Matrix Theory
Lecture Notes from September 4, 2025

taken by Joel Sleeba

1 Warm Up

Assume we know that $A \in M_n(\mathbb{C})$ is diagonalizable. Let $p_0, p_1, p_2, \ldots, p_n \in \mathbb{C}$ and consider

$$B := P(A) = p_0 I + p_1 A + p_2 A^2 + \ldots + p_n A^n$$

Is B diagonalizable?

Proof. Yes. Because if $A = S^{-1}DS$ for a diagonal D, then $A^n = S^{-1}D^nS$, and therefore

$$B = S^{-1}(p_0 I + p_1 D + p_2 D^2 + \ldots + p_n D^n)S$$

Since D is a diagonal matrix, $p_0 I + p_1 D + p_2 D^2 + \ldots + p_n D^n$ will also be a diagonal matrix. Hence we get that B is diagonalizable. □

2 Simultaneous diagonalization

Theorem 2.1. *Let A, B be diagonalizable. Then $AB = BA$ if and only if they are simultaneously diagonalizable by the same S.*

Proof. Let $D_A = S^{-1}AS$, and $B' = S^{-1}BS$, where D_A is a diagonal matrix. Without loss of generality, assume that common eigenvalues appear together in D_A. If not choose S with an additional permutation of the rows.

Assuming $AB = BA$, we get

$$\begin{aligned} D_A B' &= S^{-1}ASS^{-1}BS \\ &= S^{-1}ABS \\ &= S^{-1}BAS \\ &= S^{-1}BSS^{-1}AS \\ &= B'D_A \end{aligned}$$

If $B' = [b'_{i,j}]_{i,j=1}^n$, then by $D_A B' = B' D_A$, from the diagonal structure of D_A, we get

$$\tilde{\lambda}_i b'_{i,j} = b'_{i,j} \tilde{\lambda}_j$$

where $\tilde{\lambda}_i$ is the i-th diagonal entry on D_A. So, we have

$$(\tilde{\lambda}_i - \tilde{\lambda}_j)b'_{i,j} = 0$$

which shows that if $\tilde{\lambda}_i \neq \tilde{\lambda}_j$, then $b'_{i,j} = 0$. Thus we get that

$$B' = \begin{bmatrix} B'_1 & & & \\ & B'_2 & & \\ & & \ddots & \\ & & & B'_r \end{bmatrix}$$

Since B and B' are diagonalizable, so is each B'_i. Taking matrices $T_1, T_2, \dots, T_r$ that diagonalize $B'_1, B'_2, \dots, B'_r$ respectively, let

$$T = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_4 \end{bmatrix}$$

Then,

$$T^{-1}BT = \begin{bmatrix} T_1^{-1}B'_1T_1 & & & \\ & T_2^{-1}B'_2T_2 & & \\ & & \ddots & \\ & & & T_r^{-1}B'_rT_r \end{bmatrix} = \begin{bmatrix} D'_1 & & & \\ & D'_2 & & \\ & & \ddots & \\ & & & D'_r \end{bmatrix}$$

where each D'_i is a diagonal block. Also,

$$T^{-1}D_AT = \begin{bmatrix} T_1^{-1}\lambda_1 IT_1 & & & \\ & T_2^{-1}\lambda_2 IT_2 & & \\ & & \ddots & \\ & & & T_r^{-1}\lambda_r IT_r \end{bmatrix} = D_A$$

This implies $D_A = T^{-1}S^{-1}AST$, and $D_B = T^{-1}S^{-1}BST$ are both diagonal.

Conversely if A, B are diagonalizable by the same S, we have $A = SD_AS^{-1}$ and $B = S^{-1}D_BS$. Then

$$\begin{aligned} AB &= S^{-1}D_ASS^{-1}D_BS \\ &= S^{-1}D_AD_BS \\ &= S^{-1}D_BD_AS \\ &= S^{-1}D_BSS^{-1}D_AS \\ &= BA \end{aligned}$$

And thus we are done. $\square$

Next, we consider simultaneous diagonalization for a family of matrices.

Definition 2.1. A family $F \subset M_n$ is a commuting family if for each $A, B \in F$, $AB = BA$.

Definition 2.2. A subspace $W \subset \mathbb{C}^n$ is called an A-invariant subspace for some $A \in M_n$, if $Aw \in W$ for all $w \in W$. If $F \subset M_n$, then W is called F-invariant if for each $A \in F$, W is A-invariant.

Lemma 2.1. *If $W \subset \mathbb{C}^n$ is A-invariant for some $A \in M_n$, and suppose that dim$(W) \geq 1$, then there is an $x \in W \setminus \{\mathbf{0}\}$ such that $Ax = \lambda x$.*

Proof. We consider $B := A|_W$, the matrix representation of A restricted to the subspace W. Then $B : W \to W$ has an eigenvector since it has at least one eigenvalue as the characteristic polynomial $p_B(x)$ decomposes into linear factors by the fundamental theorem of algebra. □

Lemma 2.2. *If $F \subset M_n$ is a commuting family, then there exists an $x \in \mathbb{C}^n$ such that for each $A \in F$, $Ax = \lambda_A x$.*

Proof. Choose W to be an F-invariant subspace of minimum, non-zero dimension. Existence of W is guaranteed since $\mathbb{C}^n$ is an F-invariant subspace of non-zero dimension.

Next, we show that any $x \in W \setminus \{\mathbf{0}\}$ is an eigenvector for each $A \in \mathbb{F}$. Assume this is not true. Then there is a $y \in W \setminus \{\mathbf{0}\}$, and an $A \in F$, such that $Ay \notin \mathbb{C}y$. Since W is A-invariant by the setup, by previous lemma, we get that there is a $x \in W \setminus \{\mathbf{0}\}$ such that $Ax = \lambda_x x$ for some $\lambda_x \in \mathbb{C}$.

Let $W_0 := \{z \in W \ : \ Az = \lambda_x z\}$. Since $y \notin W_0$, we get that $W_0 \neq W$. But for any $B \in F$, by the invariance of W, $Bx \in W$, and for $u \in W_0$,

$$A(Bu) = B(Au) = \lambda_x Bu$$

We observe $Bu \in W_0$. Thus B maps W_0 to W_0, so W_0 is F-invariant. We have derived a contradiction with the minimality of W, proving our statement. □

Remark 2.1. This implies that commuting families have at least one common eigenvector

Definition 2.3. A simultaneously diagonalizable family is a family $F \subset M_n$ such that there exists $S \in M_n$ for which $S^{-1}AS$ is diagonal for each $A \in F$

Theorem 2.2. *Let $F \subset M_n$ be a family of diagonalizable matrices, then F is a commuting family if and only if it is simultaneously diagonalizable.*

We will prove this in the next lecture.