Bachelors with Mathematics as Major
3rd Semester
MMT322J: Mathematics/Applied Mathematics: Theory of Matrices

Credits: 4 THEORY + 2 TUTORIAL **Theory: 60 Hours & Tutorial: 30 Hours**

Course Objectives: (i) To understand matrix theory as a tool to solve various real life problems.
(ii) To make students aware about the properties and applications of matrices.

Course Outcome: After the completion of this course, students shall be able to (i) apply techniques of matrix theory to solve real life problems (ii) use matrix techniques in coding theory and cryptography (iii) use eigenvalues to find the stability of various systems.

Theory: 4 Credits

Unit –I

Generalization of reversal law of transpose, Hermitian and skew- Hermitian matrices, Representation of a square matrix as P + iQ, where P and Q are both Hermition. Adjoint of a matrix, For a square matrix A, A(adjA) = (adjA)A=|A|I. Commutative and associative laws in matrix operations. Necessary and sufficient condition for a square matrix to be invertible. Generalization of reversal law for the inverse of matrices under multiplication.

Unit –II

The operation of transposing and inverting are commutative, trace of a matrix, trace of AB= trace of BA and its generalization. Partitioning of matrices, Matrix polynomials and Characteristic equation of a square matrix. Cayley-Hamilton theorem, Eigen values and Eigen vector, minimal equation of a matrix.

Unit –III

Rank of a matrix. Elementary row (Column) transformations of a matrix do not alter its rank, rank of a matrix by elementary transformations, reduction of a matrix to the normal form, Elementary matrices. Every non- singular matrix is a product of elementary matrices, employment of only row (column) transformations. Rank of product of two matrices. Linear combination, linear dependence and linear independence of Row (Column) vectors, the columns of a matrix A are linearly dependent iff there exists a vector $X \neq 0$ such that AX=0. The columns of a matrix A of order $m \times n$ are linearly dependent iff rank of A< n. The matrix A has rank r if and only if it has r linearly independent columns and any s-columns (s > r) are linearly dependent (analogous results for rows).

Unit –IV

Linear homogeneous and non- homogeneous equations, the equation AX=0 has a non-zero solution if and only if rank of A < n, the number of its columns, the number of linearly independent solutions of the equation AX=0 is n-r, where r is the rank of matrix A of order $m \times n$, the equation AX=B is consistent if and only if two matrices A and [A: B] are of the same rank. Inner product of two vectors, length of a vector, normal vectors, Orthogonal and Unitary matrices, A matrix P is orthogonal (Unitary) if and only it its column vectors are normal and orthogonal in pairs.

Tutorial: 2 Credits

Unit –V

Problems based on Hermition and skew-Hermitian and inverse of matrices. Problems on characteristic roots and characteristic polynomials. Applications of Cayley Hamilton theorem for the inverse of a matrix.

Unit –VI

System of equations and their solutions, examples of determination of orthogonal matrices and examples of system of homogenous and non-homogenous equation having unique, infinite and no solution.

Books Recommended:

1. Shanti Narayan, A textbook of Matrices, Schaum S. Chand and Company, 1957.
2. A. Aziz and NA Rather and BA Zargar, Elementary Matrix theory, Kapoor Book Depot, Srinagar, 2007.
3. K. Hoffman and R. Kunze, Linear Algebra, Pearson Education, 2018.
4. S. Lipschutz & M. Lipson, Linear Algebra, Schaum"s outline series, Tata McGraw-Hill, 4th Edition 2009.

——THEORY OF MATRICES - (MMT322J)——

|||||||||||||| Author: DR. AIJAZ ||||||||||||||

ASSISTANT PROFESSOR OF MATHEMATICS

JK HIGHER EDUCATION DEPARTMENT

All stuyd materials from UG Mathematics onwards are avalaible at

Website: aijaz.vercel.app

Chapter 1

Generalizations of Operatoins on Matrices

1.1 Special Matrix Properties and Decompositions

1.1.1 Generalization of the Reversal Law of Transpose

Theorem 1.1.1 (Reversal Law for Two Matrices)**.** *If the product AB is defined, then $(AB)^T = B^T A^T$.*

Proof. To prove this, we compare the (i,j)-th element of each side. The (i,j)-th element of $(AB)^T$ is the (j,i)-th element of AB, which is $\sum_k a_{jk} b_{ki}$. The (i,j)-th element of $B^T A^T$ is the dot product of the i-th row of B^T and the j-th column of A^T. The i-th row of B^T is the i-th column of B, and the j-th column of A^T is the j-th row of A. This gives the product $\sum_k (B^T)_{ik}(A^T)_{kj} = \sum_k b_{ki} a_{jk}$. Since the resulting sums are identical, the theorem holds. ∎

This rule can be generalized to the product of any finite number of matrices.

Theorem 1.1.2 (Generalization of Reversal Law)**.** *If the product $A_1 A_2 \cdots A_n$ is well-defined, then:*

$$(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T$$

Proof. We will prove this theorem by mathematical induction on n, the number of matrices.
For $n = 2$, the theorem states $(A_1 A_2)^T = A_2^T A_1^T$. This is the standard reversal law for two matrices, which we take as established.
Assume the formula is true for a product of k matrices, where $k \geq 2$. That is, we assume:

$$(A_1 A_2 \cdots A_k)^T = A_k^T \cdots A_2^T A_1^T$$

We must now prove that the formula holds for a product of $k+1$ matrices. We can cleverly group the matrices as follows:

$$(A_1 A_2 \cdots A_k A_{k+1})^T = ((A_1 A_2 \cdots A_k) A_{k+1})^T$$

Let the product of the first k matrices be a single matrix, $B = (A_1 A_2 \cdots A_k)$. The expression becomes $(BA_{k+1})^T$. Applying the base case for two matrices gives:

$$(BA_{k+1})^T = A_{k+1}^T B^T$$

Now, we substitute back the expression for B:

$$A_{k+1}^T B^T = A_{k+1}^T (A_1 A_2 \cdots A_k)^T$$

By our inductive hypothesis, we can replace $(A_1 A_2 \cdots A_k)^T$ with its expanded form:

$$A_{k+1}^T (A_k^T \cdots A_2^T A_1^T) = A_{k+1}^T A_k^T \cdots A_2^T A_1^T$$

Thus, we have shown that if the law holds for k matrices, it also holds for $k+1$ matrices. By the principle of mathematical induction, the theorem holds for all integers $n \geq 2$. ∎

Example 1.1.3 (Generalization of Reversal Law of Transpose)**.** Verify the generalized reversal law $(ABC)^T = C^T B^T A^T$ for the matrices:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Solution. First, we compute the left-hand side, $(ABC)^T$. We calculate the product AB:

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 2 & 3 \end{pmatrix}$$

Next, we find the full product ABC:

$$(AB)C = \begin{pmatrix} 3 & 6 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 12 \\ 2 & 6 \end{pmatrix}$$

The transpose of this product is the left-hand side:

$$(ABC)^T = \begin{pmatrix} 3 & 2 \\ 12 & 6 \end{pmatrix}$$

Now, we compute the right-hand side, $C^T B^T A^T$. We first find the individual transposes:

$$A^T = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad B^T = \begin{pmatrix} -1 & 2 \\ 0 & 3 \end{pmatrix}, \quad C^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

We compute the product in reverse order, starting with $C^T B^T$:

$$C^T B^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & 6 \end{pmatrix}$$

Finally, we complete the product for the right-hand side:

$$(C^T B^T)A^T = \begin{pmatrix} -1 & 2 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 12 & 6 \end{pmatrix}$$

Since the left-hand side and right-hand side are identical, the generalized reversal law is verified for this example. ■

1.1.2 Hermitian and Skew-Hermitian Matrices

These matrices are the complex analogues of real symmetric and skew-symmetric matrices. Their definitions rely on the **conjugate transpose** (or Hermitian conjugate) of a matrix A, denoted A^*, which is defined as $A^* = (\overline{A})^T$.

Definition 1.1.4 (Hermitian Matrix)**.** A square matrix A is **Hermitian** if it is equal to its own conjugate transpose.

$$A^* = A$$

A key property of a Hermitian matrix is that its main diagonal entries must be real numbers, since for any diagonal entry a_{ii}, the condition $a_{ii} = \overline{a_{ii}}$ must hold.

Definition 1.1.5 (Skew-Hermitian Matrix)**.** A square matrix A is **skew-Hermitian** if it is equal to the negative of its conjugate transpose.

$$A^* = -A$$

A key property of a skew-Hermitian matrix is that its main diagonal entries must be purely imaginary or zero. This is because the condition $a_{ii} = -\overline{a_{ii}}$ implies that if $a_{ii} = x + iy$, then $x + iy = -(x - iy) = -x + iy$, which requires $2x = 0$, so $x = 0$.

Example 1.1.6 (Identifying Hermitian and Skew-Hermitian Matrices)**.** Determine if the matrices $A = \begin{pmatrix} 3 & 2-i \\ 2+i & -1 \end{pmatrix}$ and $B = \begin{pmatrix} i & 2+i \\ -2+i & 0 \end{pmatrix}$ are Hermitian, skew-Hermitian, or neither.

Solution. For matrix A, we compute its conjugate transpose $A^* = (\overline{A})^T$.

$$\overline{A} = \begin{pmatrix} 3 & 2+i \\ 2-i & -1 \end{pmatrix} \implies A^* = (\overline{A})^T = \begin{pmatrix} 3 & 2-i \\ 2+i & -1 \end{pmatrix}$$

Since $A^* = A$, the matrix A is **Hermitian**.

For matrix B, we compute its conjugate transpose $B^* = (\overline{B})^T$.

$$\overline{B} = \begin{pmatrix} -i & 2-i \\ -2-i & 0 \end{pmatrix} \implies B^* = (\overline{B})^T = \begin{pmatrix} -i & -2-i \\ 2-i & 0 \end{pmatrix}$$

Now we compare B^* with $-B$:

$$-B = -\begin{pmatrix} i & 2+i \\ -2+i & 0 \end{pmatrix} = \begin{pmatrix} -i & -2-i \\ 2-i & 0 \end{pmatrix}$$

Since $B^* = -B$, the matrix B is **skew-Hermitian**. ■

1.1.3 Representation of a Square Matrix as P + iQ

A fundamental result states that any square matrix with complex entries can be uniquely decomposed into a combination of two Hermitian matrices.

Theorem 1.1.7. *Every square matrix M can be uniquely expressed in the form $M = P + iQ$, where P and Q are both Hermitian matrices.*

Proof. The proof is constructive. We define the matrices P and Q as follows:

$$P = \frac{1}{2}(M + M^*) \quad \text{and} \quad Q = \frac{1}{2i}(M - M^*)$$

We must show that P and Q are Hermitian. For P, we take the conjugate transpose:

$$P^* = \left(\frac{1}{2}(M + M^*)\right)^* = \frac{1}{2}(M^* + (M^*)^*) = \frac{1}{2}(M^* + M) = P$$

Thus, P is Hermitian. For Q, we proceed similarly, noting that $(i)^* = \bar{i} = -i$:

$$Q^* = \left(\frac{1}{2i}(M - M^*)\right)^* = \frac{1}{-2i}(M^* - (M^*)^*) = \frac{1}{-2i}(M^* - M) = \frac{1}{2i}(M - M^*) = Q$$

Thus, Q is also Hermitian. To confirm the decomposition, we can substitute P and Q back:

$$P + iQ = \frac{1}{2}(M + M^*) + i\left(\frac{1}{2i}(M - M^*)\right) = \frac{1}{2}(M + M^*) + \frac{1}{2}(M - M^*) = \frac{1}{2}(2M) = M$$

The decomposition is valid. The uniqueness of this representation can also be proven. ■

Example 1.1.8 (Representation of a Square Matrix as P + iQ)**.** Express the matrix $M = \begin{pmatrix} 2+i & 4 \\ 2i & 3-2i \end{pmatrix}$ in the form $P + iQ$, where P and Q are both Hermitian matrices.

Solution. We use the formulas $P = \frac{1}{2}(M + M^*)$ and $Q = \frac{1}{2i}(M - M^*)$. First, we find the conjugate transpose of M:

$$M^* = (\overline{M})^T = \begin{pmatrix} 2-i & -2i \\ 4 & 3+2i \end{pmatrix}$$

Now we compute the sum $M + M^*$ to find P:

$$M + M^* = \begin{pmatrix} (2+i) + (2-i) & 4+4 \\ 2i + (-2i) & (3-2i) + (3+2i) \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 0 & 6 \end{pmatrix}$$

$$P = \frac{1}{2}(M + M^*) = \begin{pmatrix} 2 & 4 \\ 0 & 3 \end{pmatrix}$$

Next, we compute the difference $M - M^*$ to find Q:

$$M - M^* = \begin{pmatrix} (2+i) - (2-i) & 4-4 \\ 2i - (-2i) & (3-2i) - (3+2i) \end{pmatrix} = \begin{pmatrix} 2i & 0 \\ 4i & -4i \end{pmatrix}$$

$$Q = \frac{1}{2i}(M - M^*) = \frac{1}{2i}\begin{pmatrix} 2i & 0 \\ 4i & -4i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & -2 \end{pmatrix}$$

We can quickly verify that both P and Q are real symmetric matrices, and thus are Hermitian. The required decomposition is:

$$M = \begin{pmatrix} 2 & 4 \\ 0 & 3 \end{pmatrix} + i\begin{pmatrix} 1 & 0 \\ 2 & -2 \end{pmatrix}$$

■

1.2 The Adjoint Matrix and Fundamental Laws of Operations

1.2.1 The Adjoint of a Matrix

The adjoint (or adjugate) of a square matrix is a fundamental concept used in the theoretical definition of the inverse of a matrix. It is constructed from the matrix of cofactors.

Definition 1.2.1 (Cofactor and Adjoint)**.** Let A be an $n \times n$ square matrix.

1. The **minor** M_{ij} of the element a_{ij} is the determinant of the $(n-1)\times(n-1)$ matrix obtained by deleting the i-th row and j-th column of A.
2. The **cofactor** C_{ij} of the element a_{ij} is given by $C_{ij} = (-1)^{i+j} M_{ij}$.
3. The **matrix of cofactors** is the matrix $C = [C_{ij}]$.
4. The **adjoint** of A, denoted $\text{adj}(A)$, is the transpose of the matrix of cofactors.

$$\text{adj}(A) = C^T$$

For a simple 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the cofactors are $C_{11} = d, C_{12} = -c, C_{21} = -b, C_{22} = a$. The cofactor matrix is $\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$. The adjoint is therefore $\text{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

1.2.2 The Adjoint-Inverse Identity

The most important property of the adjoint is its relationship with the original matrix and its determinant.

Theorem 1.2.2. *For any $n \times n$ square matrix A, the following identity holds:*

$$A(\textit{adj } A) = (\textit{adj } A)A = \det(A)I$$

where I is the $n\times n$ identity matrix.

Proof. Let's consider the product $A(\text{adj } A)$. The element in the i-th row and j-th column of this product is the dot product of the i-th row of A and the j-th column of $\text{adj}(A)$. The j-th column of $\text{adj}(A)$ is the j-th row of the cofactor matrix C, which consists of the cofactors $(C_{j1}, C_{j2}, \dots, C_{jn})$. So, the (i,j)-th element of the product is:

$$(A(\text{adj } A))_{ij} = a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = \sum_{k=1}^{n} a_{ik}C_{jk}$$

We consider two cases for this sum, based on Laplace's expansion of a determinant.

- **Case 1: $i = j$ (Diagonal Elements).** The sum becomes $\sum_{k=1}^{n} a_{ik}C_{ik}$. This is precisely the definition of the determinant of A expanded along the i-th row. Thus, all diagonal elements of the product are equal to $\det(A)$.
- **Case 2: $i \neq j$ (Off-Diagonal Elements).** The sum $\sum_{k=1}^{n} a_{ik}C_{jk}$ represents the determinant of a matrix formed by replacing the j-th row of A with a copy of the i-th row. A matrix with two identical rows has a determinant of zero. Thus, all off-diagonal elements of the product are zero.

Combining these two cases, the resulting matrix has $\det(A)$ on its main diagonal and zeros everywhere else. This is exactly the matrix $\det(A)I$. The proof for $(\text{adj } A)A$ is analogous, using column expansions. ■

1.2.3 Commutative and Associative Laws in Matrix Operations

Not all laws of real number arithmetic apply directly to matrices. It is essential to know which laws hold.

Matrix Addition Matrix addition is both commutative and associative, provided the matrices have the same dimensions.

- **Commutative Law:** $A + B = B + A$
- **Associative Law:** $(A + B) + C = A + (B + C)$

Matrix Multiplication Matrix multiplication is associative, but it is **not** commutative in general.

- **Associative Law:** $A(BC) = (AB)C$ (provided the products are defined).
- **Non-Commutative Law:** In general, $AB \neq BA$.

  For a direct counterexample, let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

  $$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

  $$BA = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

  Clearly, $AB \neq BA$.

Scalar Multiplication Scalar multiplication is associative and commutative with respect to scalars and matrices. For any scalar c and matrices A, B:

$$c(AB) = (cA)B = A(cB)$$

1.3 Invertibility and the Reversal Law for Inverses

1.3.1 Necessary and Sufficient Condition for Invertibility

The invertibility of a square matrix is one of its most important properties. There are many equivalent conditions for a matrix to be invertible, but the most fundamental one relates to its determinant.

Theorem 1.3.1. *A square matrix A is invertible (or non-singular) if and only if its determinant is non-zero, i.e.,* $\det(A) \neq 0$.

Proof. We prove both directions of the "if and only if" statement.

(⇒) Assume A is invertible. By definition, there exists a matrix A^{-1} such that $AA^{-1} = I$, where I is the identity matrix. Taking the determinant of both sides, we get $\det(AA^{-1}) = \det(I)$. Using the property that the determinant of a product is the product of the determinants, we have $\det(A)\det(A^{-1}) = 1$. Since the product of these two numbers is 1, neither can be zero. Therefore, $\det(A) \neq 0$.

(⇐) Assume $\det(A) \neq 0$. We know from the properties of the adjoint matrix that $A(\text{adj } A) = \det(A)I$. Since $\det(A)$ is a non-zero scalar, we can divide the entire equation by it:

$$A\left(\frac{1}{\det(A)}\text{adj } A\right) = I$$

This equation is of the form $AB = I$, where $B = \frac{1}{\det(A)}\text{adj}(A)$. This shows that A has a right inverse. A similar argument using $(\text{adj } A)A = \det(A)I$ shows it has a left inverse. Therefore, A is invertible, and its inverse is given by the formula $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$. ■

Example 1.3.2 (Condition for Invertibility)**.** Determine if the following matrix A is invertible.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 2 & 1 & 1 \end{pmatrix}$$

Solution. A square matrix is invertible if and only if its determinant is non-zero. We calculate the determinant of A by expanding along the first column:

$$\begin{aligned} \det(A) &= 1 \cdot \det\begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} - 0 \cdot \det\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} + 2 \cdot \det\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \\ &= 1(1-4) - 0 + 2(8-3) \\ &= -3 + 2(5) = -3 + 10 = 7 \end{aligned}$$

Since $\det(A) = 7 \neq 0$, the matrix A is invertible. ■

1.3.2 Generalization of the Reversal Law for Inverses

Similar to the transpose, the inverse of a product of matrices is the product of their inverses in the reverse order.

Theorem 1.3.3 (Reversal Law for Two Matrices)**.** *If A and B are invertible matrices of the same size, then their product AB is also invertible, and its inverse is given by:*

$$(AB)^{-1} = B^{-1}A^{-1}$$

Proof. To prove that $B^{-1}A^{-1}$ is the inverse of AB, we must show that their product is the identity matrix. We multiply (AB) by $(B^{-1}A^{-1})$:

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \quad \text{(by associativity)} \\ &= A(I)A^{-1} = AA^{-1} = I \end{aligned}$$

Since their product is the identity matrix, $B^{-1}A^{-1}$ is indeed the inverse of AB. ■

This law can be generalized to the product of any finite number of invertible matrices.

Theorem 1.3.4 (Generalization of Reversal Law)**.** *If $A_1, A_2, \ldots, A_n$ are invertible matrices of the same size, then their product is also invertible, and:*

$$(A_1A_2\cdots A_n)^{-1} = A_n^{-1}\cdots A_2^{-1}A_1^{-1}$$

Proof. We prove this by mathematical induction on n.

Base Case: For $n = 2$, the theorem is $(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}$, which we have just proven.

Inductive Hypothesis: Assume the formula holds for a product of k invertible matrices.

$$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$$

Inductive Step: We prove the formula for $k+1$ matrices. We group the product as $(A_1\cdots A_k)A_{k+1}$ and apply the base case for two matrices:

$$((A_1\cdots A_k)A_{k+1})^{-1} = A_{k+1}^{-1}(A_1\cdots A_k)^{-1}$$

Now, using our inductive hypothesis on the term $(A_1\cdots A_k)^{-1}$, we get:

$$A_{k+1}^{-1}(A_k^{-1}\cdots A_2^{-1}A_1^{-1}) = A_{k+1}^{-1}A_k^{-1}\cdots A_2^{-1}A_1^{-1}$$

Thus, the formula holds for $k+1$ matrices. By the principle of mathematical induction, the theorem is true for all integers $n \geq 2$. ∎

Example 1.3.5 (Reversal Law for Inverses)**.** Verify the reversal law $(AB)^{-1} = B^{-1}A^{-1}$ for the matrices:

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

Solution. Left-Hand Side (LHS). First, we compute the product AB:

$$AB = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 5 & 4 \end{pmatrix}$$

The determinant is $\det(AB) = (4)(4) - (3)(5) = 16 - 15 = 1$. The inverse is:

$$(AB)^{-1} = \frac{1}{1}\begin{pmatrix} 4 & -3 \\ -5 & 4 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ -5 & 4 \end{pmatrix}$$

Right-Hand Side (RHS). First, we find the individual inverses of A and B.

$$\det(A) = 3 - 2 = 1 \implies A^{-1} = \frac{1}{1}\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$$

$$\det(B) = 2 - 1 = 1 \implies B^{-1} = \frac{1}{1}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Now, we compute the product $B^{-1}A^{-1}$:

$$B^{-1}A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3+1 & -2-1 \\ -3-2 & 2+2 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ -5 & 4 \end{pmatrix}$$

Conclusion. Since the LHS and RHS are identical, the reversal law is verified. ∎

Chapter 2

Matrix Operations and Eigenvalues

2.1 Matrix Operation of Transposing and Inverting are Commutative

For an invertible matrix, the order in which you perform the transpose and inverse operations does not matter.

Theorem 2.1.1. *If A is an invertible square matrix, then its transpose A^T is also invertible, and the transpose of the inverse is the inverse of the transpose.*

$$(A^{-1})^T = (A^T)^{-1}$$

Proof. To prove this, we must show that $(A^{-1})^T$ is the inverse of A^T. We can do this by showing that their product is the identity matrix. Using the reversal law for transposes, $(XY)^T = Y^T X^T$, we have:

$$A^T (A^{-1})^T = (A^{-1}A)^T$$

Since $A^{-1}A = I$, the expression becomes:

$$(I)^T = I$$

Because $A^T(A^{-1})^T = I$, it follows by the definition of an inverse that $(A^{-1})^T$ is the inverse of A^T. ■

Example 2.1.2 (Commutativity of Transpose and Inverse)**.** Verify that $(A^{-1})^T = (A^T)^{-1}$ for the matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$.

Solution. Left-Hand Side (LHS). First, we find the inverse of A. Since $\det(A) = 1$, the inverse is:

$$A^{-1} = \frac{1}{1}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

Now, we take the transpose of A^{-1}:

$$(A^{-1})^T = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

Right-Hand Side (RHS). First, we find the transpose of A:

$$A^T = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

Now, we find the inverse of A^T. Since $\det(A^T) = 1$, the inverse is:

$$(A^T)^{-1} = \frac{1}{1}\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

Conclusion. The left-hand side and right-hand side are identical, verifying the theorem. ■

2.1.1 Trace of a Matrix

The trace is a simple but useful operator that maps a square matrix to a scalar value.

Definition 2.1.3 (Trace)**.** The **trace** of an $n \times n$ square matrix A, denoted $\mathrm{tr}(A)$, is the sum of the elements on its main diagonal.

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

Properties of the Trace: The trace is a linear operator, meaning for any square matrices A, B and scalar c:

- $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$
- $\mathrm{tr}(cA) = c \cdot \mathrm{tr}(A)$

2.1.2 The Cyclic Property of the Trace

While matrix multiplication is not commutative ($AB \neq BA$), the trace of the product is surprisingly immune to the order of multiplication.

Theorem 2.1.4. *If A is an $m \times n$ matrix and B is an $n \times m$ matrix, then the trace of their product is commutative.*

$$tr(AB) = tr(BA)$$

Proof. By the definition of matrix multiplication, the (i,i)-th element of the product AB is $(AB)_{ii} = \sum_{j=1}^{n} a_{ij}b_{ji}$. The trace of AB is the sum of these diagonal elements:

$$\mathrm{tr}(AB) = \sum_{i=1}^{m}(AB)_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji}$$

Similarly, the (j,j)-th element of the product BA is $(BA)_{jj} = \sum_{i=1}^{m} b_{ji}a_{ij}$. The trace of BA is:

$$\mathrm{tr}(BA) = \sum_{j=1}^{n}(BA)_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij}$$

Since the summation of scalars is commutative, we can reorder the terms and the summations. The two expressions are equal. ■

This property can be generalized for any number of matrices in what is known as the **cyclic property of the trace**.

Theorem 2.1.5 (Generalization)**.** *For any three matrices A, B, C whose product ABC is a square matrix, the trace is invariant under cyclic permutations of the matrices.*

$$tr(ABC) = tr(BCA) = tr(CAB)$$

Proof. We can prove this by cleverly grouping the matrices and applying the base case $\mathrm{tr}(XY) = \mathrm{tr}(YX)$. Let $X = AB$ and $Y = C$. Then:

$$\mathrm{tr}(ABC) = \mathrm{tr}((AB)C) = \mathrm{tr}(XY) = \mathrm{tr}(YX) = \mathrm{tr}(C(AB)) = \mathrm{tr}(CAB)$$

Now let $X = A$ and $Y = BC$. Then:

$$\mathrm{tr}(ABC) = \mathrm{tr}(A(BC)) = \mathrm{tr}(XY) = \mathrm{tr}(YX) = \mathrm{tr}((BC)A) = \mathrm{tr}(BCA)$$

Thus, all three expressions are equal. ■

Example 2.1.6 (Trace of a Matrix)**.** Find the trace of the matrix $A = \begin{pmatrix} 5 & -1 & 0 \\ 2 & 3 & 7 \\ 1 & 6 & -4 \end{pmatrix}$.

Solution. The trace of a square matrix is the sum of its main diagonal elements.

$$\mathrm{tr}(A) = 5 + 3 + (-4) = 4$$

The trace of the matrix is 4. ■

Example 2.1.7 (Cyclic Property of the Trace: Two Matrices)**.** Verify that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for the non-square matrices:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{pmatrix}$$

Solution. First, we compute the 2×2 product AB and its trace.

$$AB = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+4+0 & 1+4+0 \\ 3+0+0 & 3+0+1 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 3 & 4 \end{pmatrix}$$

$$\mathrm{tr}(AB) = 5 + 4 = 9$$

Next, we compute the 3×3 product BA and its trace.

$$BA = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+3 & 2+0 & 0+1 \\ 2+6 & 4+0 & 0+2 \\ 0+3 & 0+0 & 0+1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ 8 & 4 & 2 \\ 3 & 0 & 1 \end{pmatrix}$$

$$\mathrm{tr}(BA) = 4 + 4 + 1 = 9$$

The traces are equal, verifying the property. ■

Example 2.1.8 (Cyclic Property of the Trace: Three Matrices)**.** Verify that $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$ for the matrices:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Solution. First, we calculate $\mathrm{tr}(ABC)$.

$$AB = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$$

$$(AB)C = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\mathrm{tr}(ABC) = 2 + 2 = 4$$

Next, we calculate $\mathrm{tr}(BCA)$.

$$BC = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$$

$$(BC)A = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 0 & 0 \end{pmatrix}$$

$$\mathrm{tr}(BCA) = 4 + 0 = 4$$

The traces are equal, verifying the cyclic property for this set of matrices. ∎

2.2 Partitioning of Matrices

Partitioning (or blocking) a matrix is the process of dividing it into smaller rectangular submatrices, called **blocks** or **submatrices**.

Definition 2.2.1 (Partitioned Matrix)**.** An $m \times n$ matrix A can be partitioned by drawing horizontal and vertical lines between its rows and columns. For example, a matrix can be represented as a 2×2 partitioned matrix:

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right)$$

Here, A_{11}, A_{12}, A_{21}, and A_{22} are the blocks (submatrices) of A.

Operations on Partitioned Matrices

Addition: If two matrices A and B have the same dimensions and are partitioned in exactly the same way, their sum is found by simply adding the corresponding blocks:

$$A + B = \left(\begin{array}{c|c} A_{11} + B_{11} & A_{12} + B_{12} \\ \hline A_{21} + B_{21} & A_{22} + B_{22} \end{array} \right)$$

Block Multiplication: If two matrices A and B are partitioned such that the number of columns in each block of A matches the number of rows in the corresponding block of B (i.e., the column partitions of A conform to the row partitions of B), their product can be computed by treating the blocks as if they were individual elements. If $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, their product is:

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Example 2.2.2 (Block Multiplication)**.** Calculate the product AB using block multiplication, where:

$$A = \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 1 \\ \hline -1 & 0 & 3 \end{array} \right) \quad \text{and} \quad B = \left(\begin{array}{c|c} 1 & 0 \\ 0 & 2 \\ \hline 3 & 1 \end{array} \right)$$

Solution. We define the partitions as follows:

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A_{21} = \begin{pmatrix} -1 & 0 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 3 \end{pmatrix}$$

$$B_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad B_{12} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad B_{21} = \begin{pmatrix} 3 & 1 \end{pmatrix}, \quad B_{22} = \begin{pmatrix} 1 \end{pmatrix}$$

The column partition of A (2 columns, 1 column) matches the row partition of B (2 rows, 1 row), so block multiplication is possible. We compute the blocks of the product $C = AB$:

Top-Left Block (C_{11}):

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 3 & 3 \end{pmatrix}$$

Top-Right Block (C_{12}):

$$C_{12} = A_{11}B_{12} + A_{12}B_{22} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

Bottom-Left Block (C_{21}):

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 3 \end{pmatrix} \begin{pmatrix} 3 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} 9 & 3 \end{pmatrix} = \begin{pmatrix} 8 & 3 \end{pmatrix}$$

Bottom-Right Block (C_{22}):

$$C_{22} = A_{21}B_{12} + A_{22}B_{22} = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \end{pmatrix} \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix} + \begin{pmatrix} 3 \end{pmatrix} = \begin{pmatrix} 3 \end{pmatrix}$$

Combining these blocks gives the final product:

$$AB = \left(\begin{array}{cc|c} 1 & 4 & 4 \\ 3 & 3 & 3 \\ \hline 8 & 3 & 3 \end{array}\right) = \begin{pmatrix} 1 & 4 & 4 \\ 3 & 3 & 3 \\ 8 & 3 & 3 \end{pmatrix}$$

This result is identical to the one obtained by standard matrix multiplication. ■

2.3 Eigenvalues, Eigenvectors, and Associated Polynomials

Definition 2.3.1 (Eigenvalue and Eigenvector)**.** For an $n \times n$ matrix A, a non-zero vector $\mathbf{v}$ is an **eigenvector** if there exists a scalar λ such that:

$$A\mathbf{v} = \lambda\mathbf{v}$$

The scalar λ is the **eigenvalue** corresponding to $\mathbf{v}$. It represents the factor by which the eigenvector $\mathbf{v}$ is stretched or shrunk.

To find the eigenvalues, we rearrange the defining equation:

$$A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0} \implies (A - \lambda I)\mathbf{v} = \mathbf{0}$$

This is a homogeneous system of linear equations. Since we require the eigenvector $\mathbf{v}$ to be non-zero, we are looking for non-trivial solutions. A non-trivial solution exists if and only if the matrix $(A - \lambda I)$ is singular (i.e., not invertible). This gives us a direct method for finding eigenvalues.

2.3.1 The Characteristic Equation

A matrix is singular if and only if its determinant is zero. This principle leads us to the characteristic equation.

Definition 2.3.2 (Characteristic Polynomial and Equation)**.** The **characteristic polynomial** of an $n \times n$ matrix A is the polynomial $p(\lambda) = \det(A - \lambda I)$. The **characteristic equation** is $p(\lambda) = 0$.

From the reasoning above, the eigenvalues of a matrix A are precisely the roots of its characteristic equation.

Example 2.3.3 (Finding Eigenvalues and Eigenvectors)**.** Find the eigenvalues and corresponding eigenvectors of the matrix $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

Solution. First, we find the eigenvalues by solving the characteristic equation $\det(A - \lambda I) = 0$.

$$\det \begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix} = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 12 - 2 = \lambda^2 - 7\lambda + 10 = 0$$

Factoring the polynomial gives $(\lambda - 5)(\lambda - 2) = 0$. The eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 2$.

Now, we find the eigenvector for each eigenvalue.

For $\lambda_1 = 5$: We solve the system $(A - 5I)\mathbf{v} = \mathbf{0}$.

$$\begin{pmatrix} 4-5 & 1 \\ 2 & 3-5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives the equation $-x + y = 0$, or $x = y$. A simple non-zero solution is $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

For $\lambda_2 = 2$: We solve the system $(A - 2I)\mathbf{v} = \mathbf{0}$.

$$\begin{pmatrix} 4-2 & 1 \\ 2 & 3-2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives the equation $2x + y = 0$, or $y = -2x$. A simple non-zero solution is $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. ■

Example 2.3.4 (Evaluating a Matrix Polynomial)**.** Given the matrix $C = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and the polynomial $p(x) = x^2 - 2x + 3$, find the matrix $p(C)$.

Solution. We need to compute $p(C) = C^2 - 2C + 3I$. First, we calculate the matrix C^2:

$$C^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$$

Now we substitute this into the polynomial expression:

$$\begin{aligned} p(C) &= C^2 - 2C + 3I \\ &= \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} + 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1-2+3 & 4-4+0 \\ 0-0+0 & 1-2+3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

Thus, $p(C) = 2I$. ■

2.3.2 The Cayley-Hamilton Theorem

This remarkable theorem states that a matrix satisfies its own characteristic equation. Before stating it, we must define a matrix polynomial.

Definition 2.3.5 (Matrix Polynomial)**.** If $p(x) = c_k x^k + \cdots + c_1 x + c_0$ is a scalar polynomial, then for a square matrix A, the matrix polynomial $p(A)$ is:

$$p(A) = c_k A^k + \cdots + c_1 A + c_0 I$$

Note that the constant term is multiplied by the identity matrix I.

Theorem 2.3.6 (Cayley-Hamilton Theorem)**.** *If $p(\lambda)$ is the characteristic polynomial of a square matrix A, then $p(A) = 0$, where 0 is the zero matrix.*

This theorem is incredibly useful for finding the inverse of a matrix. If $p(\lambda) = c_n\lambda^n + \cdots + c_1\lambda + c_0$, then $p(A) = c_n A^n + \cdots + c_1 A + c_0 I = 0$. If A is invertible, then its determinant, c_0, is non-zero. We can then multiply by A^{-1} and solve for it in terms of powers of A.

Example 2.3.7 (Application of the Cayley-Hamilton Theorem)**.** Use the Cayley-Hamilton theorem to find the inverse of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$.

Solution. First, we find the characteristic equation of A.

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 2 \\ 3 & 5-\lambda \end{pmatrix} = (1-\lambda)(5-\lambda) - 6 = 5 - 6\lambda + \lambda^2 - 6 = \lambda^2 - 6\lambda - 1 = 0$$

By the Cayley-Hamilton theorem, the matrix A satisfies this equation:

$$A^2 - 6A - I = 0$$

Since $\det(A) = -1 \neq 0$, the inverse exists. We can multiply the equation by A^{-1}:

$$A^{-1}(A^2 - 6A - I) = 0 \implies A - 6I - A^{-1} = 0$$

Solving for A^{-1} gives:

$$A^{-1} = A - 6I = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} - 6\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-6 & 2-0 \\ 3-0 & 5-6 \end{pmatrix} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$

■

2.3.3 The Minimal Polynomial of a Matrix

A matrix satisfies its characteristic polynomial, but is that the polynomial of the *smallest* degree that it satisfies? The answer is not always yes, which leads to the concept of the minimal polynomial.

Definition 2.3.8 (Minimal Polynomial)**.** The **minimal polynomial** of a square matrix A is the unique monic polynomial (leading coefficient is 1) $m(\lambda)$ of least degree such that $m(A) = 0$.

Theorem 2.3.9. *The minimal polynomial $m(\lambda)$ of a matrix A always divides the characteristic polynomial $p(\lambda)$. Furthermore, every eigenvalue of A (i.e., every root of $p(\lambda)$) is also a root of $m(\lambda)$.*

This means that the minimal polynomial and the characteristic polynomial share the same roots, but the characteristic polynomial may have these roots with higher multiplicities. For example, if $p(\lambda) = (\lambda-2)^3(\lambda-5)$, the minimal polynomial could be $(\lambda-2)(\lambda-5)$, $(\lambda-2)^2(\lambda-5)$, or the characteristic polynomial itself.

Example 2.3.10 (Finding the Minimal Polynomial)**.** Find the minimal polynomial of the matrix $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$.

Solution. First, we find the characteristic polynomial $p(\lambda)$.

$$p(\lambda) = \det(D - \lambda I) = \det \begin{pmatrix} 3-\lambda & 0 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{pmatrix} = (3-\lambda)^2(5-\lambda) = -(\lambda-3)^2(\lambda-5)$$

The characteristic polynomial is $p(\lambda) = (\lambda-3)^2(\lambda-5)$. The minimal polynomial $m(\lambda)$ must have the same roots (3 and 5) and must divide $p(\lambda)$. The candidates for the minimal polynomial are:

1. $m_1(\lambda) = (\lambda-3)(\lambda-5) = \lambda^2 - 8\lambda + 15$
2. $m_2(\lambda) = (\lambda-3)^2(\lambda-5) = p(\lambda)$

We test the polynomial of the smallest degree first. We check if $m_1(D) = 0$.

$$m_1(D) = D^2 - 8D + 15I$$

$$D^2 = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 25 \end{pmatrix}$$

$$\begin{aligned} m_1(D) &= \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 25 \end{pmatrix} - 8 \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} + 15 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 9-24+15 & 0 & 0 \\ 0 & 9-24+15 & 0 \\ 0 & 0 & 25-40+15 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Since $m_1(D) = 0$ and $m_1(\lambda)$ is a monic polynomial of the least possible degree that annihilates D, the minimal polynomial of D is $m(\lambda) = (\lambda-3)(\lambda-5) = \lambda^2 - 8\lambda + 15$. ■

Example 2.3.11. For the matrix A:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Find (a) its eigenvalues and corresponding eigenvectors, (b) its minimal polynomial, and (c) its inverse using the Cayley-Hamilton theorem.

Solution. (a) Eigenvalues and Eigenvectors Since A is an upper triangular matrix, its eigenvalues are the entries on its main diagonal. The eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 2$ (with an algebraic multiplicity of 2).

Now, we find the eigenvectors for each distinct eigenvalue by solving the system $(A - \lambda I)\mathbf{v} = \mathbf{0}$.

For $\lambda_1 = 3$: We solve the system $(A - 3I)\mathbf{v} = \mathbf{0}$.

$$\begin{pmatrix} 2-3 & 1 & 0 \\ 0 & 2-3 & 0 \\ 0 & 0 & 3-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This gives the equations $-x + y = 0$ and $-y = 0$. From the second equation, $y = 0$. Substituting into the first equation gives $x = 0$. The variable z is free. A simple non-zero choice is $z = 1$. The eigenvector corresponding to $\lambda_1 = 3$ is $\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

For $\lambda_2 = 2$: We solve the system $(A - 2I)\mathbf{v} = \mathbf{0}$.

$$\begin{pmatrix} 2-2 & 1 & 0 \\ 0 & 2-2 & 0 \\ 0 & 0 & 3-2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This gives the equations $y = 0$ and $z = 0$. The variable x is free. A simple non-zero choice is $x = 1$. The eigenvector corresponding to $\lambda_2 = 2$ is $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Note: Even though $\lambda = 2$ is an eigenvalue with algebraic multiplicity 2, we can only find one linearly independent eigenvector for it. This means the matrix is not diagonalizable.

(b) Minimal Polynomial The characteristic polynomial $p(\lambda)$ is determined by the eigenvalues:

$$p(\lambda) = (\lambda - 3)(\lambda - 2)^2 = (\lambda - 3)(\lambda^2 - 4\lambda + 4) = \lambda^3 - 7\lambda^2 + 16\lambda - 12$$

The minimal polynomial $m(\lambda)$ must have the same roots (3 and 2) and must divide $p(\lambda)$. The candidates for the minimal polynomial are:

1. $m_1(\lambda) = (\lambda - 3)(\lambda - 2) = \lambda^2 - 5\lambda + 6$
2. $m_2(\lambda) = (\lambda - 3)(\lambda - 2)^2 = p(\lambda)$

We test the polynomial of the smallest degree, $m_1(\lambda)$. We check if $m_1(A) = A^2 - 5A + 6I$ is the zero matrix. First, we compute A^2:

$$A^2 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

Now we evaluate $m_1(A)$:

$$A^2 - 5A + 6I = \begin{pmatrix} 4 & 4 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix} - 5 \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + 6 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4 - 10 + 6 & 4 - 5 + 0 & 0 \\ 0 & 4 - 10 + 6 & 0 \\ 0 & 0 & 9 - 15 + 6 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

Since $m_1(A) \neq 0$, the minimal polynomial must be the next candidate, which is the characteristic polynomial itself. The minimal polynomial is $m(\lambda) = (\lambda - 3)(\lambda - 2)^2 = \lambda^3 - 7\lambda^2 + 16\lambda - 12$.

(c) Inverse using the Cayley-Hamilton Theorem By the Cayley-Hamilton theorem, $p(A) = 0$, so:

$$A^3 - 7A^2 + 16A - 12I = 0$$

The determinant is the constant term of the characteristic polynomial evaluated at $\lambda = 0$, so $\det(A) = -(-12) = 12$. Since the determinant is non-zero, the inverse exists. Multiply the equation by A^{-1}:

$$A^2 - 7A + 16I - 12A^{-1} = 0$$

Solving for A^{-1} gives:

$$12A^{-1} = A^2 - 7A + 16I$$

Using our previously calculated A^2:

$$12A^{-1} = \begin{pmatrix} 4 & 4 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix} - 7 \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + 16 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$12A^{-1} = \begin{pmatrix} 4 - 14 + 16 & 4 - 7 + 0 & 0 \\ 0 & 4 - 14 + 16 & 0 \\ 0 & 0 & 9 - 21 + 16 \end{pmatrix} = \begin{pmatrix} 6 & -3 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Therefore, the inverse is:

$$A^{-1} = \frac{1}{12} \begin{pmatrix} 6 & -3 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/4 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

■

→ A matrix is **non-derogatory** if its minimal polynomial is identical to its characteristic polynomial.

→ A matrix is **derogatory** if its minimal polynomial is of a strictly smaller degree than its characteristic polynomial. This can only happen if the matrix has at least one repeated eigenvalue.

Example 2.3.12 (A Non-Derogatory Matrix)**.** Show that the following matrix A is non-derogatory.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Solution. Since A is an upper triangular matrix, its eigenvalues are its diagonal entries: $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$. The characteristic polynomial $p(\lambda)$ is the product of the factors corresponding to these roots:

$$p(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3) = \lambda^3 - 6\lambda^2 + 11\lambda - 6$$

The degree of the characteristic polynomial is 3.

The minimal polynomial $m(\lambda)$ must have the same distinct roots as the characteristic polynomial. In this case, the roots are 1, 2, and 3. The monic polynomial of least degree containing these roots is:

$$m(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3)$$

The degree of this minimal polynomial is 3.

The minimal polynomial is identical to the characteristic polynomial (up to a sign, though both are typically taken as monic). Since the degree of the minimal polynomial (3) is equal to the degree of the characteristic polynomial (3), the matrix A is **non-derogatory**. This is always true for matrices with distinct eigenvalues. ■

Example 2.3.13 (A Derogatory Matrix)**.** Show that the following matrix B is derogatory.

$$B = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution. The matrix B is diagonal, so its eigenvalues are the diagonal entries: $\lambda_1 = 4$ (with algebraic multiplicity 2) and $\lambda_2 = 1$. The characteristic polynomial is:

$$p(\lambda) = (\lambda - 4)^2(\lambda - 1) = (\lambda^2 - 8\lambda + 16)(\lambda - 1) = \lambda^3 - 9\lambda^2 + 24\lambda - 16$$

The degree of the characteristic polynomial is 3.

The minimal polynomial $m(\lambda)$ must have the same distinct roots (4 and 1) and must divide $p(\lambda)$. The possible candidates for the minimal polynomial are:

1. $m_1(\lambda) = (\lambda - 4)(\lambda - 1) = \lambda^2 - 5\lambda + 4$
2. $m_2(\lambda) = (\lambda - 4)^2(\lambda - 1) = p(\lambda)$

We must test the candidate with the smallest degree first. Let's check if $m_1(B) = B^2 - 5B + 4I$ is the zero matrix. First, we compute B^2:

$$B^2 = \begin{pmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, we evaluate the polynomial:

$$\begin{aligned} m_1(B) &= B^2 - 5B + 4I \\ &= \begin{pmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 5\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 4\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 16 - 20 + 4 & 0 & 0 \\ 0 & 16 - 20 + 4 & 0 \\ 0 & 0 & 1 - 5 + 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Since $m_1(B) = 0$, the minimal polynomial of B is $m(\lambda) = \lambda^2 - 5\lambda + 4$.

The degree of the characteristic polynomial is 3, while the degree of the minimal polynomial is 2. Since the degree of the minimal polynomial is strictly less than the degree of the characteristic polynomial, the matrix B is **derogatory**. ■