MATH 6304 - Theory of Matrices

Joel Sleeba
University of Houston
joelsleeba1@gmail.com

September 11, 2025

Contents

26/08/2025 Office Hours : Tuesday 10 - 11 AM, Wednesday 1 - 2 PM

1 Introduction

Let A be a $m \times n$ matrix and D a diagonal $n \times n$ matrix with entries $d_1, d_2, \dots, d_n$,

$$A = \begin{bmatrix} | & | & \cdots & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & \cdots & | \end{bmatrix}$$

Then

$$AD = \begin{bmatrix} | & | & & | \\ d_1 a_1 & d_2 a_2 & \cdots & d_n a_n \\ | & | & & | \end{bmatrix}$$

and if

$$B = \begin{bmatrix} - & b_1 & - \\ - & b_2 & - \\ & \vdots & \\ - & b_n & - \end{bmatrix}$$

then

$$DB = \begin{bmatrix} - & d_1 b_1 & - \\ - & d_2 b_2 & - \\ & \vdots & \\ - & d_n b_n & - \end{bmatrix}$$

Do the same for upper triangular matrices and add some context to the multiplications.

Notice that every time you left multiply, you play with column, and when you right multiply, you play with the rows.

Exercise 1.1. Let A be an $n \times n$ matrix over $\mathbb{C}$. Let $\omega = e^{\frac{2\pi i}{n}}$. Then prove that

$$A' = \frac{1}{n} \sum_{k=0}^{n} (U^*)^k A U^k$$

preserve all the diagonal entries of A and kills the rest of entries. That is $A' = \text{Diag}(A)$

1.1 Review of Linear Algebra

- Rank-Nullity Theorem
- Orthogonality
- Orthogonal projection is the closest point on the subspace from the given vector.

28/08/2025

Definition 1.2. If $A \in M_n$, $A = (a_{i,j})_{i,j=1}^n$, we let

$$\mathrm{trace(A)} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{a_{jj}}$$

The determinant is

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$$

Remark 1.3. If $A = [a_1 a_2 \dots a_n]$, then $det(A) = f(a_1, a_2, \dots, a_n)$ is the only function that is linear in each a_i, alternating (swapping columns doesn't alter the value), and normalized ($det(I) = 1$).

This is useful to show that for $A, B \in M_n$, $det(AB) = det(A)det(B)$.

Moreover if $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, then $det(A) = det(B)det(D)$.

We also have

$$det(A) = \sum_{i,j=1}^{n} (-1)^{i+j} a_{i,j} det(A_{i,j})$$

Where $A_{i,j}$ is the submatrix with ith row and jth column removed from A.

1.2 Eigenvalues and Eigenvectors

Definition 1.4. Eigenvalue, Eigenvector, Spectrum of a matrix

1.3 Similarity

Definition 1.5. A matrix $B \in M_n$ is similar to $A \in M_n$, if there is an invertible $S \in M_n$ such that $B = S^{-1}AS$. This defines an equivalence relation.

Theorem 1.6. *If $A, B \in M_n$ are similar. Then their characteristic polynomial $P_A = P_B$.*

Remark 1.7. characteristic polynomial i snot characteristic upto similarity, because

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

02/09/2025

1.4 Diagonazability

Can we find conditions for diagonalizability.

Theorem 1.8. *Let $A \in M_n(\mathbb{C})$, $p_A(t) = \prod_{j=1}^{n}(t - \lambda_j)$, 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. Let $x_j \in \mathbb{C}^n$ such that $Ax_j = \lambda_j x_j$. If $\{x_1, x_2, \ldots, x_n\}$ were linearly dependent, then there is a linear combination

$$\alpha_1 x_{j_1} + \alpha_2 x_{j_2} + \ldots + \alpha_r x_{j_r} = 0$$

with $r \leq n$, and all $\alpha_j \neq 0$. Let r be smallest such $r \leq n$, 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

$$\alpha_1(\lambda_1 - \lambda_r)x_1 + \alpha_2(\lambda_2 - \lambda_r)x_2 + \ldots + \alpha_r(\lambda_r - \lambda_r)x_r = 0$$

which contradicts the minimality of r. $\square$

Unfortunately this is just a sufficient condition, as it excludes the following matrix.

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

Definition 1.9. If for $A \in M_n(\mathbb{C})$,

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

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

Lemma 1.10. *If $A \in M_n$ has eigenvalue λ, and $p_A(t) = (t-\lambda)^m q(t)$, with $q(\lambda) = 0$, then $r = nul(\lambda I - A)) \leq m$*

Proof. Choose a basis $\{x_1, x_2, \ldots, x_r\}$ of veignevectors, 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 $S^{-1}AS =$verify, 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 1.11. See that the sum of all the algebraic multiplicity of the eigenvalues of $A \in M_n(\mathbb{C})$ is n.

Theorem 1.12. *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. □

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

04/09/2025

Warm Up

Assume we know that A 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 $A = S^{-1}DS$, then $A^n = S^{-1}D^nS$, and therefore

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

□

1.5 Simultaneous diagonalization

Theorem 1.13. *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 $T_1, T_2, \ldots, T_r$ that diagonalize $B'_1, B'_2, \ldots, B'_r$ respectively, let

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

Then,

$$T^{-1}BT = \begin{bmatrix} T_1^{-1}B_1'T_1 & & & \\ & T_2^{-1}B_2'T_2 & & \\ & & \ddots & \\ & & & T_r^{-1}B_r'T_r \end{bmatrix} = \begin{bmatrix} D_1' & & & \\ & D_2' & & \\ & & \ddots & \\ & & & D_r' \end{bmatrix}$$

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.

Converse is left as an exercise $\square$

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

Definition 1.14. A family $F \subset M_n$ is a commuting family if for each $A, B \in F$, $AB = BA$.

Definition 1.15. 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 1.16. *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 $B := A|_W$. Then $B : W \to W$ has an eigenvector since it has atleast one eigenvalue by the fundamental theorem of algebra. $\square$

Lemma 1.17. *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 we can choose $W = \mathbb{C}^n$.

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 $\lambda \in \mathbb{C}$.

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

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

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. $\square$

Remark 1.18. This implies that commuting families have at least one common eigenvector

Definition 1.19. 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 $A \in F$

09/09/2025

Theorem 1.20. *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.*

Proof. It is an easy exercise to show that a simultaneously diagonalizable family is commuting.

We prove the converse by induction over n. For $n = 1$, there's nothing to prove. Assume that this is true for all $F' \subset M_k$, where $k < n$. If each $A \subset F$ is of the form $A = \lambda I$, again nothing to prove. Thus, assume A is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r, r \geq 2$, and assume $AB = BA$, for each $B \in F$. Then A, B are simultaneously diagonalizable by previous theorem and hence without loss of generality, assume that A is diagonal.

verify rest from the lecture notes $\square$

Remark 1.21. Given $C \in M_n(\mathbb{C})$, we can think of function associated with C, as $\langle x, y\rangle \to \langle Cx, y\rangle$. Notice that this function preserves all information about C, since we can find the individual matrix entries. We can also associate $Q_c : x \to \langle Cx, x\rangle$. We recall Q_c determines C as

$$\langle Cx, y\rangle = \frac{1}{4}\sum_{j=1}^{4} i^j \|Q_c(x + i^j y)\|^2, \quad (i = \sqrt{-1})$$

1.6 Hermitian, normal, and unitary matrices

Definition 1.22. Let $A \in M_{n,m}$, then the adjoint $A^* \in M_{m,n}$ satisfies $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for each $x \in \mathbb{C}^m, y \in \mathbb{C}^n$.

If for $A \in M_n$, $A = A^*$, then we say that A is Hermitian or self-adjoint. If $A = -A^*$, the it is called skew-hermitian.

If $A \in M_n$, then $A = B + iC$, where $B = \frac{1}{2}(A + A^*), C = \frac{1}{2i}(A - A^*)$ have the property $B = B^*, C = C^*$. Here B is called the real part of A, and iC is called the imaginary part of A.

Proposition 1.23.

$$A = A^* \iff (iA) = -(iA)^*$$

Proposition 1.24. *A matrix $A \in M_n$ is Hermitian if and only if for all $x \in \mathbb{C}^n$, $\langle Ax, x\rangle \in \mathbb{R}$.*

Proof. If A is Hermitian, then

$$\overline{\langle Ax, x\rangle} = \langle x, Ax\rangle = \langle A^*x, x\rangle = \langle Ax, x\rangle$$

shows that $\langle Ax, x\rangle \in \mathbb{R}$.

Conversely, assume that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. Let $A = B + iC$, where $B = B^*, C = C^*$. Then

$$\langle Ax, x\rangle = \underbrace{\langle Bx, x\rangle}_{\in\mathbb{R}} + i\underbrace{\langle Cx, x\rangle}_{\in\mathbb{R}}$$

We conclude that $\langle Cx, x\rangle = 0$ for all x. Now using polarization identity, we get that $C = 0$. $\square$

We also consider and equivalent of unimodular numbers.

Definition 1.25. Let $A \in M_n$. A is unitary if $A^*A = I = AA^*$.

Proposition 1.26. *Let $A \in M_n$. The following are equivalent.*

(1) A is unitary.

(2) The columns of A form an orthonormal basis.

(3) Rows of A form an orthonormal basis.

(4) A preserves the norm for each $x \in \mathbb{C}^n$. That is $\|Ax\| = \|x\|$.

(5) A preserve the inner product. That is $\langle Ax, Ay\rangle = \langle x, y\rangle$

Proof. You know, $\square$

11/09/2025

Warm up

(1) Let $A \in M_n(\mathbb{C})$. For each $x \in \mathbb{C}^n$, $\langle Ax, x\rangle = 0$ if and only if $A = 0$.

(2) Let $A \in M_n(\mathbb{R})$ and $A = A^*$, then for each $x \in \mathbb{R}^n$, $\langle Ax, x\rangle = 0$ if and only if $A = 0$

Remark 1.27. If A is Hermitian and unitary, then $AA^* = A^2 = I$, so each eigenvalue $\lambda = \pm 1$.

Example 1.28. For $y \in \mathbb{C}^n \setminus \{0\}$, $H = I - 2\frac{y^*y}{\|y\|^2}$ is called a Householder transform

Exercise 1.29. Check that this is Hermitian and satisfies $H^2 = I$.

2 Shur Triangulizatoin and consequences of unitary equivalence

Definition 2.1. We say that $A, B \in M_n(\mathbb{C})$ are unitarily equivalent if there is $U \in M_n$, $U^*U = I$ such that $B = U^*AU$.

Remark 2.2. because of $U^{-1} = U^*$, if A and B are unitarily equivalent, then there are similar. The converse is not true in general.

Question 2.3. How can we tell if A, B are unitarily equivalent?

We first look for necessary conditions.

Proposition 2.4. *If A, B are unitarily equivalent, then*

$$\sum_{i,j=1}^{n} |a_{i,j}|^2 = \sum_{i,j=1}^{n} |b_{i,j}|^2$$

Proof. We observe that

$$\sum_{i,j=1}^{n} |a_{i,j}|^2 = \sum_{i=1}^{n} [AA^*]_{i,i} = \mathrm{tr}(AA^*)$$

The same is true for $B = U^*AU$. So,

$$\begin{aligned}\sum_{i,j=1}^{n} |b_{i,j}|^2 &= \mathrm{tr}(BB^*) \\ &= \mathrm{tr}(U^*AU(U^*AU)^*) \\ &= \mathrm{tr}(U^*AA^*U) \\ &= \mathrm{tr}(AA^*)\end{aligned}$$

□

We compare with a result that characterizes unitary equivalence

Theorem 2.5 (Carl Pearcy, 1962)**.** *If all words w_1 on the alphabet $\{A, A^*\}$, and w_2 on the alphabet $\{B, B^*\}$ of length upto $2n^2$ have*

$$tr(w_1) = tr(w_2)$$

then A, B are unitarily equivalent.

Theorem 2.6 (Shur's Triangulization Theorem)**.** *Let $AinM_n$ having eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, then thre is a unitary $U \in M_n$ such that $A = UTU^*$, where T is an upper triangular matrix with $\lambda_1, \lambda_2, \dots, \lambda_n$ in the diagonal.*

Remark 2.7. Remember that T is in genearl not unique. For example,

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Proof of Shur's Triangulization Theorem. We'll use induction on $n \in \mathbb{N}$. Assume this is true for matrices M_{n-1}. Given $A \in M_n$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Choose an eigenvector x corresponding to an eigenvalue λ_1, and take $\|x\| = 1$.

Applying the Gram-Schmidt process, we complement $\{x\}$ to an orthonormal basis $\{x, z_2, z_3, \dots, z_n\}$, Let $U_1 = [x, z_1, z_2, \dots, z_n]$. Then U_1 is unitary and $Ax = \lambda_1 x$, and

$$\begin{aligned} U_1^* A U_1 &= U_1^* [\lambda_1 x, *] \\ &= \begin{bmatrix} \lambda_1 & * \\ 0 & B \end{bmatrix} \end{aligned}$$

Now $B \in M_{n-1}$ and it follows from induction argument. □

Definition 2.8. A matrix $A \in M_n$ is called normal if $A^*A = A^*A$.

Example 2.9. Hermitian, skew-Hermitian, and Unitary matrices are normal.

Proposition 2.10. *A matrix $A \in M_n$ is normal if and only if for each $x \in \mathbb{C}^n$, $\|Ax\| = \|A^*x\|$.*

Proof. Normality is $AA^* - A^*A = 0$. Thus for all $x \in \mathbb{C}^n$, □