Matrix Theory
Lecture Notes from September 4, 2025

taken by Joel Sleeba

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?

1.7.29 Answer. Yes. Let $S \in M_n(\mathbb{C})$ be invertible such that $A = S^{-1} D S$ for a diagonal matrix $D \in M_n(\mathbb{C})$. Then $A^n = S^{-1} D^n S$ and

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

Since D is a diagonal matrix and the product of diagonal matrices are diagonal, D^n is also diagonal. Then $P(D) = p_0 I + p_1 D + p_2 D^2 + \ldots + p_n D^n$ will also be a diagonal matrix. Hence we get that B is diagonalizable.

In fact we get more, we get that B is diagonalizable by the same $S \in M_n(\mathbb{C})$ which diagonalized A. In this lecture we will be investigating the conditions on B to be diagonalized by the same matrix S which diagonalized A.

1.7.30 Remark (Easter egg). If $A, B \in M_n(\mathbb{C})$ are diagonalizable by the same S as in the example before, is there a polynomial $P \in \mathbb{C}[x]$ such that $B = P(A)$?

1.7.31 Answer. (Hint) No. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Find other examples.

1.8 Simultaneous diagonalization

1.8.32 Definition. Let $A, B \in M_n(\mathbb{C})$ be diagonalizable. We say that A and B are simultaneously diagonalizable if there exists an invertible matrix $S \in M_n(\mathbb{C})$ such that $A = S^{-1} D_A S$ and $B = S^{-1} D_B S$, where $D_A, D_B \in M_n(\mathbb{C})$ are diagonal matrices.

1.8.33 Theorem. *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. That is, if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of D_A, we are assuming that

$$D_A = \begin{bmatrix} \lambda_1 & & & & & & & & & & \\ & \lambda_1 & & & & & & & & & \\ & & \ddots & & & & & & & & \\ & & & \lambda_1 & & & & & & & \\ & & & & \lambda_2 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & \lambda_2 & & & & \\ & & & & & & & \lambda_3 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & \lambda_k \end{bmatrix}$$

If not, choose S with required additional permutations 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 is diagonalizable, by the definition of B' it follows that B' is diagonalizable. We claim that each B'_r themselves are diagonalizable. Considering B' as a linear map $\mathbb{C} \to \mathbb{C}^n$, from the block structure of B' we see that there are subspaces $W_i \subset \mathbb{C}^n$ such that

$$\mathbb{C}^n = \bigoplus_{i=1}^{r} W_i$$

and $B'(W_i) \subset W_i$. Moreover observe that B'_i is the matrix representation of the linear map B' restricted to W_i. Since B' is diagonalizable, by our characterization there is a basis of $\mathbb{C}^n$ consisting of eigenvectors of B'. Since B'_i is the restriction of B' to W_i, eigenvectors of B' which are in the subspace W_i are eigenvectors of B'_i themselves. And they form a basis for W_i as $\mathbb{C}^n$ is the direct sum of W_is. Thus from our characterization, we get that each B'_i is diagonalizable.

Taking matrices $T_1, T_2, \ldots, T_r$ that diagonalize $B'_1, B'_2, \ldots, B'_r$ respectively, let

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

Then,

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

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

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

Thus for $Q = ST$, we get that $Q^{-1}AQ = T^{-1}S^{-1}AST = D_A$, and $Q^{-1}BQ = T^{-1}S^{-1}BST = D_B$ are both diagonal, proving that A and B are simultaneously diagonalizable.

Conversely, if A and 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.

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

1.8.35 Definition. 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.

1.8.36 Lemma. *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. Since the subspace W is A invariant, A as a linear transformation restricted to W, $A|_W : W \to W$ has a matrix representation $B \in M_r$, where $r < n$. B has an eigenvector since it has at least one eigenvalue λ as the characteristic polynomial $p_B(x)$ decomposes into linear factors by the fundamental theorem of algebra. Let x be the corresponding eigenvector in W such that $Bx = \lambda x$. Now considering B as the restriction of A to W, we see that $Ax = \lambda x$. Hence we are done. □

1.8.37 Lemma. *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 $W_0 \subsetneq W$. But for any $B \in F$, by the invariance of W, $Bx \in W$. Then for $u \in W_0$,

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

and since $Bu \in W$ and it satisfies the description of the set W_0, we observe $Bu \in W_0$. Thus B maps W_0 to W_0. Since $B \in F$ was arbitrary this shows that W_0 is F-invariant. Hence have derived a contradiction with the minimality of W, proving our statement. □

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

1.8.39 Definition. 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 $A \in F$

1.8.40 Theorem. *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.