VISPAD INSTITUTE OF TECHNOLOGY

Nea onnim no sua a ohu,

Matrices

$$\begin{bmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \end{bmatrix} \begin{bmatrix} b11 & b12 \\ b21 & b22 \\ b31 & b32 \end{bmatrix} = \begin{bmatrix} c11 & c12 \\ c21 & c22 \end{bmatrix}$$

```
c11 = a11 * b11 + a12 * b21 + a13 * b31
c12 = a11 * b12 + a12 * b22 + a13 * b32
c21 = a21 * b11 + a22 * b21 + a23 * b31
c22 = a21 * b12 + a22 * b22 + a23 * b32
```

Fall 2025

(October – December Virtual Internship)

At the end of the session the students or candidates should be able to understand and work with:

- Matrices and Matrix Arithmetic
- Transposes and Symmetric Matrices
- Linear Systems and Vectors
- Image, Kernel, Rank, Nullity
- Superposition Principles for Linear Systems
- Matrix Inverses
- Linear and Affine Functions
  - Linear function
  - Affine Functions

Matrices

- **Vectors** serve as fundamental blocks, often representing a single data point or a set of features for one observation in machine learning context.
- A vector can hold features such as square footage and number of

bedrooms for one house.

- ❑Machine learning tasks almost always involve working with collections of data points, not just one.
- ❑We need a way to organize and manipulate these collections efficiently and this where matrices come in.

- A matrix is a rectangular grid of numbers, arranged in rows and columns.
- If a matrix has m rows and n rows, we say it has dimensions $m \times n$

Size (sq ft)	Bedrooms	Price ($1000s)
1500	3	300
1200	2	250
1800	4	380
1350	3	290

$$A = \begin{bmatrix} 1500 & 3 & 300 \\ 1200 & 2 & 250 \\ 1800 & 4 & 380 \\ 1350 & 3 & 290 \end{bmatrix}$$

What is the dimension of matrix A?

Why Use Matrices?

- **Compact Representation:** Matrices offer a concise way to store and refer to large amounts of structured data.
- **Standardised Format:** This structure is universally understood and forms the basis for many algorithms and software libraries.

- **Foundation for Operations:** Matrix operations allows complex data transformation and calculations to be done across an entire dataset.

❑ A matrix is a rectangular array of real numbers.

$$\begin{pmatrix} 1 & 0 & 3 \\ -2 & 4 & 1 \end{pmatrix} \qquad \begin{pmatrix} \pi & 0 \\ e & \frac{1}{2} \\ -1 & .83 \\ \sqrt{5} & -\frac{4}{7} \end{pmatrix} \qquad \begin{pmatrix} .2 & -1.6 & .32 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 3 \\ -2 & 5 \end{pmatrix}$$

❑We use the notation

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

for a general matrix of size $m \times n$, where m denotes the number of rows and n the number of columns.

- A matrix is square if $m = n$, i.e., it has the same number of rows and columns.
- A *column vector* is an $m \times 1$ matrix, while a *row vector* is a $1 \times n$ matrix.
- An $m \times n$ matrix contains m column vectors in $\mathbb{R}^n$ and n row vectors having m entries each.

- A 1×1 matrix is both a column and a rows (scalar)
- The number that lies in the $i - th$ row and the $j - th$ column of A is called the (i, j) entry of A, and is denoted by a_{ij}
- The row index always appear first and the column index second.

- Two matrices are equal, $A = B$, if and only if they have the same size and all their entries are the same: $a_{ij} = b_{ij}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$.

Basic Matrix Operations

- Matrix arithmetic involves three basic operations: *matrix addition*, *scalar multiplication*, and *matrix multiplication*.

- One is allowed to add two matrices if and only if they are of the same size.
- **Matrix addition**, is performed entry by entry.

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

- If A and B are $m \times n$ matrices, their sum $C = A + B$ is the $m \times n$ matrix whose entries are given by $c_{ij} = a_{ij} + b_{ij}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$

- **Scalar multiplication** takes a scalar $c \in \mathbb{R}$ and an $m \times n$ matrix A and computes the $m \times x$ matrix $B = cA$ by multiplying each entry of A by c.

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

- In general, $b_{ij} = ca_{ij}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$

- The product of a row vector $\boldsymbol{v}^T$ and a column vector w having the same number of entries is the scalar or 1×1 matrix defined by the following rule:

$$\mathbf{v}^T\mathbf{w} = (v_1, v_2, \ldots, v_n) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i.$$

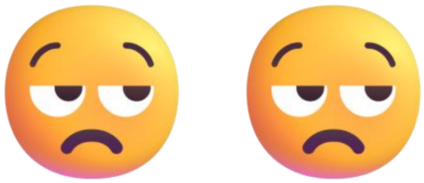
- The matrix product of a row and column vector is the same as the dot product between the corresponding column vectors.

$$\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{w}^T \boldsymbol{v} = \boldsymbol{w} \cdot \boldsymbol{v}$$

- It should be emphasized that the **matrix product** between two column vectors $v, w \in \mathbb{R}^n$ is **not defined**, except in the scalar case $n = 1$ when it coincides with the multiplication in $\mathbb{R}$.

- If A is an $m \times n$ matrix and B is an $n \times p$ matrix, so that the number of *columns* in A equals the number of *rows* in B, then the matrix product $C = AB$ is defined as the $m \times p$ matrix whose (i, j) entry equals the product of the $i - th$ row of A and the $j - th$ column of B.

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \, b_{kj}, \qquad i = 1, \ldots, m, \qquad j = 1, \ldots, n.$$



- The bad news is that matrix multiplication is not commutative - BA is not necessarily equal to AB.

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3, \qquad \begin{pmatrix} 3 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 5 \\ -4 & 11 \end{pmatrix} \neq \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

🤗 🤗 🤗

- Matrix multiplication is **associative**, so $A(BC) = (AB)C$ whenever A has size $m \times n$, B has size $n \times p$, and C has size $p \times q$; the result is a matrix of size $m \times q$.

🤗 🤗 🤗

❑Matrix multiplication is also distributive over matrix addition:

$$A(B+C)=AB+AC$$

$$(A+B)C=AC+BC,$$

for matrices of the appropriate size.

- The difference between *matrix algebra* and *ordinary algebra* is that you need to be careful not to change the order of multiplicative factors without proper justification.
- Matrix multiplication acts by multiplying rows by columns.

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

$$\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 8 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

- In general, if we use b_k to denote the $k-th$ column of B, then

$$AB = A\left(\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \ldots \ \ \boldsymbol{b}_p\right) = \left(A\boldsymbol{b}_1 \ \ A\boldsymbol{b}_2 \ \ \ldots A\boldsymbol{b}_p\right),$$

Indicating that the $k-th$ column of their matrix product is Ab_k

- Multiplying $m \times n$ matrix A by the standard basis vector $\boldsymbol{e}_j \in R^n$ produces the $j-th$ column $\boldsymbol{v}_j = A\boldsymbol{e}_j$ of A.
- Individual entries of a matrix A can be obtained by multiplying it on the left and the right by the standard basis vector:

$$a_{ij} = \boldsymbol{e}_i^T A \boldsymbol{e}_j = \boldsymbol{e}_i \cdot \left(A\boldsymbol{e}_j\right),\, e_i \in \mathbb{R}^m, e_j \in \mathbb{R}^n$$

- Matrix multiplication AB is defined by multiplying rows of A by columns of B.
- This can be also achieved by multiplying columns A by rows of B.
- Suppose that A is an $m \times n$ matrix with columns $v_1, \dots, v_n$ and B is an $n \times p$ matrix with rows $w_1^T, \dots, w_n^T$, where $w_1, \dots, w_n \in \mathbb{R}^p$.

$$AB = v_1 w_1^T + v_2 w_2^T + \cdots + v_n w_n^T,$$

where each summand is a matrix of size $m \times p$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & -3 \end{pmatrix} + \begin{pmatrix} 4 & 6 \\ 8 & 12 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 9 \end{pmatrix}$$

- There are **two** important special matrices.
- The first is the zero matrix, all of whose entries are 0.
- We use $O_{m\times n}$ to denote the $m \times n$ zero matrix
- The zero matrix is the additive unit, so $A + O = A = O + A$ when O has the same size as A.

- The ***identity matrix*** has its entries along the main diagonal – which runs from top left to bottom right – equal to 1, while the off-diagonal entries are all 0.

$$\mathrm{I} = \mathrm{I}_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

- The columns I are the standard basis vectors of $\mathbb{R}^n$.
- If A is any $m \times n$ matrix, then $I_m A = A = A I_n$. This equation can be generalized as: $IA = A = AI$.
- The identity matrix is a particular example of a **diagonal matrix**.
- A **square matrix** A is said to be **diagonal** if all its off-diagonal entries are zero: $a_{ij} = 0$ for all $i \neq j$

- $D = diag(c_1, \dots, c_n) = diag\ \boldsymbol{c}$, where $\boldsymbol{c} = (c_1, \dots, c_n)^T \in \mathbb{R}^n$ for the $n \times n$ diagonal matrix with diagonal entries $d_{ii} = c_i$

$$diag(1, 0, 3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

- The $n \times n$ identify matrix $I = diag\ \mathbf{1}$ is the diagonal matrix associated with all the ones vector $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^n$.
- A square matrix is said to be ***upper triangular*** if all its entries below the main diagonal are zero.

$$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

- A square matrix is said to be ***lower triangular*** if all its entries above the main diagonal are zero.

$$L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -8 & 1 & 0 \\ 2 & -5 & 1 \end{bmatrix}$$

Transposes and Symmetric Matrices

- Transposing a matrix is to interchange its rows and columns.
- If A is an $m \times n$ matrix, then its transpose, denoted by A^T, is the $n \times m$ matrix whose (i, j) entry equals the (j, i) entry of A; thus $B = A^T$ means that $b_{ij} = a_{ij}$, $i = 1, \dots, m, j = 1, \dots, n$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

- ❑The rows of A become the column of A^T and vice versa.
- ❑If A is a 1×1 matrix, i.e., a scalar, then $A^T = A$.
- ❑In general, transposing twice returns you to where you started: $(A^T)^T = A$

- The transpose of an upper triangular matrix is lower triangular and vice versa.
- The transpose operation does not alter a diagonal matrix.

- Transposition is compatible with ***matrix addition*** and ***scalar multiplication:***

$$(A+B)^T = A^T + B^T,$$

$$(cA)^T = cA^T$$

- It is also compatible with matrix multiplication, but reverses the order:

$$(AB)^T = B^T A^T$$

- More generally, if $(A_1 A_2 \cdots A_{k-1} A_k)^T = A_k^T A_{k-1}^T \cdots A_2^T A_1^T$.
- This is used to find the product of a row vector v^T and a column vector w with the same number of entries.

$$v^T w = (v^T w)^T = w^T v, \qquad v\ w \in \mathbb{R}^n$$

- A matrix S is called ***symmetric*** if it equals its own transpose: $S = S^T$.
- A symmetric matrix must be square.
- S is symmetric if and only if it is square and its entries satisfy $s_{ji} = s_{ij}$ for all i, j.
- In other words, entries lying in “mirror image” positions relative to the main diagonal must be equal.

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$$

- Which of these matrices is symmetric?

- **Note that all diagonal matrices, including the identity matrix, are symmetric.**

Linear Systems and Vectors

- If A is an $m \times n$ matrix, and x is a column vector in $\mathbb{R}^n$, then the product $A\boldsymbol{x}$ is a column vector in $\mathbb{R}^m$.
- Let $\boldsymbol{b} \in \mathbb{R}^m$ be another vector.

$$A\boldsymbol{x} = \boldsymbol{b}$$

- The above equation is equivalent to:

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

This consist of m linear algebraic equations in n unknowns, in which A, with entries a_{ij}, is the *coefficient matrix*, $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T$ is a column vector containing the unknowns, while $\boldsymbol{b} = (b_1, b_2, \dots, b_m)^T$ is the column vector containing the right-hand sides.

- A *homogeneous linear system* has the right sides all 0, and can be written in vectorial form as $A\boldsymbol{x} = \mathbf{0}$, where $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$.

- For the linear system

$$x + 3y + 2z - 2 = 0, \qquad 6y + z + 4w = 3,$$

$$-x - 3z + 2w = 1$$

the coefficient matrix, vector of unknowns, and right hand side are

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 6 & 1 & 4 \\ -1 & 0 & -3 & 2 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$$

- Given a set of vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in \mathbb{R}^m$, one can form an $m \times n$ matrix $A = (\boldsymbol{v}_1 \dots \boldsymbol{v}_n)$ with the indicated columns.

$$A\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n, \qquad \text{where} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 4 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 3x_2 \\ -x_1 + 2x_2 + x_3 \\ 4x_1 - x_2 - 2x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^m$, and let $A = (\boldsymbol{v}_1 \ldots \boldsymbol{v}_n)$ be the corresponding $m \times n$ matrix:

a. The vectors are linearly dependent if and only if there is a nonzero solution $\boldsymbol{x} \neq \boldsymbol{0}$ to the homogeneous linear system $A\boldsymbol{x} = \boldsymbol{0}$.

b. The vectors are linearly independent if and only if the only solution to be homogenous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is the trivial one, $\boldsymbol{x} = \boldsymbol{0}$.

c. A vector $\boldsymbol{b}$ lies in the span of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ if and only if the linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.

Given the vectors $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$, $v_3 = \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix}$, the

corresponding matrix is $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & -2 & 3 \end{pmatrix}$. Setting

$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, the linear system: $A\boldsymbol{x} = \boldsymbol{b}$

The linear systems are: $x_1 - x_3 = b_1, \quad x_2 - 2x_3 = b_2, \quad x_1 - 2x_2 + 3x_3 = b_3$

Solving the for x_1 and x_2 and substituting them into their equation we will have: $b_1 - 2b_2 + b_3 = 0$

This is a compatibility condition that needs to be imposed on the right hand side of the system in order that there be a solution.

- Setting $b_1 = b_2 = b_3 = 0$, the solution to the homogeneous system $A\boldsymbol{x} = \mathbf{0}$ is $x_1 = x_3, x_2 = 2x_3$, where x_3 is a "free variable" that can assume any value.
- The homogeneous system admits nonzero solutions, implying that the vector $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are linearly dependent.

- Not every linear system has a solution;
- Every homogeneous linear system has at least one solution, namely when all the unknowns are equal to zero.
- Gaussian elimination for solving linear system is unable to accurately and efficiently handle many of the large linear systems that arise in applications

to data science and machine learning.

Image, Kernel, Rank, Nullity

- ❑Let $A = (\boldsymbol{v}_1 \dots \boldsymbol{v}_2)$ be an $m \times n$ matrix whose columns $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ form a set of n vectors in $\mathbb{R}^m$.
- ❑The subspace spanned by its column vectors is known as the image of A, and denoted by

$$img\ A = span\{v_1, \dots, v_n\} \subset \mathbb{R}^m$$

- ❑ The image is also known as ***column space.***

- A vector $b \in \mathbb{R}^m$ belongs to $img\ A$ if can be written as a linear combination, $b = x_1 v_1 + \cdots + x_n v_n$
- The dimension of the image subspace provides an important numerical quantity called **rank.**
- The **rank** of a matrix A is the dimension of its image:

$$rank\ A = \dim img\ A$$

- Since $img\ A\ \subset \mathbb{R}^n$, we have $0 \leq rank\ A \leq n$

- The only matrix of rank 0 is the zero matrix: $rank\ O = 0$, with $img\ O = \{\mathbf{0}\}$.
- A second important subspace consists of all vectors in $\mathbb{R}^n$ that are annihilated, i.e., sent to zero, when multiplied by A.
- It is known as the *kernel* or *null space of A* and denoted by: $\ker A = \{z \in \mathbb{R}^n | A\boldsymbol{z} = \mathbf{0}\} \subset \mathbb{R}^n$

- The ***kernel*** is the set of solutions z to the homogeneous linear system $A\boldsymbol{z} = \boldsymbol{0}$.
- The ***nullity*** of a matrix A is the dimension of its kernel: $nullity\ A = \dim \ker A$
- The rank and nullity are directly related by:

$rank\ A + nullity\ A = n$ if a A is $m \times n$ matrix.

Consider the 2×3 matrix

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

What is it image, rank, kernel and nullity?

Consider a matrix $A_{4\times 3}$

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

What is it image, rank, kernel and nullity?

- Let A be a square $n \times n$ matrix. Then A is said to be nonsingular if its rank is maximal, namely $rank\ A = n$.

- Let A be a square $n \times n$ matrix. Then the following are equivalent:

  a) A is *nonsingular*;

  $b)\, rank\ A = n$

  $c)\, nullity\ A = 0$

  $d)\, img\ A = \mathbb{R}^n$

  $e)\, \ker A = \{\mathbf{0}\}$

- If A is a nonsingular square matrix, so is A^T.
- The $rank\ A = n = rank\ A^T$

Matrix Inverses

- Many problems in machine learning involve solving systems of linear equations, often represented in the compact form $Ax = b$ where A is a matrix of coefficients, x is a vector of unknowns, and b is a vector of target values or outcomes.

- In linear algebra, solving equation: $5x = 10$ will involve multiplying both sides by the reciprocal of 5, which is ${}^{1}/_{5}$ or 5^{-1}.
- This gives $(5^{-1})5x = (5^{-1})10$
- This simplifies to: $1x = 2$
- The 5^{-1} is the multiplicative inverse of 5 because $5 \times 5^{-1} = 1$.

- The ***matrix inverse*** “undoes” the effect of ***matrix***.
- For a given square matrix A, its inverse, denoted as A^{-1}, is a matrix such when multiplied by A (in either order), the result is the identity matrix I.

$$AA^{-1} = A^{-1}A = I$$

- The identity matrix I acts like the number 1 in matrix multiplication

$$AI = IA = A$$

- The matrix inverse is defined only for square matrices.
- However, not all square matrices have an inverse.
- Matrices that do have an inverse are called invertible or non-singular.

- Matrices that do not have an inverse are called non-invertible or singular.

Properties of the Inverse

- **Inverse of the inverse:** The inverse of A^{-1} is A itself.

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

- **Inverse of a Product:** The inverse of a product of two invertible matrices is the product of their inverses in reverse order. $(AB)^{-1} = A^{-1}B^{-1}$

- **Inverse of a Transpose:** The inverse of the transpose of a matrix is the transpose of its inverse.

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


Thank You