Chapter 9

Matrices

Matrices are a rectangular arrangement of numbers, expressions, symbols which are arranged as rows and columns. The numbers represented in the matrix are called as entries. Matrices find many applications in solving practical real life problems making it an indispensable concept. Matrices have wide applications in engineering analysis and design, physics, economics, and statistics. Matrices also have important applications in computer graphics for image transformations. More recently, matrices have found wide use in the field of Machine Learning (ML). Modern computers are equipped with specially designed hardware called a Graphics Processing Unit or a GPU that is used for parallel processing of matrix operations for much quicker results than ordinary sequential processing.

9.1 Definition of a Matrix

A matrix of order $m \times n$, or m by n matrix, is a rectangular array of numbers having m rows and n columns. It is represented as:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m21} & \cdots & a_{mn} \end{bmatrix}$$

9.2 Definitions & Operations Involving Matrices

9.2.1 Equality

Two matrices A and B are equal, i.e., $A = B$, if and only if they are of the same size and their corresponding entries are equal, i.e., $a_{ij} = b_{ij}$.

9.2.2 Addition (or subtraction):

If two matrices A and B have the same size, then $A + B$ has the entries $[a_{ij} \pm b_{ij}]$. Example,

$$\begin{bmatrix} 3 & 2 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 6 & 6 \end{bmatrix}$$

9.2.3 Scalar Multiplication

$cA = [c\,a_{ij}]$ where c is a number. Example,

$$2 \times \begin{bmatrix} 3 & 2 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ 8 & 4 \end{bmatrix}$$

9.2.4 Matrix Multiplication

$A\,B = C$, the entries of C are given by:

$$\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}}$$

If A is matrix of size $m \times n$, B is a matrix of size $n \times p$, then the resulting matrix C from their multiplication is of size $m \times p$. Example,

$$\begin{bmatrix} 3 & 2 \\ 4 & 2 \end{bmatrix} \times \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 13 & 11 \\ 16 & 12 \end{bmatrix}$$

Matrix addition is commutative and associative. Matrix Multiplication is not commutative.

9.2.5 Transpose of a Matrix

The transpose of matrix a_{ij} is a matrix with its elements as a_{ji}. The rows of A become the columns of A^T, i.e., the entries of $A^T = [a_{ji}]$. Example,

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{bmatrix} \quad A^T = \begin{bmatrix} 3 & 4 & 3 \\ 2 & 5 & 2 \\ 1 & 6 & 1 \end{bmatrix}$$

9.2.6 Principal Diagonal

If A is a square matrix, then the diagonal which contains all elements a_{jk} for which $j = k$ is called the *principal* or *main diagonal.* Example: *Principal Diagonal of A is* $\begin{bmatrix} 3 & 5 & 1 \end{bmatrix}$.

9.2.7 Trace of a Matrix

The sum of elements of the principal diagonal of a matrix is called the *trace* of A.

9.3 Types of Matrices

9.3.1 Diagonal Matrix

A *Diagonal* matrix is a square matrix that has non-zero entries on its diagonal while all other entries above and below the the diagonal are 0. Example,

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

9.3.2 Zero or Null Matrix

A matrix whose elements are all equal to zero is called the null or zero matrix and is often denoted by O or simply 0. Example,

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

9.3.3 Unit or Identity Matrix

All entries in the diagonal matrix are 1 and all other elements are 0. This implies $AI = IA$, where I is the *Identity Matrix*. Example,

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

9.3.4 Symmetric Matrix & Skew Symmetric Matrix

Symmetric matrices are square matrices whose transpose equals the matrix itself, i.e., $A^T = A$. Skew-symmetric matrices are square matrices whose transpose equals the negative of the matrix, i.e., $A^T = -A$. Example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 3 \end{bmatrix} \text{(Symmetric)} \qquad \begin{bmatrix} 1 & 2 & 3 \\ -2 & 4 & -5 \\ -3 & 5 & 3 \end{bmatrix} \text{(Skew Symmetric)}$$

9.3.5 Orthogonal Matrix

A square matrix A is called an *orthogonal matrix* if its transpose is the same as its inverse, i.e., $A^T = A^{-1}$ or $A^TA = I$. Example,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad A \cdot A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

9.3.6 Complex Conjugate of a Matrix

i is defined as a number whose square gives -1, something no ordinary real number can do. A complex number is just a combination of an ordinary real number and a multiple of this new unit ***i***. We write it as $a + ib$.

A complex conjugate is formed by changing the sign between two terms in a complex number. If all elements a_{jk} of a matrix A are replaced by their complex conjugates $\bar{a}_{jk}$, the matrix obtained is called the complex conjugate of A and is denoted by $\bar{A}$. Example,

$$A = \begin{bmatrix} 1 + 5i & 3 - 2i \\ 2 - 6i & 4 + 4i \end{bmatrix} \quad \bar{A} = \begin{bmatrix} 1 - 5i & 3 + 2i \\ 2 + 6i & 4 - 4i \end{bmatrix}$$

9.3.7 Hermitian & Skew-Hermitian Matrices

A square matrix A, which is the same as the complex conjugate of its transpose, i.e. if $A = \bar{A}^T$, is called *Hermitian* matrix. If $A = -\bar{A}^T$, then A is called *skew-Hermitian* matrix. If A is real, these reduce to symmetric and skew-symmetric matrices respectively. Example,

$$A = \begin{bmatrix} 3 & 1-i \\ 1+i & -2 \end{bmatrix} \quad \bar{A} = \begin{bmatrix} 3 & 1+i \\ 1-i & -2 \end{bmatrix} \quad \bar{A}^T = \begin{bmatrix} 3 & 1-i \\ 1+i & -2 \end{bmatrix} = A \text{ (Hermitian)}$$

$$A = \begin{bmatrix} 3i & 1+i \\ -1+i & -i \end{bmatrix} \quad \bar{A} = \begin{bmatrix} -3i & 1-i \\ -1-i & i \end{bmatrix} \quad \bar{A}^T = \begin{bmatrix} -3i & -1-i \\ 1-i & i \end{bmatrix} = -A \text{ (Skew Hermitian)}$$

9.3.8 Unitary Matrix

A complex square matrix A is called a *unitary matrix* if its complex conjugate transpose is the same as its inverse, i.e., $\bar{A}^T = A^{-1}$ or $\bar{A}^T A = I$. Example,

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}i & -\frac{1}{\sqrt{2}}i \end{bmatrix} \quad \bar{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}i & \frac{1}{\sqrt{2}}i \end{bmatrix} \quad \bar{A}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}i \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}i \end{bmatrix} \quad A \cdot \bar{A}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A \text{ (Unitary)}$$

The real analogue of a unitary matrix is an orthogonal matrix, i.e., if all the entries of a unitary matrix are real (i.e., their complex parts are all zero), then the matrix is orthogonal.

9.4 Linear System of Equations

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{aligned}$$

The matrix form is: $Ax = b$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_1 \\ \cdot \\ x_m \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_1 \\ \cdot \\ b_m \end{bmatrix}$$

Augmented matrix is given by:

$$\tilde{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

9.4.1 Gaussian Elimination

Consider a system of 3 equations:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3
\end{aligned}
$$

Eliminating x_1 using the 2nd and 3rd equations:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a'_{22}x_2 + a'_{23}x_3 &= b'_2 \\
a'_{32}x_2 + a'_{33}x_3 &= b'_3
\end{aligned}
$$

Eliminating x_2 using the 2nd and 3rd equations:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a'_{22}x_2 + a'_{23}x_3 &= b'_2 \\
a''_{33}x_3 &= b''_3
\end{aligned}
$$

We can then solve for x_3, then x_2 and then x_1 from the 3rd, 2nd and 1st equations in that order.

$$
\begin{aligned}
x_3 &= b''_3 / a''_{33} \\
x_2 &= (b'_2 - a'_{23}x_3)/a'_{22} \\
x_1 &= (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11}
\end{aligned}
$$

At the end of the Gauss elimination the form of the coefficient matrix and the augmented matrix is called the **row echelon form**. For the above system of 3 equations, the augmented matrix is:

$$
\tilde{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ 0 & a'_{22} & a'_{23} & b'_2 \\ 0 & 0 & a''_{33} & b''_3 \end{bmatrix}
$$

9.4.2 Jacobi's Iterative Method

Consider the linear system of equations $AX = B$ where,

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n21} & \cdots & a_{nn} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
$$

Then the solution can be obtained iteratively from:

$$\boxed{x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j\neq i} a_{ij}x_j^{(k)}\right)} \quad i = 1, 2, \cdots n, \quad x^{(k)} \ \& \ x^{(k+1)} \text{ are } k^{th} \ \& \ (k+1)^{th} \text{ iteration of } x$$

9.4.3 Gauss - Seidel Method

The Guass-Seidel method is a modification of the Jacobi method that results in higher degree of accuracy within fewer iterations. In Jacobi method the value of the variables is not modified until next iteration. In Gauss-Seidel method the value of the variables are modified as soon as new value is evaluated, i.e., in iteration $(k+1)$, use previously computed value $x_i^{(k+1)}$ if available, otherwise use $x_i^{(k)}$.

9.5 Rank of a Matrix, Linear Independence

9.5.1 Rank

Rank of a matrix A, denoted as **rank (A)**, is the maximum number of linearly independent row vectors of A. It is the number of non-zero rows in its row echelon form.

9.5.2 Existence & Uniqueness of Solutions in Linear Systems

A **consistent system of equations** has at least one solution. A linear system of n equations with n unknowns has an unique solution. This holds true when the *rank of coefficient matrix* A, r, is the same as *rank of augmented matrix* $\tilde{A}$. An **inconsistent system has no solution.** If $r < n$, then the number of solutions is ∞.

9.5.3 Null Space and Nullity

The null space of any matrix A consists of all the vectors B such that $AB = 0$ and B is not zero. It can also be thought as the solution obtained from $AB = 0$ where A is a known matrix of size $m \times n$ and B is a matrix to be found of size $n \times k$. The size of the null space of the matrix provides us with the number of linear relations among attributes. $AB = 0$ implies every row of A when multiplied by B goes to zero. This establishes the linear relationships between the variables. Every null space vector corresponds to one linear relationship. **Nullity** is number of vectors in the null space of matrix A.

9.5.4 Rank Nullity Theorem

Rank of A + Nullity of A = Total number of columns of A

Example,

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix} \implies \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 3 & 6 & 1 \end{bmatrix}$$

The rank of the matrix A which is the number of non-zero rows in its echelon form is 2.

With $AB = 0$,

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0 \implies b_1 + 2b_2 = 0, b_3 = 0 \implies B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \implies b_1 \begin{bmatrix} 1 \\ -\frac{1}{2} \\ 0 \end{bmatrix}$$

Thus nullity, i.e., the dimension of the null space is 1. Thus, the sum of the rank and the nullity of A is $2 + 1 = 3$ which is equal to the number of columns of A.

9.6 Determinant

A **determinant** of order n is a scalar of an $n \times n$ (square) matrix $A[ij]$ is given by:

$$D = det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}$$

$$D = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \quad (j = 1, 2, \cdots, n) \text{ where } M_{ij} \text{ is a determinant of order } n - 1$$

The determinant M_{ij} is obtained by removing the row and column in A corresponding to the element a_{ij}. M_{ij} is called the **minor** of a_{ij}. C_{ij}, called the **cofactor** of a_{ij} , is defined as $(-1)^{i+j} M_{ij}$. Hence, $D = \sum_{i=1}^{n} a_{ij} C_{ij} \ \ (j = 1, 2, \cdots, n)$ where C_{ij} is a determinant of order $n - 1$. **Adjoint** of a matrix, written as $adj(A)$, is defined as the transpose of the cofactor matrix of A. Example,

$$det \begin{vmatrix} 2 & -3 & 1 \\ 2 & 0 & -1 \\ 1 & 4 & 5 \end{vmatrix} = 2\,det \begin{vmatrix} 0 & -1 \\ 4 & 5 \end{vmatrix} - (-3)\,det \begin{vmatrix} 2 & -1 \\ 1 & 5 \end{vmatrix} + 1\,det \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix}$$
$$= 2(0 + 4) + 3(10 + 1) + 1(8 - 0) = 49$$

9.6.1 Properties of Determinants

1. The value of the determinant is unchanged if the rows and columns are interchanged.
2. Addition of a multiple of a row to another row does not alter the value of the determinant.
3. A zero row or column renders the value of a determinant zero.
4. A determinant with two identical rows or columns has the value zero. Proportional rows or columns render the value of a determinant zero.
5. Interchange of two rows multiplies the value of the determinant by -1.

6. Multiplication of a row by a non zero constant c multiplies the value of the determinant by c. $det(cA) = c\,det(A)$.

7. A $m \times n$ matrix A has rank $r \geq 1$ *iff* A has a $r \times r$ submatrix whose determinant $\neq 0$.

8. An $n \times n$ square matrix A has rank n *iff* $det A \neq 0$.

9. $det(AB) = det(BA) = det(A)det(B)$

9.6.2 Cramer's Rule

$$
\begin{aligned}
&a_{11}x_1 + a_{11}x_1 + \ldots + a_{1n}x_n = b_1\\
&a_{21}x_1 + a_{21}x_1 + \ldots + a_{2n}x_n = b_2\\
&\vdots\\
&a_{n1}x_1 + a_{n1}x_1 + \ldots + a_{nn}x_n = b_n
\end{aligned}
$$

$$\boxed{x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, \cdots, x_n = \frac{D_n}{D}} \text{ (Cramer's Rule)}$$

where D_k is the determinant obtained by replacing the k^{th} column by the entries $b_1, b_2, \cdots, b_n$.

The proof is simple:

Let $A = [a_1 \;\; a_2 \;\; \cdots a_n]$

where a_i is a column vector.

$$
\text{Let } I_i(X) = \begin{bmatrix} 1 & 0 & \cdots & x_1 & 0 & \cdots & 0\\ 0 & 1 & 0 & x_2 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 1 & 0 & x_n & 0 & \cdots & 1 \end{bmatrix} = [e_1 \;\; e_2 \;\; \cdots \;\; x_i \;\; e_{i+1} \;\; e_n]
$$

$$
\begin{aligned}
AI_i(X) &= [Ae_1 \;\; Ae_2 \;\; \cdots \;\; Ax_i \;\; Ae_{i+1} \;\; Ae_n]\\
&= [a_1 \;\; a_2 \cdots \;\; a_{i-1} \;\; b \;\; a_{i+1} \;\; \cdots \;\; a_n] = A_i(b) \text{ (replace } i^{th} \text{ column of } A \text{ with } b)\\
det(A_i(b)) &= det(A)\, I_i(X) = det(A)\; det(I_i(X)) = det(A)\, x_i\\
&\implies \boxed{x_i = \frac{det(A_i(b))}{det(A)}}
\end{aligned}
$$

Example,

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

$$
det\begin{bmatrix} 1 & 1\\ 2 & -3 \end{bmatrix} = -5, \quad x = -\frac{1}{5}det\begin{bmatrix} 5 & 1\\ -4 & -3 \end{bmatrix} = 11/5, \quad y = -\frac{1}{5}det\begin{bmatrix} 1 & 5\\ 2 & -4 \end{bmatrix} = 14/5
$$

9.7 Inverse of a Matrix

The inverse of a **square** matrix A, denoted by A^{-1} is a $n \times n$ matrix that satisfies the following:

$$AA^{-1} = A^{-1}A = I \,(I \text{ is an } n \times n \text{ unit matrix})$$

If A^{-1} exists,$A$is called a **non-singular** matrix, else it is called a **singular** matrix. If the inverse exists, it is always **unique**. A has an inverse *iff* rank $A = r$.

9.7.1 Inverse by Gauss Jordan method

To determine A^{-1},

1. Create augmented matrix $\tilde{A} = [A\ I]$ of size $n \times 2n$.
2. Apply Gauss elimination to $\tilde{A}$ to reduce to upper triangular form $[UH]$.
3. Eliminate the entries of U above the diagonal and make the diagonal entries 1 to get to arrive at the form $[I\,K]$.
4. Then, $A^{-1} = K$

9.7.2 Inverse by Cofactors

$$\begin{aligned}
A\,adj(A) &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{vmatrix} \\
&= \begin{vmatrix} det(A) & 0 & \cdots & 0 \\ 0 & det(A) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & det(A) \end{vmatrix} = det(A)\,I \\
\implies A\left[\frac{adj(A)}{det(A)}\right] &= I \\
\implies \boxed{A^{-1} = \frac{adj(A)}{det(A)}}
\end{aligned}$$

Example,

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

$$det(A) = 10,\quad cof(A) = \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix},\quad adj(A) = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix},\quad A^{-1} = \frac{1}{10}\begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$$

9.7.3 Property of Matrix Inverse

$(AB)^{-1} = B^{-1}A^{-1}$ because $(AB)(AB)^{-1} = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$ or $I = I$

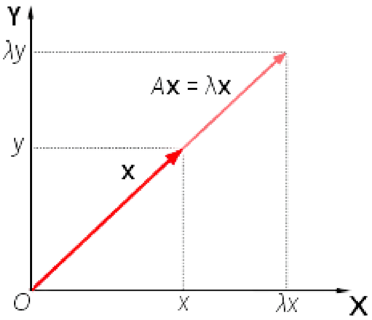
Generalizing, $\boxed{(ABC \cdots PQR)^{-1} = R^{-1}Q^{-1}P^{-1} \cdots C^{-1}B^{-1}A^{-1}}$

9.8 Eigenvalue &Eigenvector

Consider the following system of equations in matrix form.

$$AX = \lambda X \quad \text{(where } A \text{ is a } n \times n \text{ matrix and } \lambda \text{ is a scalar)}$$
$$(A - \lambda I)X = 0$$

The number, i.e., the scalar value λ is an **eigenvalue** of A and X, a non zero vector, is called an **eigenvector** of A. Geometrically, an eigenvector, corresponding to a real nonzero eigenvalue, points in a direction in which it is stretched by the transformation and the eigenvalue is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed. A^T has the same eigenvalues as A.



Using Cramer's rule:

$$\boxed{det(A - \lambda I) = 0}$$

Solve for λ, substitute in equation, and determine x. $det(\lambda)$ is called the **characteristic determinant** and the polynomial is called the **characteristic polynomial.** A $n \times n$ matrix as at least 1 eigenvalue, at most n different eigenvalues.

9.8.1 Algebraic Multiplicity

The algebraic multiplicity of an eigenvalue, μ, is the number of times it appears, i.e., repeated, as a root of the characteristic polynomial. Example,

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 2 \end{bmatrix} \quad det(A - \lambda I) = \begin{bmatrix} 4-\lambda & 2 \\ 1 & 2-\lambda \end{bmatrix} = \lambda^2 - 6\lambda + 6 = 0, \quad \lambda_1 = 3 + \sqrt{3}, \quad \lambda_1 = 3 - \sqrt{3}$$
$$\mu(\lambda_1) = 1, \quad \mu(\lambda_2) = 1$$

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad det(A - \lambda I) = \begin{bmatrix} 1-\lambda & 0 \\ 2 & 1-\lambda \end{bmatrix} = (\lambda - 1)^2 = 0, \quad \lambda_1 = 1, \quad \lambda_2 = 1$$
$$\mu(\lambda_1) = 2, \quad \mu(\lambda_2) = 2$$

9.8.2 Geometric Multiplicity

Eigenspace is the collection of eigenvectors associated with each eigenvalue for the linear transformation applied to the eigenvector. The geometric multiplicity of an eigenvalue is the dimension of the linear space of its associated eigenvectors (i.e., its eigenspace). Example,

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \quad det(A - \lambda I) = \begin{bmatrix} 2-\lambda & 0 \\ 1 & 1-\lambda \end{bmatrix} = (\lambda - 2)(\lambda - 1) = 0, \quad \lambda_1 = 2, \quad \lambda_2 = 1$$
$$\begin{bmatrix} 2-\lambda_1 & 0 \\ 1 & 1-\lambda_1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This yields $x_{11} = x_{21}$ and is non-zero. Hence, the eigenspace of λ_1 is the linear space that contains all vectors of the form $X_2 = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ where α is any non-zero scalar. Thus, the eigenspace of λ_1 is generated by the single vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and so has dimension 1 and the geometric multiplicity of λ_1 is 1.

Note that the second set of equations, corresponding to $\lambda_2 = 1$, yields $x_{12} = x_{22} = 0$ and hence the vector X_2 is not non-zero and is of no use.

Now consider,

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad det(A - \lambda I) = \begin{bmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix} = (\lambda - 2)(\lambda - 2) = 0, \quad \lambda_1 = 2, \quad \lambda_2 = 2$$
$$\begin{bmatrix} 2-\lambda_1 & 0 \\ 1 & 2-\lambda_1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This system of equations is satisfied for any value of x_{11} and x_{21}.

Hence, the eigenspace of λ_1 is the linear space that contains all vectors x_1 are:

$$X_1 = x_{11} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_{21} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

where x_{11} and x_{21} are scalars that can be arbitrarily chosen. Thus, the eigenspace of λ_1 is generated by the two linearly independent vectors $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Hence, it has dimension 2. As a consequence, the geometric multiplicity of λ_1 is 2, equal to its algebraic multiplicity.

9.8.3 Defective eigenvalues

The algebraic and geometric multiplicity of an eigenvalue do not necessarily coincide. When the geometric multiplicity of a repeated eigenvalue is strictly less than its algebraic multiplicity, then that eigenvalue is said to be defective.

An eigenvalue that is not repeated has an associated eigenvector which is different from zero. Therefore, the dimension of its eigenspace is equal to 1, its geometric multiplicity is equal to 1 and equals its algebraic multiplicity. Thus, an eigenvalue that is not repeated is also non-defective.

9.8.4 Real Eigenvalues

Let A be a real symmetric matrix and let λ be a complex eigenvalue of A.

$Ax = \lambda x, x \neq 0$

Taking complex conjugates of both sides, and since A is real we have,

$A\bar{x} = \bar{\lambda}\bar{x}$

Taking transpose and with A as symmetric we have,

$\bar{x}^T A = \bar{\lambda}\bar{x}^T$

$\bar{x}^T Ax = \bar{\lambda}\bar{x}^T x$

$\bar{x}^T \lambda x = \bar{\lambda}\bar{x}^T x$

$\lambda = \bar{\lambda}$

The eigenvalues of a symmetric matrix are real. Similarly, we can establish that the **eigenvalues of a skew-symmetric matrix are pure imaginary or zero.**

9.8.5 Matrix Diagonalization

Two square matrices A and B are said to be **similar** if there exists an invertible P such that,

$$B = P^{-1}AP$$

If two matrices are similar, then they have the same rank, trace, determinant and eigenvalues. Not only two similar matrices have the same eigenvalues, but their eigenvalues have the same algebraic and geometric multiplicities. When A is diagonalizable, then there exists an invertible matrix P such that,

$$D = P^{-1}AP$$

where D is a diagonal matrix.

This is because multiplying the above with P we have,

$$AP = PD$$

Since D is diagonal, P_k is an eigenvector associated with D_{kk}. The matrix P used in the diagonalization must be invertible. Therefore, its columns must be linearly independent. Stated differently, there must be k linearly independent eigenvectors of A.

For some matrices, called defective matrices, it is not possible to find k linearly independent eigenvectors. A matrix is defective when it has at least one repeated eigenvalue whose geometric multiplicity is strictly less than its algebraic multiplicity (called a defective eigenvalue). Therefore, **defective matrices cannot be diagonalized.**

Matrix A is diagonalizable if and only if it does not have any defective eigenvalue. If all the eigenvalues of A are distinct, then A does not have any defective eigenvalue. Therefore, possessing **distinct eigenvalues is a sufficient condition for diagonalizability.**

9.8.6 Positive Definite Matrix

A square matrix A is positive definite if pre-multiplying and post-multiplying it by the same vector x always gives a positive number as a result, independently of how we choose the vector, i.e., $x^T A\ x > 0$. Positive definite symmetric matrices have the property that all their eigenvalues are positive.

9.8.7 Quadratic Form & Positive Definiteness

A quadratic form in A is a transformation $x^T Ax$ and is a scalar. When A is symmetric, we can also write the transformation as $x^T(\frac{A}{2} + \frac{A^T}{2})x$. A is said to be positive definite *iff* $x^T Ax > 0$ for any non-zero x. It is said to be semi positive definite iff $x^T Ax \geq 0$ for any non-zero x.

If A is positive definite, then it is full-rank. A matrix is said to have full rank if its rank equals the largest possible for a matrix of the same dimensions, which is the lesser of the number of rows and columns.