R Programming for Data Science Workshop

Part 01 – Fundamentals of R

H.M. Samadhi Chathuranga Rathnayake

B.Sc(Hons).Special in Industrial Statistics (1st Class) (UOC),
B.Eng (Hons) in Software Engineering (LMU),
CLSSWB, Dip SE, Dip IT, Dip IT & E-Com, Dip B.Mgt, Dip HRM, Dip Eng

Topics will be discussed

1. Basic Operations & Data Types
2. Compound Data Structures
3. Decisions & Loops
4. Functions & libraries
5. Handling Strings

Mathematical Operations

Addition	+
Subtraction	-
Multiplication	*
Division	/
Integer Division	%/%
Power	** or ^
Modulus Operator	%%

Relational Operations

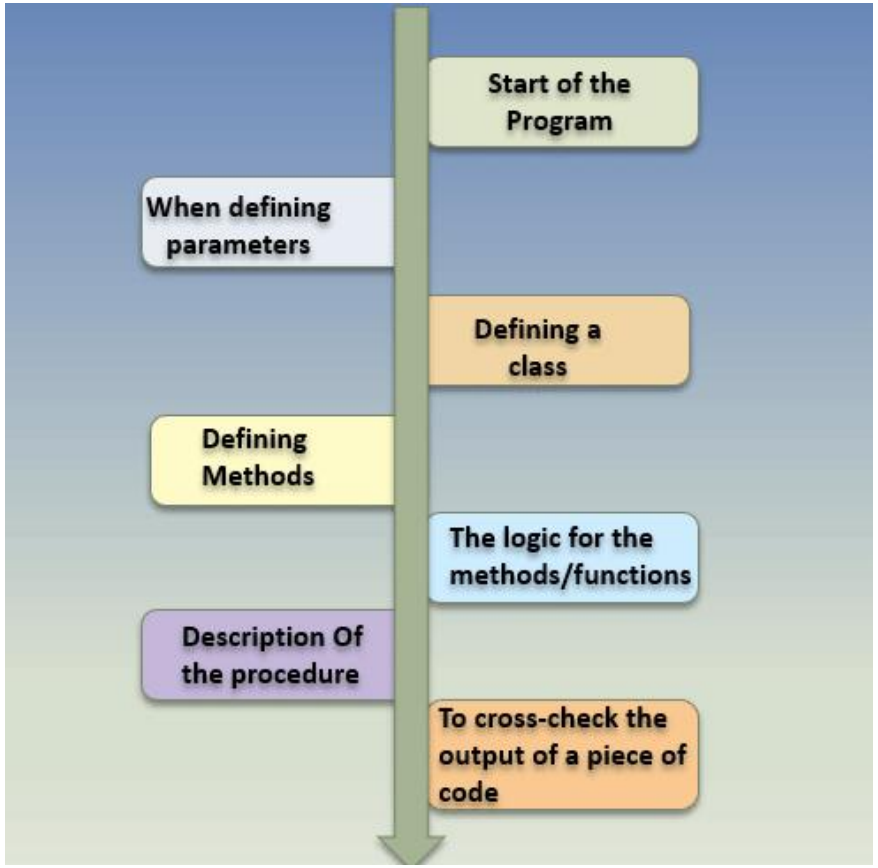
Greater than	>
Less than	<
Greater than or equal	>=
Less than or equal	<=
Equal	==
Not equal	!=

Logical Operations

And	&&
Or	||
Not	!

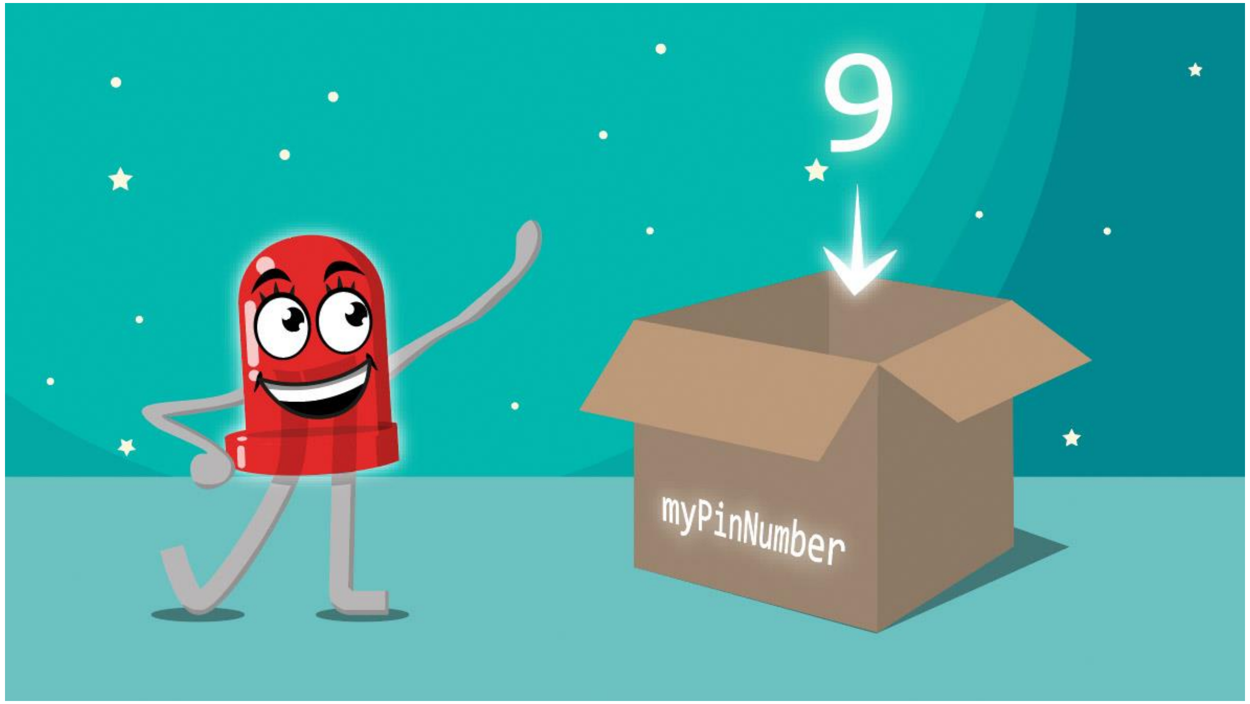
Comments

In computer programming, a comment is a programmer-readable explanation or annotation in the source code of a computer program. They are added with the purpose of making the source code easier for humans to understand, and are generally ignored by compilers and interpreters.



Variables

A variable is a reserved memory location to store values. In other words, a variable in a R program gives data to the computer for processing.



Primitive Data Types

R has 5 primitive data types:

- Numeric
- Integer
- Character
- Logical
- Complex

Data Type	Examples
Numeric	2, 3, 4.456
Integer	3L, 2L
Character	“Cat”, “Dog”, “A”, “Rainy Day”
Logical	TRUE, FALSE, T, F
Complex	2+3i, 4-5i

Compound Data Types

R has 5 compound data structures:

- Vectors
- Matrices & Arrays
- Lists
- Factors
- Data Frames

Compound Data Types

Vectors

- Array of values.

Ex:-

1st Index	2nd Index	3rd Index	4th Index	5th Index
21	32	45	10	50

Compound Data Types

Matrices & Arrays

- Collection of values with rows and columns.

Ex:-

$$\begin{array}{cc} & \text{Columns} \\ & \downarrow \\ \text{Rows} \rightarrow & \begin{bmatrix} 3 & 2 & 5 \\ 3 & 7 & 8 \\ -2 & 10 & 6 \end{bmatrix} \end{array}$$

Compound Data Types

Lists

- It can contain any type of data or any data structure together.

Ex:-

A	$[2, 3, 4, 6]$
B	"Cat"
C	$\begin{bmatrix} 3 & 2 & 5 \\ 3 & 7 & 8 \\ -2 & 10 & 6 \end{bmatrix}$

Compound Data Types

Factors

- This is the way of storing data in categorical format.

Ex:-

<table>
<tr><td>“Male”</td><td rowspan="3">4 values in Male category</td></tr>
<tr><td>“Female”</td></tr>
<tr><td>“Male”</td></tr>
<tr><td>“Male”</td><td rowspan="3">2 values in Female category</td></tr>
<tr><td>“Female”</td></tr>
<tr><td>“Male”</td></tr>
</table>

Compound Data Types

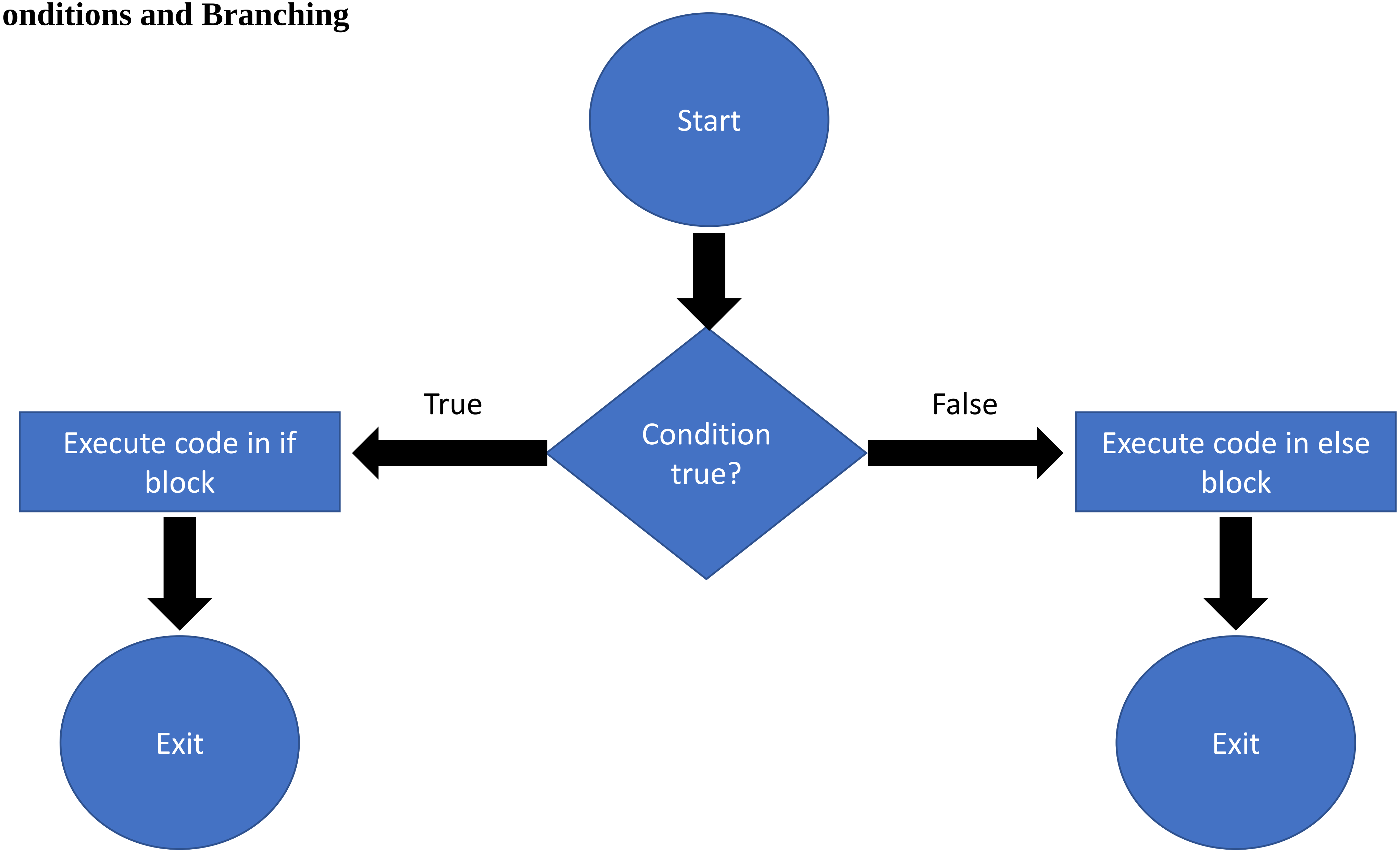
Data Frames

- A table or a two-dimensional array-like structure in which each column contains values of one variable and each row contains one set of values from each column.
- This is a structured way of storing data.
- Simply the structure of a data set.

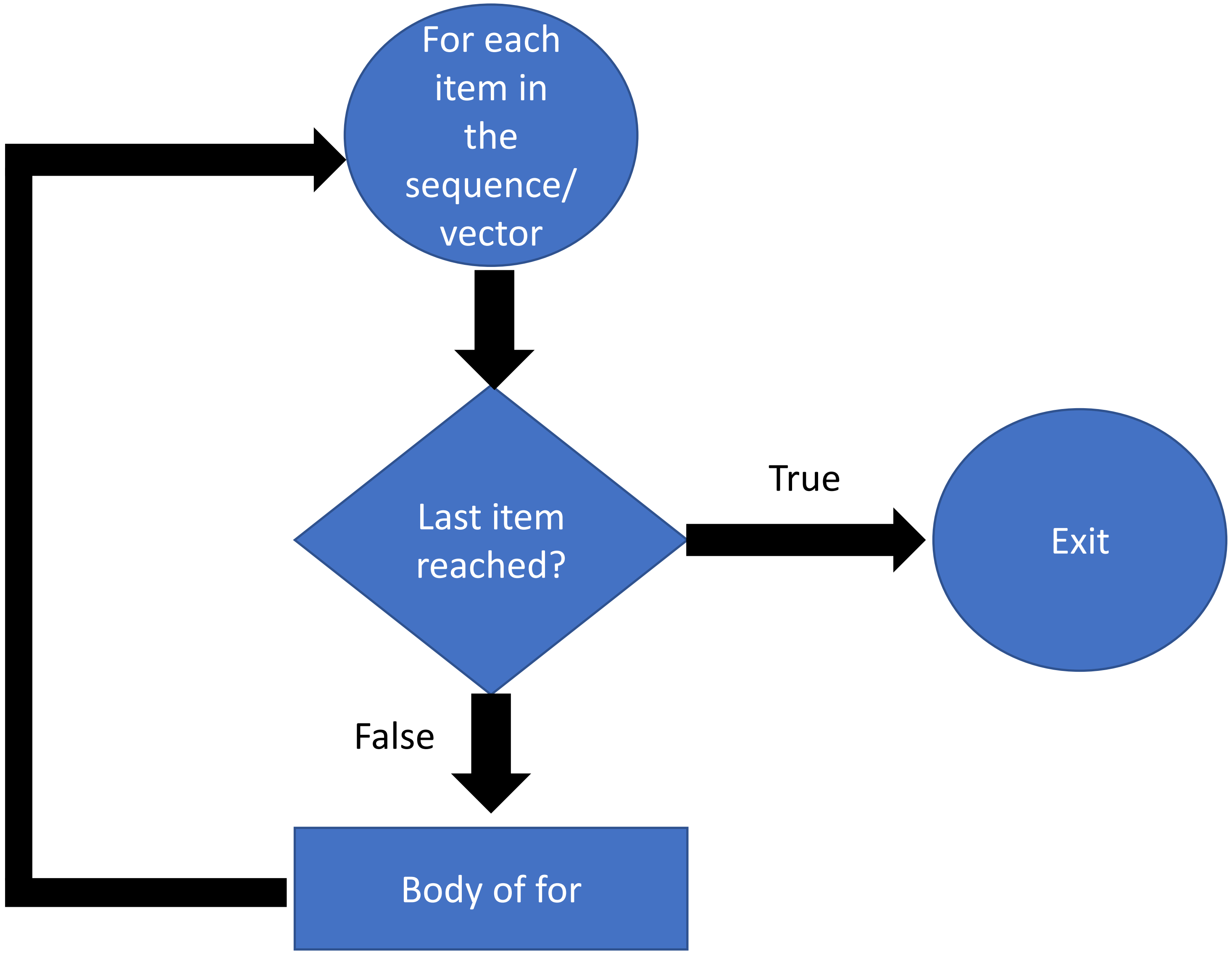
Ex:-

Name	Age	Gender
Sam	26	Male
Kane	24	Male
Jane	23	Female
Peter	21	Male

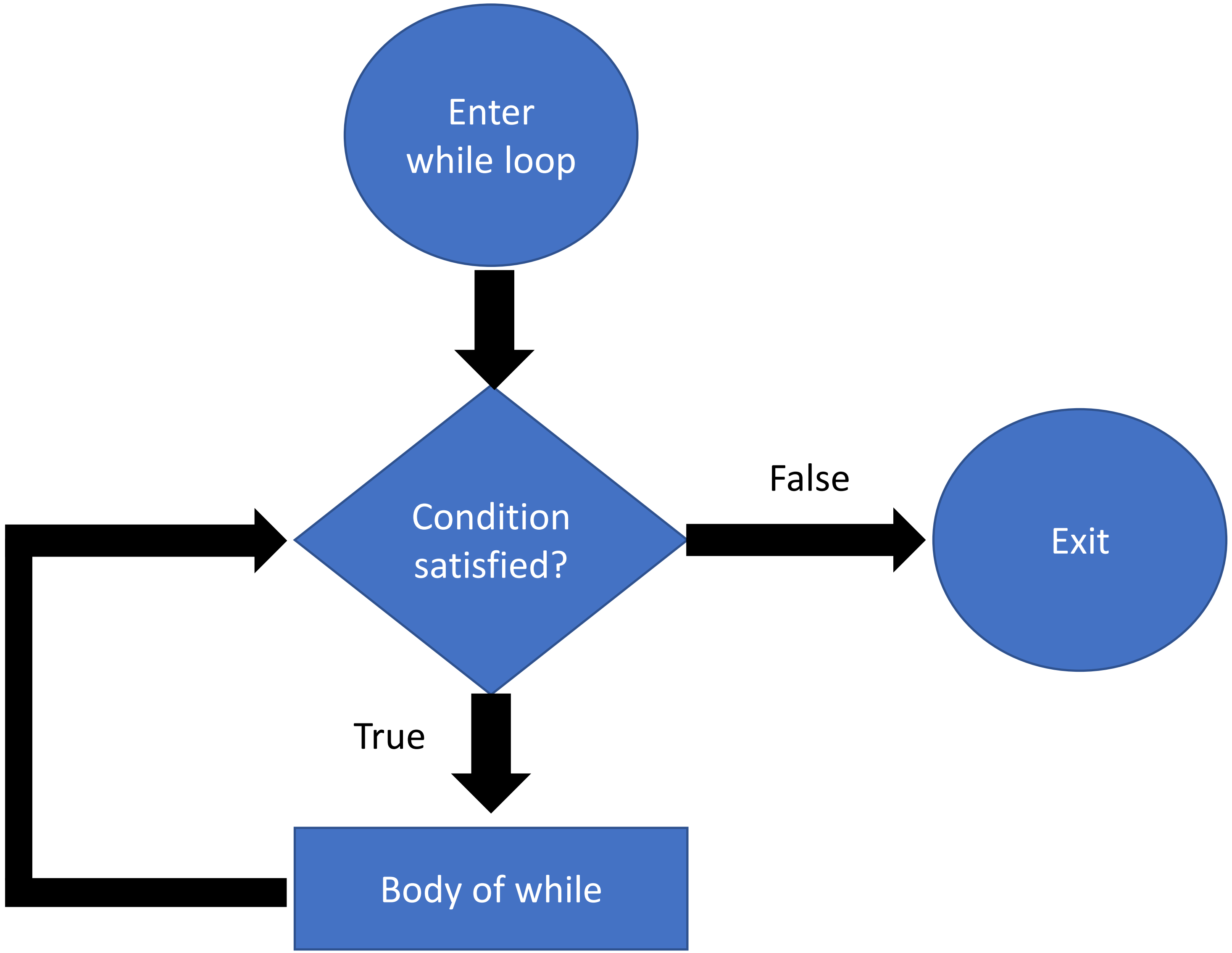
Conditions and Branching



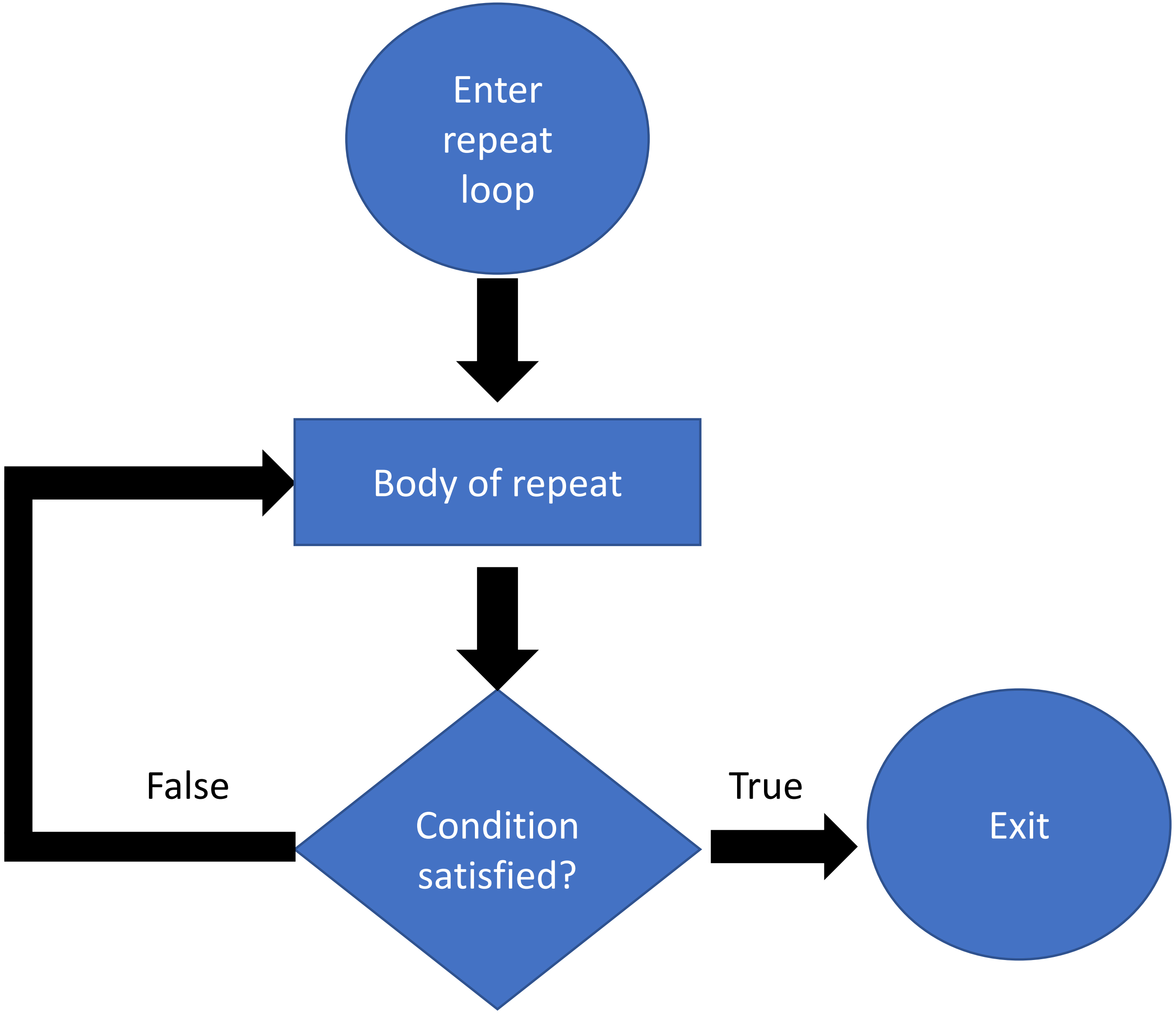
For Loop


For each item in the sequence/ vector
Last item reached?
True
Exit
False
Body of for

While Loop


Enter
while loop
Condition
satisfied?
False
Exit
True
Body of while

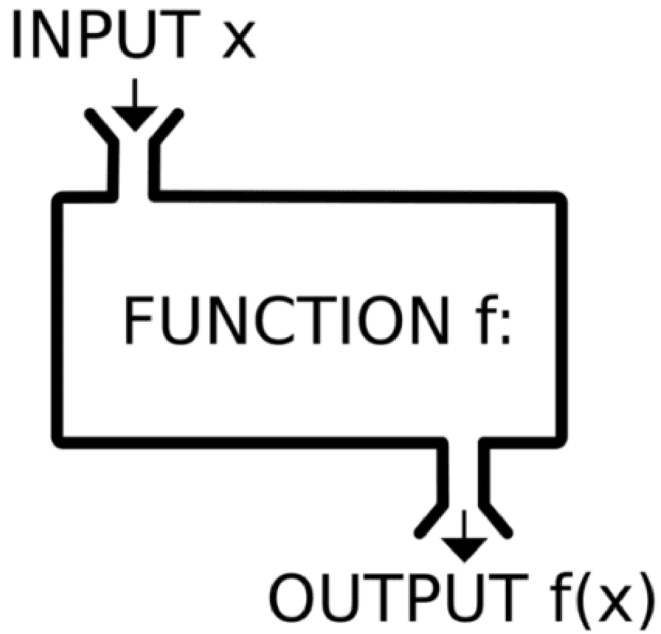
Repeat Loop



Enter repeat loop
Body of repeat
False
Condition satisfied?
True
Exit

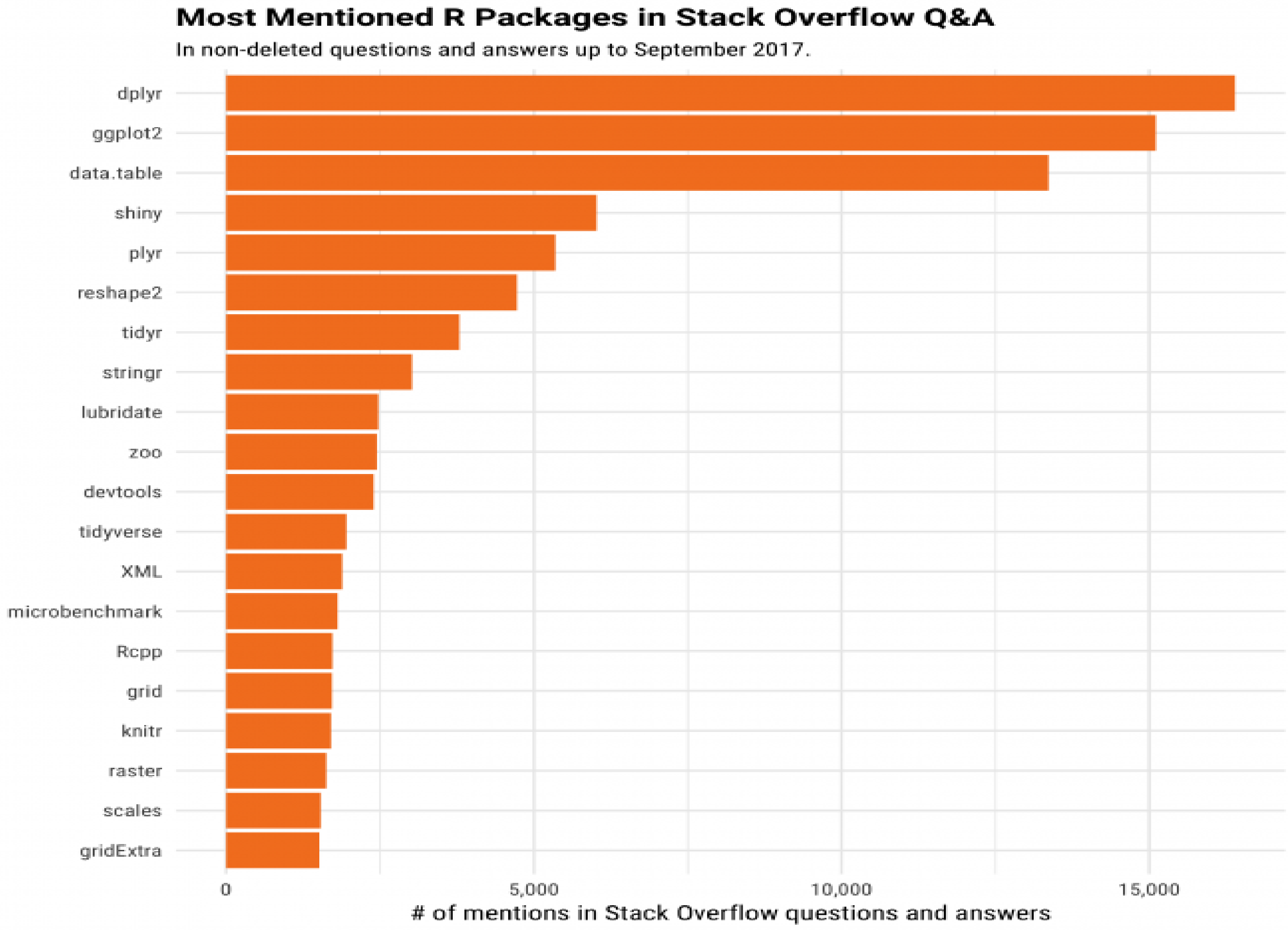
Functions

Functions are repeated blocks of codes in several places in a program.



Libraries in R

Packages are collections of R functions, data, and compiled code in a well-defined format, created to add specific functionality. The directories in R where the packages are stored are called the libraries.



Matrices
(Additional Part)

A matrix is a collection of numbers by rows and columns.

$$\boldsymbol{A} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}_{m\times n}$$

Types of matrices

- A vector is a matrix which has only one row and only one column. If there is only one row it is called as a row vector and if there is only one column it is called as a column vector.

$$\boldsymbol{a} = \begin{bmatrix} 10 \\ 8 \\ 7 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} -2 & 7 & 4 \end{bmatrix}$$

- A scalar matrix is a matrix with only one row and only one column.

$$\boldsymbol{a} = 5$$

- A square matrix is a matrix in which the number of rows equal to the number of columns.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 6 \\ 3 & 2 \end{bmatrix}$$

Types of matrices

- A null matrix which is also called a zero matrix is any matrix in which all the elements are 0.

$$\boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- A diagonal matrix is a square matrix where all the off diagonal elements are 0.

$$\boldsymbol{C} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

- An identity matrix is a diagonal matrix where all the diagonal elements are 1. It is generally denoted by I_n when the dimension is $n \times n$.

$$\boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- A symmetric matrix is a square matrix in which $a_{ij} = a_{ji}$ for all *i* and *j*.

$$\boldsymbol{D} = \begin{bmatrix} 9 & 1 & 5 \\ 1 & 6 & 2 \\ 5 & 2 & 7 \end{bmatrix}$$

Matrix addition & subtraction

$$\boldsymbol{A}_{m\times n} \pm \boldsymbol{B}_{m\times n\times} = \boldsymbol{R}_{m\times n}$$

where $r_{ij} = a_{ij} + b_{ij}$

Ex:-

$$\begin{bmatrix} 2 & 4 & 6 \\ -1 & 3 & 2 \\ 1 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 6 & 9 \\ 10 & -5 & 2 \\ 3 & 4 & 0 \end{bmatrix} =$$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ -2 & 6 \end{bmatrix} \; \boldsymbol{B} = \begin{bmatrix} 4 & 3 \\ 1 & 5 \end{bmatrix} \; \boldsymbol{C} = \begin{bmatrix} 6 & 2 \\ 1 & 3 \end{bmatrix}$$

Find **A**+**B-C** =

Answers

Ex:-

$$\begin{bmatrix} 2 & 4 & 6 \\ -1 & 3 & 2 \\ 1 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 6 & 9 \\ 10 & -5 & 2 \\ 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 10 & 15 \\ 9 & -2 & 4 \\ 4 & 10 & 2 \end{bmatrix}$$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ -2 & 6 \end{bmatrix} \; \boldsymbol{B} = \begin{bmatrix} 4 & 3 \\ 1 & 5 \end{bmatrix} \; \boldsymbol{C} = \begin{bmatrix} 6 & 2 \\ 1 & 3 \end{bmatrix}$$

Find **A**+**B-C** = $\begin{bmatrix} 0 & 4 \\ -2 & 8 \end{bmatrix}$

Matrix multiplication

- Scalar multiplication

$$k\boldsymbol{A} = \boldsymbol{B}$$

such that for all i, j $\quad b_{ij} = ka_{ij}$

Ex:-

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

- Multiplying 2 matrices

$$\boldsymbol{C}_{m\times p} = \boldsymbol{A}_{m\times n} \times \boldsymbol{B}_{n\times p}$$

such that,

$$c_{ij} = \sum_{k=1}^{p} a_{ik} \times b_{kj}$$

Answers

Ex:-

$$2 \times \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ -6 & 0 \end{bmatrix}$$

Ex:-

$$\begin{bmatrix} 2 & 8 & 1 \\ 3 & 6 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 7 \\ 9 & -2 \\ 6 & 3 \end{bmatrix} =$$

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

- Find **AB** and **BA**

Ex:-

$$\begin{bmatrix} 2 & 8 & 1 \\ 3 & 6 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 7 \\ 9 & -2 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 80 & 1 \\ 81 & 21 \end{bmatrix}$$

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

- $\boldsymbol{AB} = \begin{bmatrix} 8 & 5 & 3 \\ 16 & 7 & 5 \\ -4 & -1 & -1 \end{bmatrix}$
- $\boldsymbol{BA} = \begin{bmatrix} 6 & 10 \\ 7 & 8 \end{bmatrix}$

Properties of matrix addition and subtraction

Let **A**, **B** and **C** are $m \times n$ matrices.

- $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$ (Commutative law)
- $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$ (Associative law)
- There is a unique $m \times n$ matrix 0 with $\boldsymbol{A} + \underline{0} = \boldsymbol{A}$ (Additive identity property)
- For any $\boldsymbol{A}_{m\times n}$ matrix there is a unique $\boldsymbol{B}_{m\times n}$ which is $-\boldsymbol{A}_{m\times n}$ matrix such that $\boldsymbol{A} + \boldsymbol{B} = \underline{0}$ (Additive inverse property)

Properties of matrix scalar multiplication

Let A and B are matrices with appropriate dimensions and r and s are real numbers.

- $r(s\boldsymbol{A}) = (rs)\boldsymbol{A}$
- $(r + s)\boldsymbol{A} = r\boldsymbol{A} + s\boldsymbol{A}$
- $r(\boldsymbol{A} + \boldsymbol{B}) = r\boldsymbol{A} + r\boldsymbol{B}$
- $\boldsymbol{A}(r\boldsymbol{B}) = r(\boldsymbol{A}\boldsymbol{B}) = (r\boldsymbol{A})\boldsymbol{B}$

Properties of matrix multiplication

Let **A**, **B** and **C** are matrices with appropriate dimensions.

- $\boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}) = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}$ (Associative law)
- $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$ (Distributive law)
- $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}\boldsymbol{C} + \boldsymbol{B}\boldsymbol{C}$ (Distributive law)
- There are unique matrices $\boldsymbol{I}_m$ and $\boldsymbol{I}_n$ such that $\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$ (Multiplicative identity)

Matrix transpose

Transpose of a matrix is where the rows and columns are interchanged.

$$\boldsymbol{B} = \boldsymbol{A}^T \dashrightarrow b_{ij} = a_{ji}$$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 2 & 7 & 1 \\ 8 & 6 & 4 \end{bmatrix} \qquad \boldsymbol{A}^T =$$

Properties of matrix transpose

- $(\boldsymbol{A}^T)^T = \boldsymbol{A}$
- For a symmetric matrix, $\boldsymbol{A}^T = \boldsymbol{A}$
- $(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$
- $(\boldsymbol{AB})^T = \boldsymbol{B}^T \boldsymbol{A}^T$
- $(r\boldsymbol{B})^T = r\boldsymbol{B}^T$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 2 & 7 & 1 \\ 8 & 6 & 4 \end{bmatrix} \qquad \boldsymbol{A}^T = \begin{bmatrix} 2 & 8 \\ 7 & 6 \\ 1 & 4 \end{bmatrix}$$

Determinant of a matrix

Determinant is defined only for a square matrix.

$$|\boldsymbol{A}| = det(\boldsymbol{A})$$

- For a scalar matrix,

$$\boldsymbol{A} = a \dashrightarrow |\boldsymbol{A}| = a$$

- In 2×2 case,

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \dashrightarrow |\boldsymbol{A}| = ad - bc$$

- In 3×3 case

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \dashrightarrow |\boldsymbol{A}| = a(ei - fh) - b(di - gf) + c(dh - ge)$$

Ex:- Find the determinant of following matrices.

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

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

Ex:- Find the determinant of following matrices.

$$\boldsymbol{A} = \begin{bmatrix} 2 & 4 \\ -3 & 2 \end{bmatrix} -----\rightarrow \det(A) = 16$$

$$\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 0 & 1 \\ 1 & -1 & 3 \end{bmatrix} -----\rightarrow \det(B) = 1(0+1) - 2(9-1) + 1(-1-0) = -16$$

Trace of a matrix

Sum of the diagonal elements of a matrix is called the trace of the matrix.

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 3 & 7 & 2 \\ -1 & 6 & 4 \\ 9 & 0 & -5 \end{bmatrix} \qquad tr(\boldsymbol{A}) =$$

Properties of the trace of a matrix

- If A I scalar $tr(\boldsymbol{A}) = \boldsymbol{A}$
- $tr(\boldsymbol{A}^T) = tr(\boldsymbol{A})$
- $tr(k\boldsymbol{A}) = k \times tr(\boldsymbol{A})$ when k is a real value
- $tr(\boldsymbol{I_n}) = n$
- $tr(\boldsymbol{A} \pm \boldsymbol{B}) = tr(\boldsymbol{A}) \pm tr(\boldsymbol{B})$
- $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 3 & 7 & 2 \\ -1 & 6 & 4 \\ 9 & 0 & -5 \end{bmatrix} \qquad tr(\boldsymbol{A}) = 4$$

Inverse of a matrix

The inverse of a matrix A is denoted by $\boldsymbol{A}^{-1}$, such that,

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$$

To have an inverse of a matrix, the column vectors should be independent.

Ex:- Check whether matrix **A** has independent column vectors.

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

Adjoint of a matrix

Let $\boldsymbol{A} = [a_{ij}]$ be a square matrix of order n. The adjoint of a matrix **A** is the transpose of cofactor matrix of **A**

$$Adj(\boldsymbol{A}) = [cof(a_{ij})]^T$$

$$cof(a_{ij}) = (-1)^{i+j}\det(M_{ij})$$

Ex:- Find the adjoint of the following matrix.

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

Ex:- Find the adjoint of the following matrix.

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

$$\boldsymbol{Cofactor\ matrix\ of\ A} = \begin{bmatrix} (-1)^{1+1}(0-6) & (-1)^{1+2}(-4-0) & (-1)^{1+3}(4-0) \\ 1 & -1 & -1 \\ -6 & 2 & 4 \end{bmatrix}$$

$$\boldsymbol{Cofactor\ matrix\ of\ A} = \begin{bmatrix} -6 & 4 & 4 \\ 1 & -1 & -1 \\ -6 & 2 & 4 \end{bmatrix}$$

$$\boldsymbol{Adj(A)} = \begin{bmatrix} -6 & 1 & -6 \\ 4 & -1 & 2 \\ 4 & -1 & 4 \end{bmatrix}$$

Matrix inverse through adjoint

The inverse of the matrix can be found through,

$$A^{-1} = \frac{Adj(A)}{|A|}$$

To have an inverse the matrix should be a non-singular matrix.

Ex:- Find the inverse of following matrix,

$$\boldsymbol{B} = \begin{bmatrix} 1 & 5 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Ex:-

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

$$\boldsymbol{det}(\boldsymbol{A}) = -2$$

$$\boldsymbol{Adj}(\boldsymbol{A}) = \begin{bmatrix} -6 & 1 & -6 \\ 4 & -1 & 2 \\ 4 & -1 & 4 \end{bmatrix}$$

$$A^{-1} = \frac{Adj(A)}{|A|}$$

$$A^{-1} = \begin{bmatrix} 3 & -0.5 & 2 \\ -2 & 0.5 & -1 \\ -2 & 0.5 & -2 \end{bmatrix}$$

- For a 2×2 matrix the inverse is particularly simple.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

The inverse is,

$$\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Ex

$$\boldsymbol{B} = \begin{bmatrix} 12 & 10 \\ 3 & 7 \end{bmatrix}$$

- If A is a diagonal matrix, Suppose **A** is given as follows,

$$\boldsymbol{A} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}$$

Properties of matrix inverse

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