Matrix Theory
Lecture Notes from September 2, 2025

taken by Joel Sleeba

1 Warm Up

Let $A, B \in M_n(\mathbb{C})$, and A be an invertible matrix. Consider the function $F : \mathbb{R} \to M_n(C) := t \to A + tB$. Since F is continuous, and $F(0) = A$ is invertible, by the inverse function theorem, F is invertible in some neighborhood $B_\varepsilon(0)$ of 0. Let $G : B_\varepsilon(0) \to M_n(\mathbb{C})$, be the inverse. Clearly $G(0) = A^{-1}$. We are interested to have a good approximation G.

If we can find G', the derivative of G, $H(t) = G(0) + tG'(t)$ would be a good approximation for G. We know that $F(t)G(t) = I_n$, then using the product rule of differentiation,

$$\begin{aligned} F(t)G'(t) + F'(t)G(t) &= 0 \\ F(0)G'(0) &= -F'(0)G(0) \\ AG'(0) &= -BA^{-1} \\ G'(0) &= -A^{-1}BA^{-1} \end{aligned}$$

Thus we get that $H(t) = A - tA^{-1}BA^{-1}$ is a good approximation for G.

2 Conditions for Diagonazability

Now we look for some more conditions for diagonalizability.

Theorem 2.1. *Let $A \in M_n(\mathbb{C})$, with its characteristic polynomial $p_A(t) = \prod_{j=1}^{n}(t - \lambda_j)$, and $\lambda_i \neq \lambda_j$ for $j \neq k$, then A is diagonalizable.*

Proof. We'll show that there's a linearly independent set of n eigenvectors. Then by what we've proved in the last lecture, we'll be done. Let $\{x_1, x_2, \ldots, x_n\}$ be such that $x_j \in \mathbb{C}^n$ with $Ax_j = \lambda_j x_j$. If $\{x_1, x_2, \ldots, x_n\}$ were linearly dependent, then there is a linear combination

$$\alpha_1 x_{j_1} + \alpha_2 x_{j_2} + \ldots + \alpha_s x_{j_s} = 0$$

with $s \leq n$, and all $\alpha_j \neq 0$. Let r be smallest such $s \leq n$, and assume with possible renumbering that $j_i = i$. Then applying A to the linear combination gives us

$$A(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n) = \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \ldots + \alpha_n \lambda_n x_n = 0$$

multiplying the previous equation with λ_r and then subtracting gives us

$$\begin{aligned} 0 &= (\alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2 + \ldots + \alpha_n\lambda_n x_n) - (\alpha_1\lambda_r x_1 + \alpha_2\lambda_r x_2 + \ldots + \alpha_r\lambda_r x_r) \\ &= \alpha_1(\lambda_1 - \lambda_r)x_1 + \alpha_2(\lambda_2 - \lambda_r)x_2 + \ldots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)x_{r-1} + \alpha_r(\lambda_r - \lambda_r)x_r \\ &= \alpha_1(\lambda_1 - \lambda_r)x_1 + \alpha_2(\lambda_2 - \lambda_r)x_2 + \ldots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)x_{r-1} \end{aligned}$$

which contradicts the minimality of r. □

Unfortunately this is just a sufficient condition, as in the next example.

Example 2.1. Consider the matrix

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Clearly A is diagonalizable. But the characteristic polynomial $p_A(x) = x^2(1-x)$ does not satisfy the conditions of Theorem 2.1

Definition 2.1. If for $A \in M_n(\mathbb{C})$, with characteristic polynomial

$$p_A(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\ldots(t-\lambda_r)^{m_r}$$

then we say that the eigenvalue λ_j has algebraic multiplicity m_j. We call $\text{null}(\lambda_j I - A)$, the geometric multiplicity of λ_j

Lemma 2.1. *If $A \in M_n$ has eigenvalue λ, and characteristic polynomial $p_A(t) = (t-\lambda)^m q(t)$, with $q(\lambda) = 0$, then $r = nul(\lambda I - A)) \leq m$*

Proof. Choose a basis $\{x_1, x_2, \ldots, x_r\}$ of eigenvectors, spanning $E_\lambda = \{x \in \mathbb{C}^n : Ax = \lambda x\}$. Complete it to a basis $\{x_1, x_2, \ldots, x_n\}$ of $\mathbb{C}^n$. Let $S = [x_1, x_2, \ldots, x_n]$.

Then $AS = [\lambda x_1, \lambda x_2, \ldots \lambda x_r, y_{r+1}, \ldots y_n]$ with some vectors $y_{r+1}, \ldots, y_n$. Then

$$S^{-1}AS = \begin{bmatrix} \lambda I_r & 0 \\ 0 & C \end{bmatrix}$$

and we get

$$\begin{aligned} \det(tI - A) &= \det(tI - S^{-1}AS) \\ &= (t-\lambda)^r \det(t - C) \end{aligned}$$

Thus we conclude that algebraic multiplicity of λ is at least equal to r. □

Remark 2.1. See that the sum of all the algebraic multiplicity of the eigenvalues of $A \in M_n(\mathbb{C})$ is n. This is a direct consequence of the fundamental theorem of algebra.

Theorem 2.2. *The matrix $A \in M_n(\mathbb{C})$ is diagonalizable if and only if the algebraic and geometric multiplicities are equal for each eigenvalue.*

Proof. We note that given two eigenvalues $\lambda_j \neq \lambda_k$, then their eigenspaces E_i, E_j intersect trivially. Thus if $\{v_1, v_2, \ldots, v_{r_1}\}$ and $\{u_1, u_2, \ldots, u_{r_2}\}$ form a basis for E_{λ_1} and E_{λ_2} respectively, then $\{v_1, v_2, \ldots, v_{r_1}, u_1, u_2, \ldots, u_{r_2}\}$ is linearly independent. Iterating this way, we get a basis for $E_1 + E_2 \ \ldots + E_n$ with dimension $r = \sum_{i=1}^{k} r_i$.

If algebraic and geometric multiplicities equal then $r = n$, and we have a basis of eigenvectors. Otherwise if $r < n$, then we do not have such a basis of eigenvectors. And since existence of a basis of eigenvectors characterizes diagonalizability, this characterizes diagonalizability. $\square$

Next lecture, we'll look when multiple matrices can be simultaneously diagonalizable with the same S matrix.