition will justify this

Definition Let $T \in L(V,V)$.
Then a basis $\mathcal{B}$ for V is a Jordan basis relative to A if

(1) There is an T-invariant direct sum decomposition
$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_k \quad \text{with } \dim U_j = n_j,\ 1 \le j \le k$$

(2) The basis $\mathcal{B}$ is partitioned
$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$$
such that $\mathcal{B}_j$ is a basis for U_j, $1 \le j \le k$

(3) The matrix of $T|U_j$ relative to $\mathcal{B}_j$ is a Jordan λ_j-block of length n_j for some complex number λ_j

Existence and Uniqueness of Jordan Normal Form

Theorem J

1. Let A be a complex n by n matrix. Then there exists an invertible n by n matrix P such ~~that~~ PAP^{-1} has Jordan normal form.
Equivalently there is a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ such that if we rewrite A in terms of the new basis $\mathcal{B}$ the resulting ~~matrix~~ B will be in Jordan normal form so $\mathcal{B}$ is a Jordan basis for B.

2. Two complex n by n matrices A and B are similar if and only if their associated Jordan normal forms have the same Jordan blocks with the same multiplicities possibly arranged in a different order.

Proof
This theorem is very hard. It will be proved in the last two lectures.

Permuting the blocks

Here is a simple example of two different Jordan normal forms that are similar namely

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ and } \begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

First note that (check it)

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

The matrix $P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ is the matrix of the permutation σ of the standard basis vectors such that

$$\sigma(e_1) = e_3, \quad \sigma(e_2) = e_2, \quad \sigma(e_3) = e_1$$

So it has order 2, that is, $P^2 = I$ so

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

so the two matrices above are similar