LINEAR ALGEBRA

* A linear equation is one where all the variables such as x, y, z have index (power) of 1 or 0 only.

  - $x = 3$ (✓)
  - $3x + y + z + w = -8$ (✓)
  - $x + 2y + z = 5$ (✓)
  - $\cos x + \sin y = 1$ (x)
  - $\sqrt{x} + y + z = 6$ (x)
  - $e^{x+y+z} = 1$ (x)
  - $x = -3y$ (✓)
  - $x - 2y + 5z = \sqrt{3}$ (✓)

Vectors

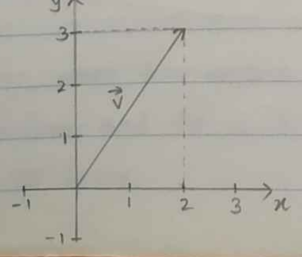
* A vector is a quantity that has both size (magnitude) and direction.
  - Walk due North 5 kilometres.
  - Velocity, acceleration, force and displacement are all vector quantities.

* A scalar is a number that measures the size of a particular quantity.
  - Length, area, volume, mass and temperature are all scalar quantities.

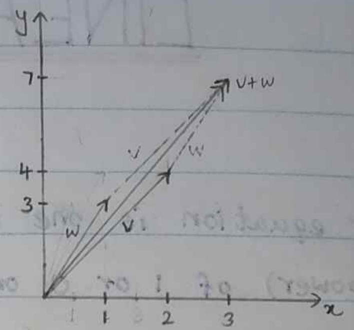
* In 2-dimensional space, a vector can be written as;
  - $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

* In 3-dimensional space, a vector can be written as;
  - $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$

* The graphical representation of $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$;



Vector Addition



$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ $\qquad$ $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$v + w = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$

Scalar Multiplication

* The scalar can be a real number. $(\lambda \in \mathbb{R})$

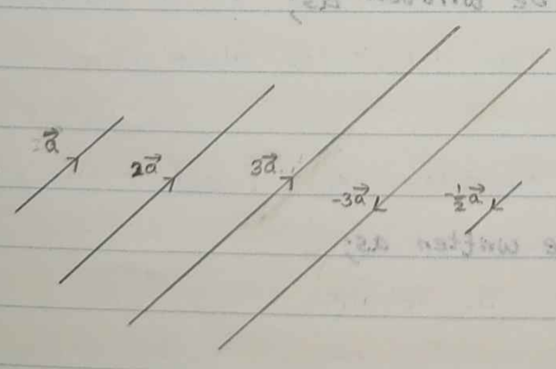
* If $\lambda > 1$, the vector will keep the same direction but stretch.

* If $0 < \lambda < 1$, the vector will keep the same direction but shrink.

* If $\lambda < -1$, the vector will change direction and stretch.

* If $-1 < \lambda < 0$, the vector will change direction and shrink.

$3v = 3\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$



Linear Combination

* Linear algebra is fundamentally built upon two basic operations: vector addition and scalar multiplication.

* Combining those two operations gives linear combinations.

* A linear combination of vectors is an expression constructed from a set of vectors by multiplying each vector by a scalar and then adding the results.

* The linear combinations of v and w are the vectors $cv + dw$ where $c, d \in \mathbb{R}$.
    - $c\begin{bmatrix}2\\4\end{bmatrix} + d\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}2c+d\\4c+3d\end{bmatrix}$

* The linear combinations fill the xy plane which implies that if you take all possible linear combinations of v and w (assuming they are not parallel), you can reach any point in the 2-dimensional xy-plane by shrinking/stretching/flipping the vectors. This is because v and w form a 'basis' for $\mathbb{R}^2$ (xy-plane).

* v and w being a 'basis for $\mathbb{R}^2$' simply means that they are two non-parallel vectors that, through linear combinations, can generate every single other vector (or point) in the 2D xy-plane. They fill up the plane.

* All combinations $c\begin{bmatrix}4\\1\end{bmatrix} + d\begin{bmatrix}2\\-1\end{bmatrix}$ fill the xy plane. They produce every $\begin{bmatrix}x\\y\end{bmatrix}$.

* The vectors $c\begin{bmatrix}1\\1\\2\end{bmatrix} + d\begin{bmatrix}0\\0\\4\end{bmatrix}$ fill a plane in xyz space.

* When there are only two vectors and they are in 3D space, their linear combinations will not be able to fill the entire 3D space because you would typically need three non-coplanar vectors for that.

* However, when two vectors that are in 3D space are not parallel to each other, they will define and fill a 2-dimensional plane that passes through the origin within that 3D space.

* Linear algebra is not limited to 2 vectors in 2-dimensional and 3-dimensional spaces. Linearity allows for higher dimensions and more vectors.

* Instead of just 2 components (like in 2D plane), vectors can have 'm' components, meaning they exist in an m-dimensional space.

* There can be 'n' vectors denoted as $v_1, v_2, v_3, \dots, v_n$.

* When you have 'n' vectors, each with 'm' components ('n' vectors in an 'm' dimensional space), they can be organized into an $m \times n$ matrix A.
    - Each row corresponds to a component of the vectors. Since each vector has m components, the matrix will have m rows.
    - Each column represents one of the n vectors. So, the vectors $v_1, v_2, v_3, \dots, v_n$ become the columns of the matrix.
    - The resulting matrix A will have m rows and n columns.
    - $A = [v_1 \ v_2 \ v_3 \dots v_n]$

$$\underset{m \times n}{A} = \begin{bmatrix} v_1 & v_2 & v_3 & \dots & v_n \end{bmatrix}$$

Describe all the combinations $Ax = x_1 v_1 + x_2 v_2 + x_3 v_3 + \dots + x_n v_n$ of the columns.

* When you multiply matrix A by a vector x (where x is a column vector $[x_1 \ x_2 \ x_3 \dots x_n]^T$), the result Ax is a linear combination of the columns of A. Specifically, it's x_1 times the first column (v_1), plus x_2 times the second column (v_2), and so on, up to x_n times the n^{th} column (v_n). This problem is about understanding the span of the columns of A - what set of all possible vectors can be formed by taking linear combinations of the columns of A.

Find the numbers x_1 to x_n that produce a desired output vector $Ax = b$.

* This is the inverse problem. Given a matrix A and a target vector b, the goal is to find the coefficients $(x_1, x_2, \dots, x_n)$ of the linear combination of the columns of A that equal b. This is the objective of solving a system of linear equations. If such x_i values exist, it means that the vector b is in the column space (or span) of matrix A.

Row Way / Row Picture (System of Scalar Equations)

* In row way, each row of the matrix equation is treated as a separate scalar equation.

  - $v_1 c + w_1 d = b_1$
    $v_2 c + w_2 d = b_2$

* The row picture focuses on each individual equation within the system. Each equation, when plotted in a coordinate system, represents a geometric object (a line in 2D, a plane in 3D, or a hyperplane in higher dimensions). For a system with n variables, each equation corresponds to an n-1 dimensional flat surface in n-dimensional space.

  - If you have 2 variables (x, y), each equation $ax+by=c$ is a line in the 2D plane.
  - If you have 3 variables (x, y, z), each equation $ax+by+cz=d$ is a plane in 3D space.

* The solution to the system is the point (or set of points) where all these geometric objects (lines, planes, hyperplanes) intersect.

* Row picture is very intuitive for 2D and 3D systems, as it directly corresponds to familiar geometric concepts. It gives a clear visual of where the solution lies.

* However, it becomes very difficult, if not impossible, to visualize directly for systems with more than three variables.

Column Way / Column Picture (Linear Combination)

* Column way is expressing the right-hand side vector as a linear combination of the column vectors of the coefficient matrix.
  - $c\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + d\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

* The solution (the values of the variables) tells you how much of each column vector you need to 'mix' or 'scale' to arrive at the target right-hand side vector. Geometrically, you are looking for the path formed by scaling and adding the column vectors to reach the destination vector.

* Column picture is excellent for understanding concepts like linear independence, span, column space and matrix multiplication as a combination of columns.
  - It clarifies when a solution exists. If the target vector b can be reached by a linear combination of column vectors, it means b lies in the column space (the span) of the matrix A. If b is outside this space, there's no solution.
  - This is useful in visualizing linear independence / dependence. If the column vectors are linearly independent (not parallel in 2D), they span the entire plane (or higher-dimensional space), meaning you can reach any b. If they are linearly dependent (parallel in 2D), they only span a line, and you can only reach b vectors that lie on that same line.
  - It builds intuition by helping to visualize how transformations work and how vectors combine.

* While the vectors can be drawn, precisely determining the scaling factors (the solution values) purely by drawing is generally not feasible for accurate results. Algebraic methods are usually needed for computation.

Matrix Form (Matrix - Vector Product)

* Matrix way is representing the entire system compactly as a product of a matrix and a vector in the form $Ax = b$.
  - $\begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

* Finding the solution means finding the vector x that, when multiplied by matrix A, yields the vector b. from a transformation perspective, matrix A can be seen as a 'function' or 'linear transformation' that takes an input vector x and transforms it into an output vector b.

* This method is highly compact and efficient for representing large systems. It's the basis for almost all computational algorithms like Gaussian Elimination, LU decomposition used to solve linear systems. It naturally leads to concepts like matrix inverse, determinants, eigenvalues, and eigenvectors.

Ex

$2x + y = 3$ —— ①

$x - 2y = -1$ —— ②

① × 2 + ② ⇒

$5x = 5$

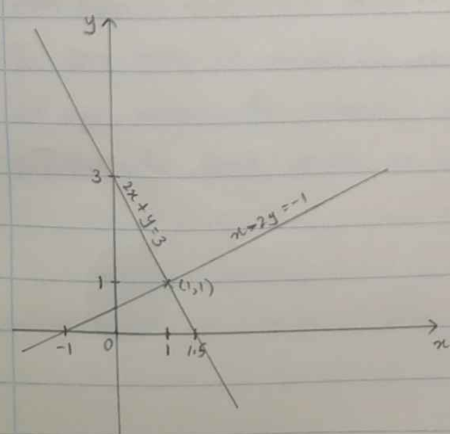
$\underline{\underline{x = 1}}$

$x = 1$ in ①;

$2 + y = 3$

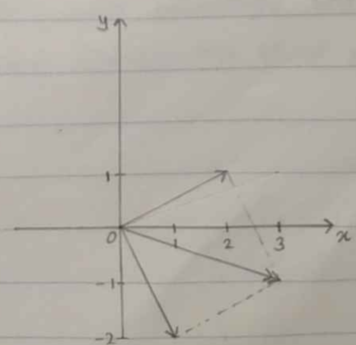
$\underline{\underline{y = 1}}$

Row Picture



Column Picture

$x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$



Matrix Form

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

$\boxed{Ex}$ $2x - y = 0$ —— ①

$-x + 2y = 3$ —— ②

① × 2 + ② ⇒

$3x = 3$

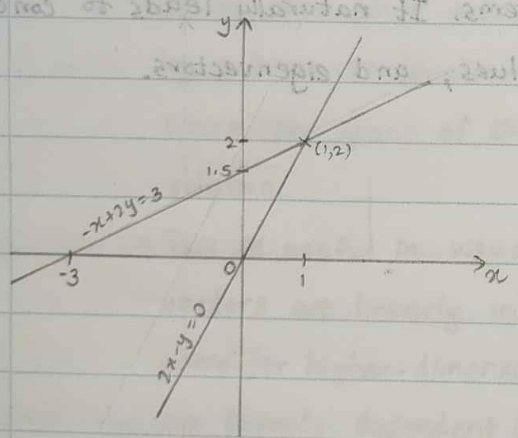
$\underline{\underline{x = 1}}$

$x = 1$ in ①,

$2 - y = 0$

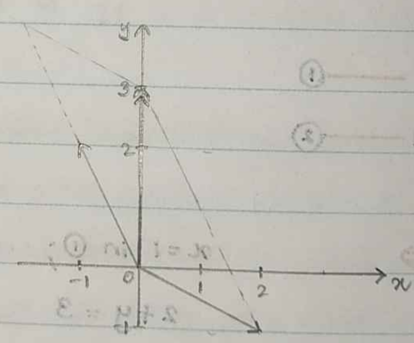
$\underline{\underline{y = 2}}$

Row Picture



Column Picture

$$x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$



Matrix form

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

Zero Vector

* Zero vector is a vector that has a zero magnitude and no direction.

* Suppose two people are pulling a rope from its two ends with equal force but in opposite directions. So, the net force applied to the rope will be a zero vector (null vector) as the two equal forces balance each other out because they are in opposite directions.

Linear Independence

* Two or more vectors are said to be linearly independent if none of them can be written as a linear combination of the others.

* To check for linear independence, we form a linear combination of the vectors and set it equal to the zero vector. This is expressed as;

$$\begin{bmatrix}1\\1\\3\\3\end{bmatrix}\begin{bmatrix}-1\\1\\-1\\-3\end{bmatrix}\begin{bmatrix}2\\5\\9\\6\end{bmatrix} \xrightarrow[\text{combination}]{\text{linear}} c_1\begin{bmatrix}1\\1\\3\\3\end{bmatrix} + c_2\begin{bmatrix}-1\\1\\-1\\-3\end{bmatrix} + c_3\begin{bmatrix}2\\5\\9\\6\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

* If the only solution for the scalar coefficients c_1, c_2 and c_3 is $c_1 = 0$, $c_2 = 0$ and $c_3 = 0$ (the trivial solution), then the vectors are linearly independent. This means no vector can be formed by combining the others.

* If there are other solutions where at least one of c_1, c_2 and c_3 is non-zero (a non-trivial solution), then the vectors are linearly dependent. In this case, it would be possible to express one vector as a linear combination of the others. for example, if $c_1 \neq 0$, you could rearrange the equation to express the first vector in terms of the second and third vectors.

$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $v_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

* Since $v_2 = 2v_1$, they are linearly dependent.
* $2v_1 + (-1)v_2 = 0$

$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

* No scalar multiple of one gives the other. So, they are linearly independent.
* Only solution to $av_1 + bv_2 = 0$ is $a = 0$ and $b = 0$.

$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $v_3 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$

* Since $v_3 = 2v_1 + v_2$, they are linearly dependent.
* $2v_1 + v_2 + (-1)v_3 = 0$

$v_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}$ $v_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ -3 \end{bmatrix}$ $v_3 = \begin{bmatrix} 2 \\ 5 \\ 9 \\ 6 \end{bmatrix}$

* Since $2v_3 = 7v_1 + 3v_2$, they are linearly dependent.
* $7v_1 + 3v_2 + (-2)v_3 = 0$

Geometrical Representation of vectors when dependent and independent.

Scenario	Independent	Dependent
i. 2 2D vectors in 2D space	* The two vectors are not collinear. They point in different directions. Geometrically, they will span the entire 2D plane. Any point in the 2D plane can be reached by a linear combination of these two vectors.	* The two vectors are collinear. They lie on the same line passing through the origin. One vector is a scalar multiple of the other. Geometrically, they will span a 1D line within the 2D space.
ii. 2 2D vectors in 3D space.	* The two vectors are not collinear. They point in different directions within the 3D space. Geometrically, they will span the xy plane.	* The two vectors are collinear. They will span a 1D line that passes through the origin and lies within the xy-plane.
iii. 3 or more 2D vectors in 2D space.	* This scenario is not possible. In a 2D space, you can have a maximum of two linearly independent vectors. If you have three vectors in 2D space, they must always be linearly dependent. This is because the 'dimension' of the space they can span is limited to 2. Once you have two independent vectors spanning the 2D plane, any additional vector in that plane can be written as a linear combination of the first two.	* Any three 2D vectors in 2D space are always linearly dependent. Geometrically, they will span at most the 2D plane. At least one of them can be expressed as a linear combination of the other two.
iv. 3 or more 2D vectors in 3D space	* Same as when the vectors were in the 2D plane but the vectors are confined into the xy plane.	* Any three 2D vectors, even when placed conceptually within 3D space will always be linearly dependent. They will collectively span at most the xy plane (or a 1D line if they are collinear) in 3D space.

v.	2 3D vectors in 3D space	* The two vectors are not collinear. They point in different directions in 3D space. Geometrically, they will span a 2D plane passing through the origin in 3D space.	* The two vectors are collinear. They lie on the same line passing through the origin in 3D space. One is a scalar multiple of the other. Geometrically, they will span a 1D line within the 3D space.
vi.	3 3D vectors in 3D space	* The three vectors are not coplanar. They point in distinct directions such that they do not all lie in the same plane passing through the origin. Geometrically, they will span the entire 3D space. Any point in 3D space can be reached by a linear combination of these three vectors.	* The three vectors are coplanar. They all lie in the same plane passing through the origin. This means one of the vectors can be expressed as a linear combination of the other two. Geometrically, they will span at most a 2D plane (or a 1D line if they are also collinear) within the 3D space, not the entire 3D space.

Invertibility

* An invertible matrix is a square matrix that has a corresponding inverse matrix. When a matrix is multiplied by its inverse, the result is the identity matrix.

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

* An invertible matrix represents a transformation that can be 'undone' or reversed. This property is crucial because it guarantees:

- Unique Solutions – For a system of linear equations $Ax = b$, if A is invertible, there's always a unique solution for x ($x = A^{-1}b$).
- Non-singularity – The determinant of an invertible matrix is always non-zero.
- Linear independence – Its columns and rows are linearly independent, meaning they don't lie on the same line (in 2D) or plane (in 3D), and span the entire space. (Full rank. Square matrix)

* If a matrix is not invertible, it's called singular, implying its transformation collapses some dimension, making it impossible to uniquely reverse.

Square Matrices and Types

* A typical matrix is a rectangle of $m \times n$ numbers where m is the number of rows and n is the number of columns. If m is equal to n, then that matrix is a 'square matrix'.

i. Diagonal matrix

* Let $d_{ij} \in D_n$. Then $d_{ij} = 0$ for $i \neq j$. Therefore, all the elements which are not in the diagonal (top left to right bottom) should be zero.

$$D_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

* $\text{Trace}(D) = \sum_{i=1}^{n} a_{ij}$
  * $\text{Trace}(D_3) = a_{11} + a_{22} + a_{33}$
  $= -1 + 5 + 0$
  $= 4$

ii. Identity (Unit) matrix

* I_n is a diagonal matrix where;
  * $a_{ij} = 1$ for $i = j$
  * $a_{ij} = 0$ for $i \neq j$
  * $\text{Trace}(I_n) = n$

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Properties

* $I^n = I$
* $AI = IA = A$
* $AA^{-1} = A^{-1}A = I$

iii. Upper triangular matrix

* This is a square matrix where $a_{ij} = 0$ for $i > j$ if $a_{ij} \in U_n$

$$\bullet\; U_3 = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 5 & 6 \\ 0 & 0 & 7 \end{bmatrix} \leftarrow \text{upper triangle} \qquad \left.\begin{matrix} a_{21} \\ a_{31} \\ a_{32} \end{matrix}\right\} = 0$$

iv. Lower triangular matrix.

* This is a square matrix where $a_{ij} = 0$ for $i < j$ if $a_{ij} \in L_n$.

$$\bullet\; L_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 0 & 2 & 5 & 0 \\ 5 & 0 & 6 & 0 \end{bmatrix} \qquad \left.\begin{matrix} a_{12} \\ a_{13} \\ a_{14} \\ a_{23} \\ a_{24} \\ a_{34} \end{matrix}\right\} = 0$$

lower triangle $\rightarrow$

v. Symmetric matrix

* This is a square matrix where $a_{ij} = a_{ji}$ for $i \neq j$ if $a_{ij} \in A_n$.

$$\bullet\; A_3 = \begin{bmatrix} -1 & 5 & -3 \\ 5 & \frac{1}{2} & 0 \\ -3 & 0 & 8 \end{bmatrix}$$

vi. Skew symmetric matrix

* This is a square matrix where $a_{ij} = -a_{ij}$ for $i \neq j$ and $a_{ij} = 0$ for $i = j$ if $a_{ij} \in A_n$

$$A_3 = \begin{bmatrix} 0 & -2 & \frac{1}{3} \\ 2 & 0 & 0 \\ -\frac{1}{3} & 0 & 0 \end{bmatrix}$$

* Any square matrix can be written as an addition of a symmetric matrix and a skew symmetric matrix.
    - for any square matrix A, $A = S + K$
    - S is the symmetric part: $S = \frac{1}{2}(A + A^T)$
    - K is the skew-symmetric part: $K = \frac{1}{2}(A - A^T)$

Matrix-Vector Multiplication

* When multiplying a matrix A by a vector x to get Ax, there are two equivalent ways to view this operation.
    i. Row picture (Dot products)
    ii. Column picture (Linear combination)

i. Row picture (Dot products)

* Each component of the result comes from the dot product of row A with vector x.

* For each row i:
    - (row i of A) $\cdot x$ = component i of result

* This view emphasizes how each equation in a system contributes to the solution.

Column picture (Linear combination)

* Ax represents a linear combination of the columns of A.

* The components of x serve as weights for combining the columns.
  - $Ax = x_1(\text{column } 1) + x_2(\text{column } 2) + \dots + x_n(\text{column } n)$

* This view reveals the geometric structure of the solution space.

Matrix Vectors and Dot Product

* Both rows and columns of a matrix can be treated as vectors. The interpretation depends on context.

* The same matrix can be viewed in two different ways.
  - Consider the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$
  - Row vectors are $[1\ 2\ 3]$ and $[4\ 5\ 6]$ (two 3D vectors).
  - Column vectors are $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$ (three 2D vectors).

* Matrix multiplication can be viewed as a systematic way of computing many dot products and arranging the results in a matrix.
  - Each element C_{ij} = (row i of A) $\cdot$ (column j of B)
  - We temporarily adopt "rows and columns as vectors" view for multiplication.

* The context determines the interpretation.
  - Linear transformations - Columns as basis vectors
  - Data analysis - Rows as data points, columns as features
  - Matrix multiplication - Rows of first matrix, columns of second matrix
  - System of equations - Rows as coefficient vectors

Matrix Multiplication

Essential rules and properties

Dimension compatibility

* For AB to exist: number of columns in A = number of rows in B
* If A is $(m \times r)$ and B is $(r \times n)$, then AB is $(m \times n)$.

Entry calculation

* Entry (i,j) of AB = (row i of A) $\cdot$ (column j of B)
* This is the fundamental definition from which all other methods derive.

Column perspective

* Column j of AB = $A \times$ (column j of B)
* Each column of the result is a linear transformation of the corresponding column in B.

Non-commutative

* Generally, $AB \neq BA$.
* Order matters significantly in matrix operations.

Associativity

* $(AB)C = A(BC)$
* Parantheses can be moved without changing the result.
* This property is crucial for efficient computation of matrix chains.

Matrix factorization insight

* If A has r independent columns in C, then $A = CR$.
* where C is $(m \times r)$ and R is $(r \times n)$.
* This connects matrix multiplication to fundamental concepts like rank and linear independence.

Four methods of matrix multiplication

i. Element-by-element (Standard definition)

* Each entry C_{ij} = (row i of A) · (column j of B)
* Requires computing individual dot products for each position.

ii. Column-by-column processing

* A × (column k of B) = column k of C
* Process each column of B independently through matrix A.
* Useful for understanding how transformations affect individual vectors.

iii. Row-by-row processing

* (Row i of A) × B = row i of C
* Process each row of A independently with entire matrix B.
* Efficient for certain computational implementations.

iv. Sum of rank-1 matrices (Outer products)

* $AB = \Sigma$ (column j of A) (row j of B)
* Each term creates a rank-1 matrix; sum them all.
* Reveals the fundamental structure of matrix multiplication.

Geometric interpretation

* Matrix multiplication represents composition of linear transformations:
  - Each matrix represents a geometric transformation (rotation, scaling, shearing, etc.)
  - AB applies transformation B first, then transformation A.
  - The column picture shows how the transformation affects the coordinate system.
  - The row picture shows how each output coordinate depends on all input coordinates.

Given $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 & 7 & 8 \\ 7 & 8 & 9 & 10 \end{bmatrix}$, find $C = AB$ using each method.

Method I

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times2} \times \begin{bmatrix} 5 & 6 & 7 & 8 \\ 7 & 8 & 9 & 10 \end{bmatrix}_{2\times4}$$

$$C = \begin{bmatrix} (1\times5)+(2\times7) & (1\times6)+(2\times8) & (1\times7)+(2\times9) & (1\times8)+(2\times10) \\ (3\times5)+(4\times7) & (3\times6)+(4\times8) & (3\times7)+(4\times9) & (3\times8)+(4\times10) \\ (5\times5)+(6\times7) & (5\times6)+(6\times8) & (5\times7)+(6\times9) & (5\times8)+(6\times10) \end{bmatrix}$$

$$C = \begin{bmatrix} 19 & 22 & 25 & 28 \\ 43 & 50 & 57 & 64 \\ 67 & 78 & 89 & 100 \end{bmatrix}$$

Method II

$$\text{Column } 1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} (1\times5)+(2\times7) \\ (3\times5)+(4\times7) \\ (5\times5)+(6\times7) \end{bmatrix} = \begin{bmatrix} 19 \\ 43 \\ 67 \end{bmatrix}$$

$$\text{Column } 2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} (1\times6)+(2\times8) \\ (3\times6)+(4\times8) \\ (5\times6)+(6\times8) \end{bmatrix} = \begin{bmatrix} 22 \\ 50 \\ 78 \end{bmatrix}$$

$$\text{Column } 3 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 9 \end{bmatrix} = \begin{bmatrix} (1\times7)+(2\times9) \\ (3\times7)+(4\times9) \\ (5\times7)+(6\times9) \end{bmatrix} = \begin{bmatrix} 25 \\ 57 \\ 89 \end{bmatrix}$$

$$\text{Column } 4 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 8 \\ 10 \end{bmatrix} = \begin{bmatrix} (1\times8)+(2\times10) \\ (3\times8)+(4\times10) \\ (5\times8)+(6\times10) \end{bmatrix} = \begin{bmatrix} 28 \\ 64 \\ 100 \end{bmatrix}$$

$$C = \begin{bmatrix} 19 & 22 & 25 & 28 \\ 43 & 50 & 57 & 64 \\ 67 & 78 & 89 & 100 \end{bmatrix}$$

Method III

$$\text{Row } 1 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 & 8 \\ 7 & 8 & 9 & 10 \end{bmatrix} = \begin{bmatrix} (1\times5)+(2\times7) & (1\times6)+(2\times8) & (1\times7)+(2\times9) & (1\times8)+(2\times10) \end{bmatrix}$$

$$= \begin{bmatrix} 19 & 22 & 25 & 28 \end{bmatrix}$$

$$\text{Row } 2 = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 & 8 \\ 7 & 8 & 9 & 10 \end{bmatrix} \quad \begin{bmatrix} (3\times5)+(4\times7) & (3\times6)+(4\times8) & (3\times7)+(4\times9) & (3\times8)+(4\times10) \end{bmatrix}$$

$$\begin{bmatrix} 43 & 50 & 57 & 64 \end{bmatrix}$$

$$\text{Row } 3 = \begin{bmatrix} 5 & 6 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 & 8 \\ 7 & 8 & 9 & 10 \end{bmatrix} \quad \begin{bmatrix} (5\times5)+(6\times7) & (5\times6)+(6\times8) & (5\times7)+(6\times9) & (5\times8)+(6\times10) \end{bmatrix}$$

$$\begin{bmatrix} 67 & 78 & 89 & 100 \end{bmatrix}$$

$$C = \begin{bmatrix} 19 & 22 & 25 & 28 \\ 43 & 50 & 57 & 64 \\ 67 & 78 & 89 & 100 \end{bmatrix}$$

Method IV

$$\text{First outer product} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 7 & 8 \\ 15 & 18 & 21 & 24 \\ 25 & 30 & 35 & 40 \end{bmatrix}$$

$$\text{Second outer product} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \begin{bmatrix} 7 & 8 & 9 & 10 \end{bmatrix} = \begin{bmatrix} 14 & 16 & 18 & 20 \\ 28 & 32 & 36 & 40 \\ 42 & 48 & 54 & 60 \end{bmatrix}$$

$$C = \begin{bmatrix} 5 & 6 & 7 & 8 \\ 15 & 18 & 21 & 24 \\ 25 & 30 & 35 & 40 \end{bmatrix} + \begin{bmatrix} 14 & 16 & 18 & 20 \\ 28 & 32 & 36 & 40 \\ 42 & 48 & 54 & 60 \end{bmatrix} = \begin{bmatrix} 19 & 22 & 25 & 28 \\ 43 & 50 & 57 & 64 \\ 67 & 78 & 89 & 100 \end{bmatrix}$$

Gaussian Elimination

* Consider a system of linear equations:

- $$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

* This can be written in the matrix form as $Ax = b$ where;
  - A is the $m \times n$ coefficient matrix.
  - x is the $n \times 1$ solution vector.
  - b is the $m \times 1$ right-hand side vector.

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{m1} & a_{m2} \cdots a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{b}$$

* If we include a column that represents the constants to the coefficient matrix, we obtain what is called an augmented matrix.

- $$\left[\begin{array}{ccc|c} a_{11} & a_{12} \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \\ a_{m1} & a_{m2} \cdots & a_{mn} & b_n \end{array}\right]$$

* The following elementary row operations result in augmented matrices that represent equivalent linear systems.
  - Any two rows may be interchanged.
  - Each element in a row can be multiplied by a non-zero constant.
  - Any row can be replaced by the sum of that row and a multiple of another.

* To efficiently solve a system of linear equations, construct an augmented matrix and then apply the appropriate elementary row operations to obtain an augmented matrix in upper triangular form (or row echelon form). In this form, the equivalent linear system can easily be solved using back substitution. This process is called Gaussian elimination.

Ex

$$2x - 9y + 3z = -18$$
$$x - 2y - 3z = -8$$
$$-4x + 23y + 12z = 47$$

$$\left[\begin{array}{ccc|c} 2 & -9 & 3 & -18 \\ 1 & -2 & -3 & -8 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$R_1 \leftrightarrow R_2$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 2 & -9 & 3 & -18 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$R_2 \leftarrow -2R_1 + R_2$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$R_3 \leftarrow 4R_1 + R_3$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 15 & 0 & 15 \end{array}\right]$$

$R_3 \leftarrow 3R_2 + R_3$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 0 & 27 & 9 \end{array}\right]$$

Back substitution;

$$x - 2y - 3z = -8 \text{ — (1)}$$
$$-5y + 9z = -2 \text{ — (2)}$$
$$27z = 9 \text{ — (3)}$$

From (3);

$$z = \frac{1}{3}$$

$z = 1/3$ in (2);

$$-5y + 3 = -2$$
$$5y = 5$$
$$y = 1$$

$y = 1$ and $z = 1/3$ in (1);

$$x - 2 - 1 = -8$$
$$x = -5$$

∴ Solution;

$$x = -5, \; y = 1, \; z = \frac{1}{3}$$

* Dependent and inconsistent systems can be identified in an augmented coefficient matrix when the coefficients in one row are all zero. If a row of zeros has a corresponding constant of zero, then the matrix represents a dependent system. If the constant is non-zero, then the matrix represents an inconsistent system.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

$0 = 0$ Dependent

$0 \neq 2$ Inconsistent

Row Echelon form (REF) Properties

* All zero rows are at the bottom.
* The leading entry (pivot) of each non-zero row is to the right of the leading entry of the row above.
* All entries below a leading entry are zero.

Reduced Row Echelon Form (RREF) Properties

* All the properties of REF.
* Each leading entry is 1.
* Each leading entry is the only non-zero entry in its column.

Rank of a Matrix

* The rank of a matrix is the maximum number of linearly independent rows or columns in the matrix.

* It is equal to the number of non-zero rows in the row echelon form of the matrix.

* Rank is always equal or less than the smaller of the number of rows or columns.

* Rank of matrix A can be expressed as, rank (A).

Key properties

* Rank(A) = Row rank = Column rank
* $0 \leq \text{Rank}(A) \leq \min(m, n)$ for an m, n matrix
* Rank is invariant under elementary row operations.
* The rank of a zero matrix or a null matrix is zero.

Linear independence and dependence using rank

* The general rule for linear independence is;
  - Rows are linearly independent if and only if Rank(A) = number of rows.
  - Columns are linearly independent if and only if Rank(A) = number of columns.

Case 1: Square matrices $(m \times m)$

Rank	Row Independence	Column Independence	Matrix Type
Rank = m	Independent	Independent	Full rank, invertible
Rank < m	Dependent	Dependent	Rank deficient, singular

Case 2: Tall matrices ($m \times n$ where $m > n$)

Rank	Row Independence	Column Independence	Notes
Rank = n	Dependent	Independent	Full column rank, more rows than columns
Rank < n	Dependent	Dependent	Both sets are dependent.

Case 3: Wide matrices ($m \times n$ where $m<n$)

Rank	Row Independence	Column Independence	Notes
Rank $= m$	Independent	Dependent	Full row rank, more columns than rows
Rank $< m$	Dependent	Dependent	Both sets are dependent

Properties of rank

* Rank $(A) = $ Rank (A^T)
* Rank $(AB) \leq \min$ [Rank(A), Rank (B)]
* Rank $(A+B) \leq$ Rank (A) + Rank (B)
* If A is invertible,
  - Rank (AB) = Rank (B)
  - Rank (BA) = Rank (B)

Rank and solutions

* For a system $Ax = b$ where A is $m \times n$ (m equations, n unknowns):
  - Rank (A) = Rank$([A|b]) = n$
  - Unique solutions (when system is consistent and we have as many independent equations as unknowns)
  - Rank(A) = Rank$([A|b]) < n$
  - Infinite solutions (fewer independent equations than unknowns, creating free variables)
  - Rank $(A) <$ Rank$([A|b])$
  - No solution (Inconsistent system)

Gauss - Jordan Elimination

Gaussian vs. Gauss - Jordan elimination

Gaussian Elimination	Gauss - Jordan Elimination
* Transforms the matrix to row echelon form.	* Transforms the matrix to reduced row echelon form.
* Uses forward elimination only.	* Uses both forward and backward elimination.
* Requires back substitution to find the solution.	* Solution can be read directly from the final matrix. (no back substitution needed)

* Gauss-Jordan elimination extends Gaussian elimination by continuing the elimination process upward to create zeros above the pivots as well.

Complete algorithm

* Forward elimination by creating zeros below each pivot.
* Normalize pivots by making each leading entry equal to 1.
* Backward elimination by creating zeros above each pivot.

When to use each method

Gaussian elimination

* Solving a single system $Ax = b$
* You need the LU decomposition.
* Computational efficiency is important.

Gauss - Jordan elimination

* You want the solution without back substitution
* Finding the inverse of a matrix: $[A|I] \rightarrow [I|A^{-1}]$
* Working with multiple right-hand sides simultaneously.

Ex From previous example, REF; $\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 0 & 27 & 9 \end{array}\right]$

$R_2 \leftarrow -\frac{1}{5}R_2$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & 1 & -9/5 & 2/5 \\ 0 & 0 & 27 & 9 \end{array}\right]$$

$R_3 \leftarrow \frac{1}{27}R_3$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & 1 & -9/5 & 2/5 \\ 0 & 0 & 1 & 1/3 \end{array}\right]$$

$R_2 \leftarrow \frac{9}{5}R_3 + R_2$

$$\left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1/3 \end{array}\right]$$

$R_1 \leftarrow 3R_3 + R_1$

$$\left[\begin{array}{ccc|c} 1 & -2 & 0 & -7 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1/3 \end{array}\right]$$

$R_1 \leftarrow 2R_2 + R_1$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1/3 \end{array}\right]$$

$\therefore$ Solution;

$x = -5,\ y = 1,\ z = \frac{1}{3}$

Ex Given $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$, find A^{-1}.

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right]$$

$R_2 \leftarrow -2R_1 + R_2$

$R_3 \leftarrow -4R_1 + R_3$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right]$$

$R_2 \leftrightarrow R_3$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & -1 & -1 & -2 & 1 & 0 \end{array}\right]$$

$R_3 \leftarrow R_2 + R_3$

$R_3 \leftarrow -R_3$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

$R_1 \leftarrow -2R_3 + R_1$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

$\therefore A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$

Elementary Matrices

* An elementary matrix is always a square matrix.

* Any elementary matrix, which we often denote by E, is obtained from applying one row operation to the identity matrix of the same size.

* Elementary matrix can be constructed from any row operation, but only one operation can be applied.

Definition

* Let E be an $n \times n$ matrix. Then E is an elementary matrix if it is the result of applying one row operation to the $n \times n$ identity matrix I_n. Those which involve switching rows of the identity matrix are called permutation matrices.

* To perform any of the three row operations on a matrix A, it suffices to take the product EA, where E is the elementary matrix obtained by using the desired row operation on the identity matrix.

Types of elementary matrices

Type I (Row interchange)

* Let P_{ij} denote the elementary matrix which involves switching the i^{th} and the j^{th} rows. Then P_{ij} is a permutation matrix and;

$$P_{ij} A = B$$

where B is obtained from A by switching the i^{th} and the j^{th} rows.

Ex Let $P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} a & b \\ g & d \\ e & f \end{bmatrix}$. Find B where $B = P_{12} A$.

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ g & d \\ e & f \end{bmatrix} = \begin{bmatrix} g & d \\ a & b \\ e & f \end{bmatrix}$$

* The matrix P_{12} is obtained by switching the first and second rows of the 3×3 identity matrix I. B is the matrix obtained by switching rows 1 and 2 of A. Therefore, by multiplying A by P_{12}, the row operation which was applied to I to obtain P_{12} is applied to A to obtain B.

Type II (Row scaling)

* Let $E_{(k,i)}$ denote the elementary matrix corresponding to the row operation in which the i^{th} row is multiplied by the nonzero scalar, k. Then;

$$E_{(k,i)}A = B$$

where B is obtained from A by multiplying the i^{th} row of A by k.

Ex Let $E_{(5,2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Find the matrix B where $B = E_{(5,2)}A$.

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ 5c & 5d \\ e & f \end{bmatrix}$$

* The matrix $E_{(5,2)}$ is obtained by multiplying the second row of the identity matrix by 5. B is obtained by multiplying the second row of A by the scalar 5.

Type III (Row addition)

* Let $E_{(k\times i+j)}$ denote the elementary matrix obtained from I by adding k times the i^{th} row to the j^{th}. Then;

$$E_{(k\times i+j)}A = B$$

where B is obtained from A by adding k times the i^{th} row to the j^{th} row of A.

Ex Let $E_{(2\times 1+3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Find B where $B = E_{(2\times 1+3)}A$.

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \\ 2a+e & 2b+f \end{bmatrix}$$

* The matrix $E_{(2\times 1+3)}$ was obtained by adding 2 times the first row of I to the third row of I. B is the matrix obtained by adding 2 times the first row of A to the third row.

* In an elimination context, E_{ij} means that it is an elimination matrix that zeros out the (i,j) entry of a target matrix.

Ex

$$2x - 9y + 3z = -18$$
$$x - 2y - 3z = -8$$
$$-4x + 23y + 12z = 47$$

$$[A|b] = \left[\begin{array}{ccc|c} 2 & -9 & 3 & -18 \\ 1 & -2 & -3 & -8 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$P_{12} \times [A|b]$ $(R_1 \leftrightarrow R_2) \Rightarrow$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 2 & -9 & 3 & -18 \\ 1 & -2 & -3 & -8 \\ -4 & 23 & 12 & 47 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 2 & -9 & 3 & -18 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$E_{21} \times P_{12}[A|b]$ $(R_2 \leftarrow -2R_1 + R_2) \Rightarrow$

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 2 & -9 & 3 & -18 \\ -4 & 23 & 12 & 47 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ -4 & 23 & 12 & 47 \end{array}\right]$$

$E_{31} \times E_{21} P_{12}[A|b]$ $(R_3 \leftarrow 4R_1 + R_3) \Rightarrow$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ -4 & 23 & 12 & 47 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 15 & 0 & 15 \end{array}\right]$$

$E_{32} \times E_{31} E_{21} P_{12}[A|b]$ $(R_3 \leftarrow 3R_2 + R_3) \Rightarrow$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 15 & 0 & 15 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -2 & -3 & -8 \\ 0 & -5 & 9 & -2 \\ 0 & 0 & 27 & 9 \end{array}\right]$$

$$x - 2y - 3z = -8$$
$$-5y + 9z = -2$$
$$27z = 9$$

$\therefore$ Solution;

$$x = -5, \quad y = 1, \quad z = \frac{1}{3}$$

Inverses of elementary matrices

* Every elementary matrix is invertible and its inverse is also an elementary matrix.

* The inverse of an elementary matrix is constructed by doing the reverse row operation on I. E^{-1} will be obtained by performing the row operation which would carry E back to I.
    - If E is obtained by switching rows i and j, the E^{-1} is also obtained by switching rows i and j. Therefore, for permutation matrices, $E = E^{-1}$.
    - If E is obtained by multiplying row i by scalar k, then E^{-1} is obtained by multiplying row i by the scalar $1/k$.
    - If E is obtained by adding k times row i to row j, then E^{-1} is obtained by substracting k times row i from row j.

* Let A be an $m \times n$ matrix and let B be the reduced row echelon form of A. Then we can write $B = EA$ where E is the product of all elementary matrices representing the row operations done to A to obtain B.

Ex Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \end{bmatrix}$. Find B, the reduced row echelon form of A and write it in the form $B = EA$.

$P_{12} \times A \Rightarrow$

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

$E_{(-2\times1+3)} \times P_{12}A \Rightarrow$

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

$\therefore B = E_{(-2\times1+3)}(P_{12}A)$

$B = [E_{(-2\times1+3)}P_{12}]A$

$B = EA$

$E = E_{(-2\times1+3)}P_{12}$

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

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

Inverses of elementary matrices

* Every elementary matrix is invertible and its inverse is also an elementary matrix.

* The inverse of an elementary matrix is constructed by doing the reverse row operation on I. E^{-1} will be obtained by performing the row operation which would carry E back to I.
    - If E is obtained by switching rows i and j, the E^{-1} is also obtained by switching rows i and j. Therefore, for permutation matrices, $E = E^{-1}$.
    - If E is obtained by multiplying row i by scalar k, then E^{-1} is obtained by multiplying row i by the scalar $1/k$.
    - If E is obtained by adding k times row i to row j, then E^{-1} is obtained by substracting k times row i from row j.

* Let A be an $m \times n$ matrix and let B be the reduced row echelon form of A. Then we can write $B = EA$ where E is the product of all elementary matrices representing the row operations done to A to obtain B.

$\boxed{\text{Ex}}$ Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \end{bmatrix}$. Find B, the reduced row echelon form of A and write it in the form $B = EA$.

$P_{12} \times A \Rightarrow$

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

$E_{(-2\times1+3)} \times P_{12}A \Rightarrow$

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

$\therefore B = E_{(-2\times1+3)}(P_{12}A)$

$B = [E_{(-2\times1+3)}P_{12}]A$

$B = EA$

$E = E_{(-2\times1+3)}P_{12}$

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

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

* While the process used in the above example is reliable and simple when only a few row operations are used, it becomes cumbersome in a case where many row operations are needed to carry A to B. The following method provides an alternate way to find the matrix E.

* Let A be an $m \times n$ matrix and let B be its reduced row-echelon form. Then $B = EA$ where E is an invertible $m \times m$ matrix found by forming the matrix $[A|I_m]$ and row reducing to $[B|E]$.

Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \end{bmatrix}$. Using the above method, find E such that $B = EA$.

$$[A|I_3] = \left[\begin{array}{cc|ccc} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$P_{12} \times [A|I_3] \Rightarrow$

$$\left[\begin{array}{cc|ccc} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$E_{(-2\times1+3)} \times P[A|I_3] \Rightarrow$

$$\left[\begin{array}{cc|ccc} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 \end{array}\right]$$

* The left side of this matrix is B and the right side is E. Comparing this to the matrix E found above example, it can be seen that the same matrix is obtained regardless of which process is used.

* An $n \times n$ matrix A is invertible if and only if A can be carried to the $n \times n$ identity matrix using the usual row operations. This leads to an important consequence related to the above discussion.

* Suppose that A is an $n \times n$ invertible matrix. Then set up the matrix $[A|I_n]$ as done above, and row reduce until it is of the form $[B|E]$. In this case, $B = I_n$ because A is invertible.

  - $B = EA$

    $I_n = EA$

    $A = E^{-1}$

* Now suppose that $E = E_1 E_2 \dots E_k$ where each E_i is an elementary matrix representing a row operation used to carry A to I. Then;

  - $E^{-1} = (E_1 E_2 \dots E_k)^{-1}$

    $E^{-1} = E_k^{-1} \dots E_2^{-1} E_1^{-1}$

* If E_i is an elementary matrix, so too is E_i^{-1}. It follows that;

  - $A = E^{-1}$

    $A = E_k^{-1} \dots E_2^{-1} E_1^{-1}$

  and A can be written as a product of elementary matrices.

* There fore, let A be an $n \times n$ matrix. Then A is invertible if and only if it can be written as a product of elemetary matrices.

Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$. Write A as a product of elementary matrices.

$$[A|I] = \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{array}\right]$$

$E_1 \times [A|I]$ where $E_1 = P_{12} \Rightarrow$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 & 1 \end{array}\right] ; \quad E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$E_2 \times E_1 [A|I]$ where $E_2 = E_{(2 \times 2 + 3)} \Rightarrow$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \end{array}\right] ; \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$E_3 \times E_2 E_1 [A|I]$ where $E_3 = E_{(-1 \times 2 + 1)} \Rightarrow$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \end{array}\right] ; \quad E_3 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

* Since the reduced row echelon form of A is I, $I = EA$ where E is the product of the above elementary matrices. It follows that $A = E^{-1}$. Since we want to write A as a product of elementary matrices, we wish to express E^{-1} as a product of elementary matrices.

$$\begin{aligned} A &= E^{-1} \\ &= (E_3E_2E_1)^{-1} \\ &= E_1^{-1}E_2^{-1}E_3^{-1} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Transpose Matrices

* The transpose of a matrix is obtained by interchanging its rows and columns. If A is an $m \times n$ matrix, then its transpose A^T is an $n \times m$ matrix.

* The transpose of a matrix A is denoted as A^T.

* Let $A_{2\times3} = \begin{bmatrix} -1 & 5 & 0 \\ 6 & 1 & 7 \end{bmatrix}$. Then $A^T_{3\times2} = \begin{bmatrix} -1 & 6 \\ 5 & 1 \\ 0 & 7 \end{bmatrix}$.

* The element at position (i,j) in matrix A becomes the element at position (j,i) in A^T.

Properties

* $(A^T)^T = A$
* $(kA)^T = kA^T$
* $(A+B)^T = A^T + B^T$
* $(AB)^T = B^TA^T$

* Let A be an invertible matrix, then the transpose of the matrix, A^T, is also invertible and;
    * $(A^T)^{-1} = (A^{-1})^T$

* This means you can interchange the order of inverse and transpose operations.
    - $(A^{-1})^T A^T = I$
    - $A^T (A^{-1})^T = I$

* We can verify this by showing that $(A^{-1})^T$ behaves as the inverse of A^T.
    - $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$
    - $A^T (A^{-1})^T = (A^{-1}A)^T = I^T = I$

Determinant

* The determinant is a scalar value that can be computed from the elements of a square matrix. It provides crucial information about the matrix and its properties, particularly regarding;
    - Whether the matrix has an inverse.
    - Whether a system of linear equations has a unique solution.
    - The linear independence of the matrix's columns / rows.

* For a matrix A, the determinant is written as $\det(A)$ or $|A|$.

Determinant of a 2x2 matrix.

* For a 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant is calculated as;
    - $\det(A) = ad - bc$
    - $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Important properties

* Swapping two rows or columns changes the sign of the determinant.
* If $\det(A) = 0$, the matrix has no inverse and its columns are linearly dependent.

Determinant of a 3x3 matrix

* For a 3x3 matrix, we use the Rule of Sarrus or cofactor expansion.

Rule of Sarrus (Diagonal Method)

* Consider a matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

* Write the first two columns again to the right.
    - $\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$

* Calculate the positive diagonals (↘).
    - $\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$
    - $aei + bfg + cdh$

* Calculate the negative diagonals (↙).
    - $\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$
    - $ceg + afh + bdi$

* $\det(A) = (aei + bfg + cdh) - (ceg + afh + bdi)$

  $= aei + bfg + cdh - ceg - afh - bdi$

Minors

* The minor M_{ij} of element a_{ij} is the determinant of the matrix obtained by deleting row i and column j.

* For a 3×3 matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$;
  - Minor of element 'a' (M_{11}) is $\det \begin{bmatrix} e & f \\ h & i \end{bmatrix}$.
  - Minor of element 'b' (M_{12}) is $\det \begin{bmatrix} a & f \\ g & i \end{bmatrix}$

Visual representation of M_{ij}

$$\begin{array}{cc} & j^{th}\text{ column} \\ M_{ij} = \det & \begin{bmatrix} a_{11} & \cdots & \cdot & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \cdot & \cdots & a_{ij} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \cdot & \cdots & a_{nn} \end{bmatrix} & i^{th}\text{ row} \end{array}$$

Cofactor

* The cofactor C_{ij} is the signed minor.
  - $C_{ij} = (-1)^{i+j} \times M_{ij}$

* The sign pattern for a cofactor matrix is $\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$

$\boxed{\text{Ex}}$ Determine the cofactor of 5 in $\begin{bmatrix} 3 & 5 & 7 \\ -1 & 2 & 3 \\ -4 & 4 & -9 \end{bmatrix}$.

$$C_{12} = (-1)^3 \times [(-1 \times -9) - (3 \times -4)]$$

$$= -1 \times 21$$

$$= \underline{\underline{-21}}$$

Cofactor Expansion (Laplace Expansion)

* Consider a matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

* The determinant can be calculated by expanding along any row or column.
  - Along row i: $\det(A) = \sum_{j=1}^{n} a_{ij} \times C_{ij}$
  - Along column j: $\det(A) = \sum_{i=1}^{n} a_{ij} \times C_{ij}$

* Choose the row or column with the most zeros to minimize calculations.

Ex Compute the determinant of $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

Using 1st row,

$\det(A) = 1(-2) + 5(0) + 0(-4) = -2$

Using 2nd row,

$\det(A) = 2(0) + 4(0) + (-1)(2) = -2$

Using 3rd row,

$\det(A) = 0(-5) + (-2)(1) + 0(-6) = -2$

Using 1st column,

$\det(A) = 1(-2) + 2(0) + 0(-5) = -2$

Using 2nd column,

$\det(A) = 5(0) + 4(0) + (-2)(1) = -2$

Using 3rd column,

$\det(A) = 0(-4) + (-1)(2) + 0(-6) = -2$

Ex Find det (A) where $A = \begin{bmatrix} -1 & 5 & -2 \\ -6 & 6 & 0 \\ 3 & -7 & 1 \end{bmatrix}$.

$$\det(A) = -6(9) + 6(5)$$
$$= -24$$

Inverse Matrices

* For a square matrix A, we look for an 'inverse matrix' A^{-1} of the same size, such that A^{-1} times A equals I (the identity matrix). Whatever A does, A^{-1} undoes.

* The matrix A is invertible if there exists a matrix A^{-1} that 'inverts' A.
  - $A^{-1}A = AA^{-1} = I$

* The product $A^{-1}Ax = x$, meaning A^{-1} undoes what A does to any vector.

* Not all matrices have inverses. The columns must be linearly independent.

* A matrix cannot have two different inverses. The inverse is unique when it exists.

* For invertible matrices, the unique solution to $Ax = b$ is $x = A^{-1}b$.

* The matrix A is invertible if and only if;
  - The determinant is non-zero.
  - All columns are linearly independent.
  - The matrix has full rank.
  - Gaussian elimination process produces n pivots. (for nxn matrix)

* For a 2x2 matrix, there's a direct formula;
  - $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$
  - $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Properties of Inverse of Matrix Products

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

* The inverses appear in reverse order. For example, in $(AB)^{-1}$, we're finding what undoes the combined action of A then B. To undo 'first A, then B', we must do 'first undo B, then undo A'. Hence, $(AB)^{-1} = B^{-1}A^{-1}$.

* A matrix is non-invertible (singular) if we cannot convert the augmented matrix $[A|I]$ into $[I|A^{-1}]$ through row operations.

* If row reduction produces at least one row of zeros, the matrix is non-invertible.

$$\left[\begin{array}{cccc|cccc} 1 & -2 & 3 & 5 & 1 & 0 & 0 & 0 \\ 2 & 5 & 6 & 9 & 0 & 1 & 0 & 0 \\ -3 & 1 & 2 & 3 & 0 & 0 & 1 & 0 \\ 1 & 13 & -30 & -49 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\Downarrow \text{To RREF}$$

$$\left[\begin{array}{cccc|cccc} 1 & -2 & 3 & 5 & 1 & 0 & 0 & 0 \\ 0 & 9 & 0 & -1 & -2 & 1 & 0 & 0 \\ 0 & -5 & 11 & 18 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 3 & 1 \end{array}\right]$$

- This matrix is singular because its reduced row echelon form contains a zero row.

Finding Matrix Inverses : The Cofactor Method

* Verify that $\det(A) \neq 0$. If the determinant is zero, the matrix has no inverse.

* For each entry a_{ij} in matrix A, calculate the cofactor C_{ij} and get the cofactor matrix C consisting of all these cofactors.

* Find the adjoint (adjugate) where the adjoint of A is the transpose of the cofactor matrix.
  - $\text{adj}(A) = C^T$

* Apply the inverse formula.

$$\boxed{A^{-1} = \frac{1}{\det(A)} \times \text{adj}(A) = \frac{C^T}{\det(A)}}$$

Ex Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & 5 \\ -2 & 3 & 7 \\ 6 & -1 & 0 \end{bmatrix}$.

$$C = \begin{bmatrix} +7 & -(-42) & +(-16) \\ -5 & +(-30) & -(-1) \\ +(-15) & -17 & +3 \end{bmatrix} = \begin{bmatrix} 7 & 42 & -16 \\ -5 & -30 & 1 \\ -15 & -17 & 3 \end{bmatrix}$$

$$\text{adj}(A)\ (C^T) = \begin{bmatrix} 7 & -5 & -15 \\ 42 & -30 & -17 \\ -16 & 1 & 3 \end{bmatrix}$$

$$\det(A) = 1(7) + 5(-16) = -73$$

$$\therefore A^{-1} = \frac{1}{-73} \begin{bmatrix} 7 & -5 & -15 \\ 42 & -30 & -17 \\ -16 & 1 & 3 \end{bmatrix}$$

Properties of Determinants

Property 1: Identity Matrix

* The determinant of an $n \times n$ identity matrix is 1.
* $\det(I) = 1$

Property 2: Row Exchange

* Exchanging (swapping) two rows of matrix A reverses the sign of the determinant.
* If A' is A with two rows swapped, then $\det(A') = -\det(A)$.

Property 3: Linearity in Rows

* The determinant is a linear function of each row separately.
* If row i of A is $cv + dw$ where A_1 has v in row i, A_2 has w in row i, and all other rows are identical;
  - $\det(A) = c.\det(A_1) + d.\det(A_2)$
  - $\det \begin{bmatrix} cv+dw \\ \text{row } 2 \\ \vdots \\ \text{row } n \end{bmatrix} = c.\det \begin{bmatrix} v \\ \text{row } 2 \\ \vdots \\ \text{row } n \end{bmatrix} + d.\det \begin{bmatrix} w \\ \text{row } 2 \\ \vdots \\ \text{row } n \end{bmatrix}$
* The determinant is linear in each row when other rows are held constant.

Property 4: Equal Rows

* If two rows of A are equal, then $\det(A) = 0$.
* This follows from property 2: swapping equal rows gives $\det(A) = -\det(A)$, so $\det(A) = 0$.

Property 5: Row Operations

* Subtracting/adding a multiple of one row from another leaves $\det(A)$ unchanged.
* This is crucial for Gaussian elimination while preserving determinant relationships

Property 6: Zero Row

* A matrix with a row of zeros has $\det(A) = 0$.
* If we multiply the zero row by any scalar t, the matrix is unchanged, so $\det(A) = t \cdot \det(A)$, which is only true for all t if $\det(A) = 0$.

Property 7: Triangular Matrices

* If A is triangular (upper or lower), then $\det(A)$ = product of diagonal entries.
* $\det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn}$
* This can be proven using elimination to make all non-diagonal entries 0, then applying linearity n times.

Property 8: Singular vs Invertible

* If A is singular (non-invertible), then $\det(A) = 0$.
* If A is invertible, then $\det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$.
* Connection to eigenvalues: $\det(A - \lambda I) = 0$ gives the characteristic equation for eigenvalues.

Property 9: Product Rule

* The determinant of a product equals the product of determinants.
* $\det(AB) = \det(A) \cdot \det(B)$
* This is a remarkable property that greatly simplifies calculations.

Property 10: Transpose

* The transpose has the same determinant as the original matrix.
* $\det(A^T) = \det(A)$
* This means all row properties also apply to columns.

Cramer's Rule

* Cramer's rule provides a method to solve the linear system $Ax = b$ using determinants, provided $\det(A) \neq 0$.

* For the system $Ax = b$ where $\det(A) \neq 0$;

$$x_j = \frac{\det(B_j)}{\det(A)}$$

Where B_j is the matrix A with its j^{th} column replaced by the vector b.

Step-by-step process

i. Calculate $\det(A)$. If $\det(A) = 0$, the system has no unique solution.

ii. For each unknown x_j;
   - Create matrix B_j by replacing column j of A with vector b.
   - Calculate $\det(B_j)$.
   - Set $x_j = \dfrac{\det(B_j)}{\det(A)}$

* For a 3x3 system;
   - $x_1 = \dfrac{\det(B_1)}{\det(A)}$ where $B_1 = [b \ a_2 \ a_3]$
   - $x_2 = \dfrac{\det(B_2)}{\det(A)}$ where $B_2 = [a_1 \ b \ a_3]$
   - $x_3 = \dfrac{\det(B_3)}{\det(A)}$ where $B_3 = [a_1 \ a_2 \ b]$

Practical notes

* Cramer's rule works better in theory than in computational practice.
* More efficient for small systems or theoretical proofs.
* Useful for understanding the relationship between determinants and linear systems.
* For large systems, Gaussian elimination is typically more efficient.

Ex Use Cramer's rule to solve for the vector x.

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

$\det(A) = 2(4) + 1(2) = 10$

$$B_1 = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 2 & -4 & 4 \end{bmatrix} \qquad \therefore x_1 = \frac{\det(B_1)}{\det(A)} = \frac{8}{10} = 0.8$$

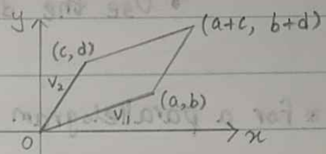
$$B_2 = \begin{bmatrix} -1 & 1 & -3 \\ 2 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix} \qquad \therefore x_2 = \frac{\det(B_2)}{\det(A)} = \frac{-20+5}{10} = -1.5$$

$$B_3 = \begin{bmatrix} -1 & 2 & 1 \\ 2 & 0 & 0 \\ 3 & -4 & 2 \end{bmatrix} \qquad \therefore x_3 = \frac{\det(B_3)}{\det(A)} = \frac{2(-8)}{10} = -1.6$$

Geometric Interpretation : Areas and Volumes

2D: Area of Parallelogram

* For a parallelogram with vertices at $(0,0)$, $\underbrace{(a,b)}_{v_1}$, $\underbrace{(c,d)}_{v_2}$ and $(a+c,\ b+d)$;
  - Area = $\left|\det\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right| = |ad-bc|$
    $\begin{array}{cc} \uparrow & \uparrow \\ v_1 & v_2 \end{array}$



3D: Volume of Parallelepiped

* For a 'box' with edges e_1, e_2, e_3 starting from origin;
  - Volume = $|\det(E)| = |\det[e_1\ e_2\ e_3]|$ (e_1, e_2, e_3 are column vectors)

General n- Dimensional Rule

* For an n-dimensional 'box' with edges $e_1, e_2, \dots, e_n$, Volume = $\det(E)$ where E has the edge vectors as columns. This generalizes the concept of area (2D) and volume (3D) to higher dimensions.

* The absolute value of the determinant gives the n-dimensional volume of the parallelepiped formed by the column (or row) vectors of the matrix.

* When forming the matrix, edge vectors can either be columns or rows because $\det(A) = \det(A^T)$. But the column interpretation is more common.

* The parallelogram or parallelepiped doesn't need to start from the origin to use determinants for areas and volumes. However, the standard formulas using determinants do assume that one vertex is at the origin.
  - If the parallelogram doesn't start from the origin, you can translate it so that one vertex is at the origin, then apply the formula.
  - For parallelepipeds, similarly, if it doesn't start from the origin, translate it first.

* Translation doesn't change areas or volumes. So, if you have a parallelogram with vertices at points A, B, C, D, you can;
    - Translate everything so that vertex A is at the origin.
    - The other vertices become B-A, C-A, D-A.
    - Two adjacent sides from A become vectors (B-A) and (C-A).
    - Use the determinant formula with these vectors.

* For a parallelogram with vertices at (1,2), (4,3), (3,6), (6,7);
    - Translate so (1,2) is at origin.
    - Adjacent sides become vectors (3,1) and (2,4).
    - Area $= \left|\det\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\right| = 10.$

Eigenvalues and Eigenvectors

* When a square matrix A acts upon a vector x, it typically produces a new vector Ax that is both stretched and rotated. However, there are special vectors that remain in the same direction when transformed by A. These are called eigenvectors.

* Consider the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.
    - When applied to vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$; $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
    - When applied to vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
    - When applied to vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$; $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

* The matrix A leaves the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ unchanged. In particular, it does not rotate the vector. When a matrix A acts upon a vector and does not rotate it, we have $Ax = \lambda x$ where λ is a scaling factor. In our example, $\lambda = 1$. We call such a vector an eigenvector for the matrix A, and the associated scaling factor λ an eigenvalue.

Definition

* Let A be an $n \times n$ matrix. If there exists a real value λ and a non-zero $n \times 1$ vector x satisfying;

  - $Ax = \lambda x$

Then we refer to λ as an eigenvalue of A, and x as an eigenvector of A corresponding to λ.

Matrix Powers

$Ax = \lambda x$

$A^2 x = \lambda A x$

$A^2 x = \lambda^2 x$

In general;

$A^n x = \lambda^n x$

* The eigenvectors of A^n are the same as the eigenvectors of A; while the eigenvalues of A^n are the eigenvalues of A raised to the n^{th} power.

Finding Eigenvalues

* From the equation $Ax = \lambda x$, we can rearrange to;

  - $Ax - \lambda x = 0$

    $(A - \lambda I)x = 0$

* For this system to have a non-trivial solution, the matrix $(A - \lambda I)$ must be singular.

  - $\det(A - \lambda I) = 0$

* So, if λ is an eigenvalue of A, then $\det(A - \lambda I) = 0$. It turns out that this equation can be used to calculate every eigenvalue. The equation $\det(A - \lambda I) = 0$ will be a polynomial equation in λ, and its roots will give us all the eigenvalues. We call this polynomial equation the characteristic equation of A.

* In triangular matrices, eigenvalues are on the diagonal.

* When A is singular, $\lambda = 0$ is one of the eigenvalues.

Finding Eigenvectors

* To find the eigenvectors, for each eigenvalue λ, solve the system;
  - $(A - \lambda I)x = 0$

* An $n \times n$ matrix has n eigenvalues (repeated λs are possible). Each λ leads to x.

Common cases when solving for eigenvectors

Case 1: One free variable

$$\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$2x - 4y = 0$
$x - 2y = 0$ $\Longrightarrow$ $x = 2y$ (rows are dependent) $\therefore$ x is dependent on y.
$\therefore$ y is a free variable.

$\therefore$ Solution; $x = \begin{bmatrix} 2y \\ y \end{bmatrix}$ where $y \neq 0$

Case 2: Multiple free variables (Higher dimensional eiganspace)

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Only constraint; $x + 2z = 0$
$x = -2z$

* Since there are no other constraints on y or z, they are considered free variables. This means the system has an infinite number of solutions, and the eiganspace has a dimension of 2. We can find two linearly independent eigenvectors by choosing simple non-zero values for our free variables.

The general solution vector; $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2z \\ y \\ z \end{bmatrix}$

* To find a set of basis vectors for this eiganspace, we can make two separate choices for the free variables. While any two values can be chose that yield linearly independent vectors, a simple and reliable method is to set one free variable to '1' and the other to '0' for each basis vector you need.

For the eigenvector v_1, let $y=1$ and $z=0$.

$$v_1 = \begin{bmatrix} -2(0) \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

For the eigenvector v_2, let $y=0$ and $z=1$.

$$v_2 = \begin{bmatrix} -2(1) \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

The eigenvectors for this system are any non-zero linear combination of these two linearly independent vectors. The eigenspace is spanned by the basis vectors v_1 and v_2.

Case 3: Chain substitution

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$x + y + z = 0$ $\qquad$ $y + 2z = 0$

$x - 2z + z = 0$ $\qquad$ $y = -2z$

$x = z$

$\therefore$ Solution; $x = \begin{bmatrix} z \\ -2z \\ z \end{bmatrix}$ where $z \neq 0$

Case 5: Zero eigenvalue (Singular matrix)

* This is the case of solving $(A - 0I)x = Ax = 0$ when $\lambda = 0$.

* Find the null space of A directly. All solutions form the eigenspace for $\lambda = 0$.

Tips for eigenvector calculations.

* Always check your answer: verify that $Ax = \lambda x$.
* Choose the simplest form (often integers) for final eigenvectors.
* When multiple free variables exist, find a basis for the eigenspace.
* Set free variables to convenient values (like 1) to get clean answers.

* Eigenvectors are not unique. Any non-zero multiple cx $(c \neq 0)$ of an eigenvector is also an eigenvector.

* The zero vector is never an eigenvector (by definition, eigenvectors must be non-zero).

* Zero can be a eigenvalue.

* We can multiply eigenvectors by any non-zero constant. $A(cx) = \lambda(cx)$ is still true.

* MATLAB and most software produce eigenvectors of length $\|x\| = 1$.

* We cannot use elimination to calculate eigenvalues. If we use elimination to convert a matrix A into an upper triangular matrix U, the eigenvalues of U could be different than the eigenvalues of A. However, the two will not be completely unrelated.

* What relates them is the fact that the determinant of a matrix is equal to the product of its eigenvalues. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A;
    - $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$

* Since A and U have the same determinant, the product of their eigenvalues will be the same.

* The trace of a matrix is defined as being the sum of the diagonal elements.
  - $\text{trace}(A) = a_{11} + a_{22} + \dots + a_{nn}$

* The trace of a matrix will be equal to the sum of the eigenvalues of the matrix.
  - $\text{trace}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n$

Matrix Diagonalization

* A nxn matrix A can be diagonlized if and only if it has n linearly independent eigenvectors.

* If all eigenvalues are distinct (no repeated eigenvalues), then the matrix can always be diagonalized.

* If eigenvalues are repeated, check if there are enough independent eigenvectors.

Diagonalization process

* Form the invertible eigenvector matrix $P = [x_1 \; x_2 \; \dots \; x_n]$ by putting the n linearly indepedent eigenvectors into the columns.

* Form the diagonal eigenvalue matrix $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

* Then, $A = P \Lambda P^{-1}$ or equivalently $\Lambda = P^{-1} A P$.

* The eigenvalues in Λ must appear in the same order as their corresponding eigevectors in P or vice versa.

* Diagonalization of a matrix refers to the process of transforming any matrix A into its diagonal form Λ. Then;

  - $\Lambda = P^{-1}AP$

where Λ is a diagonal matrix and P is a modal matrix.

* The matrix P has an inverse because its columns (the eigenvectors of A) were assumed to be linearly independent.

* A and Λ have the same eigenvalues but the eigenvectors are different.

Benefits of Diagonalization

* Easy matrix powers
  - $A^k = P\Lambda^k P^{-1}$
* Calculations are simplified because powers of diagonal matrices are trivial to compute since the power of a diagonal matrix just exponentiate the diagonal elements while keeping the other elements zero.
* Diagonalization is useful in understanding the long-term behaviour of linear systems.

* There is no connection between invertibility and diagonalizability.

* Repeated eigenvalues may prevent diagonalization if there aren't enough independent eigenvectors.

Checking whether a matrix is diagonalizable when there are repeated eigenvalues

* For a matrix to be diagonalizable, its set of eigenvectors must form a basis for the vector space. For an nxn matrix, this means there must be n linearly independent eigenvectors.

* The key concepts are;
  - Algebraic Multiplicity (AM) - The number of times an eigenvalue appears as a root of the characteristic polynomial.
  - Geometric Multiplicity (GM) - The dimension of the eigenspace corresponding to that eigenvalue. This is the number of linearly independent eigenvectors associated with that eigenvalue.

* A matrix is diagonalizable if and only if for every eigenvalue, its algebraic multiplicity equals its geometric multiplicity. That is, for every eigenvalue λ;
  - $AM(\lambda) = GM(\lambda)$

* Let's consider a case of a 3×3 matrix with 2 distinct eigenvalues, say λ_1 (repeated) and λ_2.

* The characteristic polynomial is a cubic polynomial. The sum of algebraic multiplicities must be 3.
  - $AM(\lambda_1) = 2$ and $AM(\lambda_2) = 1$

* For the matrix to be diagonalizable, we must have;
  - $AM(\lambda_1) = GM(\lambda_1) = 2$ (means there are 2 linearly independent eigenvectors for λ_1)
  - $AM(\lambda_2) = GM(\lambda_2) = 1$ (means there is one eigenvector for λ_2)

* Since for any eigenvalue, $1 \leq GM(\lambda) \leq AM(\lambda)$, the condition $GM(\lambda_2) = 1$ is always satisfied because $AM(\lambda_2) = 1$. The real question is whether $GM(\lambda_1)$ can be equal to 2. If it is, then the matrix is diagonalizable. If $GM(\lambda) = 1$, then the matrix is not diagonalizable.

Example of a diagonalizable matrix.

Consider the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

* The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$ where $AM(\lambda_1) = 2$ and $AM(\lambda_2) = 1$.
(For diagonal matrices, the eigenvalues are simply the diagonal entries)

* The eigenspace for $\lambda_1 = 1$ is spanned by the two vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.
So, $GM(\lambda_1) = 2$.

* The eigenspace for $\lambda_2 = 2$ is spanned by the vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. So, $GM(\lambda_2) = 1$.

* Since $AM(\lambda_1) = GM(\lambda_1)$ and $AM(\lambda_2) = GM(\lambda_2)$, the matrix is diagonalizable.

Example of a non-diagonalizable matrix

Consider the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

* The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$ where $AM(\lambda_1) = 2$ and $AM(\lambda_2) = 1$.

* For $\lambda_1 = 1$, we solve $(A - 1I)x = 0$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

* This gives $y = 0$ and $z = 0$, and x is a free variable. The eigenspace is spanned by $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. So, $GM(\lambda_1) = 1$.

* Since $AM(\lambda_1) = 2 \neq GM(\lambda_1) = 1$, the matrix is not diagonlizable.

Date: __/__/__

Ex Let $A = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

a) Determine the eigenvalues of matrix A.

$$A - \lambda I = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 6-\lambda & 2 & 0 \\ 2 & 3-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{bmatrix}$$

$$\begin{aligned} \det(A-\lambda I) &= (-1-\lambda)\left[(6-\lambda)(3-\lambda) - 4\right] \\ &= (-1-\lambda)(14 - 9\lambda + \lambda^2) \\ &= -14 - 5\lambda + 8\lambda^2 - \lambda^3 \\ &= -(\lambda^3 - 8\lambda^2 + 5\lambda + 14) \\ &= -\left[(\lambda+1)(\lambda^2 - 9\lambda + 14)\right] \\ &= -(\lambda+1)(\lambda-2)(\lambda-7) \end{aligned}$$

To find the eigenvalues;

$$\det(A - \lambda I) = 0$$

$$-(\lambda+1)(\lambda-2)(\lambda-7) = 0$$

$\underline{\underline{\lambda = -1}}$, $\underline{\underline{\lambda = 2}}$, $\underline{\underline{\lambda = 7}}$

b) Find the linearly independent eigenvectors corresponding to each eigenvalue obtained in part (a).

For $\lambda = -1$;

$$(A - \lambda I)x = 0 \qquad\qquad 7x + 2y = 0$$

$$(A + I)x = 0 \qquad\qquad 2x + 4y = 0$$

$$\begin{bmatrix} 7 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad x, y = 0 \qquad z \text{ is a free variable.}$$

$$x = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = z\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \; ; \; z \in \mathbb{R}$$

For $\lambda = 2$;

$$(A-\lambda I)x = 0 \qquad 4x+2y=0 \qquad -3z=0$$

$$(A-2I)x = 0 \qquad 2x+y=0 \qquad z=0$$

$$\begin{bmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad y = -2x$$

$$x = \begin{bmatrix} x \\ -2x \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} ; \ x \in \mathbb{R}$$

For $\lambda = 7$;

$$(A-\lambda I)x = 0 \qquad -x+2y=0 \qquad -8z=0$$

$$(A-7I)x = 0 \qquad x = 2y \qquad z=0$$

$$\begin{bmatrix} -1 & 2 & 0 \\ 2 & -4 & 0 \\ 0 & 0 & -8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x = \begin{bmatrix} 2y \\ y \\ 0 \end{bmatrix} = y\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} ; \ y \in \mathbb{R}$$

The eigenvectors take the following forms;

$$z\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ x\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \ y\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} ; \ x, y, z \in \mathbb{R}$$

c) If the matrix is diagonalizable, determine the invertibe matrix P and diagonal matrix D such that $P^{-1}AP = D$.

By substituting $x=y=z=1$, we can obtain the following eigenvectors.

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

The eigenvectors we obtained are linearly independent. Hence, we can construct the P matrix as follows.

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

Note that the first column corresponds to the eigenvalue $\lambda = -1$, the second column corresponds to $\lambda = 2$, and the third column corresponds to $\lambda = 7$. We must follow the same order when constructing the diagonalized matrix D.

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

d) If $P^{-1} = \begin{bmatrix} 2/5 & 1/5 & 0 \\ 1/5 & -2/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, determine A^2.

To obtain the P corresponding to the given P^{-1}, use $P = (P^{-1})^{-1}$ property.

$$C_{P^{-1}} = \begin{bmatrix} +(-2/5) & -(1/5) & +(0) \\ -(1/5) & +(2/5) & -(0) \\ +(0) & -(0) & +(-1/5) \end{bmatrix} = \begin{bmatrix} -2/5 & -1/5 & 0 \\ -1/5 & 2/5 & 0 \\ 0 & 0 & -1/5 \end{bmatrix}$$

$$\det(P^{-1}) = 1 \times -1/5 = -1/5$$

$$(P^{-1})^{-1} = \frac{1}{\det(P^{-1})} \cdot C_{P^{-1}}^T$$

$$P = -5 \begin{bmatrix} -2/5 & -1/5 & 0 \\ -1/5 & 2/5 & 0 \\ 0 & 0 & -1/5 \end{bmatrix}$$

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

Observe the columns of P. Using the corresponding eigenvalues, obtain the matrix D.

$$D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

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

$$A^2 = P D^2 P^{-1}$$

$$A^2 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 49 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 5 \end{bmatrix} \frac{1}{5}$$

$$A^2 = \begin{bmatrix} 40 & 18 & 0 \\ 18 & 13 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e) Calculate the determinant det(A) and the trace tr(A) of matrix A using the eigenvalues.

$$\det(A) = \text{Product of eigenvalues}$$
$$= -1 \times 2 \times 7$$
$$= -14$$

$$tr(A) = \text{Sum of eigenvalues}$$
$$= -1 + 2 + 7$$
$$= 8$$

Deriving det(A) using eigenvalues

* Since D is a triangular matrix, the determinant det(D) is the product of the diagonal values according to the properties of the determinants. Since the diagonal of D consists of eigenvalues, det(D) becomes the product of eigenvalues.

$$\det(D) = \lambda_1 \lambda_2 \dots \lambda_n$$

$$A = PDP^{-1}$$
$$\det(A) = \det(PDP^{-1})$$
$$\det(A) = \det(P) . \det(D) . \det(P^{-1})$$
$$\det(A) = \frac{1}{\det(P^{-1})} \times \det(D) \times \det(P^{-1})$$
$$\det(A) = \det(D) = \lambda_1 \lambda_2 \dots \lambda_n$$

Orthogonal and Orthonormal Vectors

* Two vectors are said to be orthogonal if they are perpendicular to each other, meaning they form a 90° angle when placed in the same coordinate system

* In two-dimensional space, if two vectors are orthogonal, they meet at a right angle. In three dimensional-space, this means they lie on mutually perpendicular planes. The idea of orthogonality extends to higher dimensions as well.

* If A and B are orthogonal;
  - $A \cdot B = 0$

  where A and B are vectors, and $\cdot$ represents the dot product between the two.

* Since $0 . x = 0$ for any vector x, the zero vector is orthogonal to every vector in $\mathbb{R}^n$.

* Consider two vectors $u = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $v = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$. Their dot product is;
  - $(3 \times -4) + (4 \times 3) = -12 + 12 = 0$

  Since the dot product is zero, u and v are orthogonal.

* Dot product (scalar product) of two n-dimensional vectors A and B, is given by
  - $A . B = \sum_{i=1}^{n} a_i b_i = A^T B$

* Since we typically represent vectors as column matrices (or row), direct multiplication is impossible due to dimensional incompatibility. To fix this, we transpose the first vector A. A transpose operation flips a matrix over its diagonal, turning a column vector into a row vector or vice versa.

* The vectors A and B are orthogonal to each other if and only if;
  - $A \cdot B = \sum_{i=1}^{n} a_i b_i = A^T B = 0$

Ex Check whether the vectors $v_1 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix}$ are orthogonal to each other.

$$v_1^T v_2 = \begin{bmatrix} 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} = 2 - 10 + 8 = 0$$

Hence, the vectors are orthogonal to each other.

* An orthogonal set of non-zero vectors is linearly independent.

Unit Vector

* Let's consider a vector A. The unit vector of the vector A can be defined as
  - $\underline{a} = \frac{A}{|A|}$; where $|A|$ is the magnitude of A.

* Unit vectors are used to define directions in a coordinate system.

* Any vectors can be written as a product of a unit vector and a scalar magnitude

Ex Consider a vector A in 2D space, $A = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Calculate the unit vector.

$$|A| = \sqrt{3^2 + 4^2} = 5$$

$$\therefore \text{unit vector, } \underline{a} = \frac{A}{|A|} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$$

Orthonormal Vectors

* A set of vectors A is orthonormal if every vector in A has magnitude 1 and the set of vectors are mutually orthogonal.

* Now, take the same 2 vectors which are orthogonal to each other and when the dot product is taken between these two vectors, it is going to be zero. So, if we also impose the condition that we want each of these vectors to have unit magnitude, we can divide the vector by its magnitude as we see in the unit vector.

* Now, we can write v_1 and v_2 as;

- $$v_1 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} \div \sqrt{1^2 + (-2)^2 + 4^2} = \begin{bmatrix} 1/\sqrt{21} \\ -2/\sqrt{21} \\ 4/\sqrt{21} \end{bmatrix}$$

- $$v_2 = \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} \div \sqrt{2^2 + 5^2 + 2^2} = \begin{bmatrix} 2/\sqrt{33} \\ 5/\sqrt{33} \\ 2/\sqrt{33} \end{bmatrix}$$

* So, what we do is we have taken the vectors from the previous example and converted them into unit vectors. by dividing them by their magnitudes. So, these vectors will still be orthogonal to each other and now individually they also have unit magnitude. Such vectors are known as orthonormal vectors.

* All orthonormal vectors are orthogonal by the definition itself.

Gram - Schmidt Process

* The Gram-Schmidt process is a method for converting a set of linearly independent vectors into an orthonormal set, meaning that the vectors are both orthogonal to each other and have unit length.

* First, we define the projection operator. The projection of a vector v onto u is given by;

- $$\text{Proj}_u v = \frac{v \cdot u}{\|u\|^2} u$$; where $v \cdot u$ = dot product of v and u
$\|u^2\|$ = the square of the magnitude of u

* The projection of vector v onto u ($\text{Proj}_u v$) is the component of v that lies in the direction of u. It's the 'piece' of v that points in the same direction as u.

Ex Given the vectors $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $u = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^2$, find the orthogonal component v_1 by subtracting the projection from the original vector v.

$$\text{Proj}_u v = \frac{(1)(1)+(1)(0)}{1^2+0^2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Subtracting the projection from v to get an orthogonal vector v_1;

$$v_1 = v - \text{Proj}_u v = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Thus, v_1 is orthogonal to u.

Gram - Schmidt Algorithm

* To orthonormalize a set of linearly independent vectors $v_1, v_2, \dots, v_n$;
  - Let $u_1 = v_1$. Then normalize it to get $e_1 = \frac{u_1}{\|u_1\|}$.
  - Let $u_2 = v_2 - \text{Proj}_{u_1} v_2$. Then normalize it to get $e_2 = \frac{u_2}{\|u_2\|}$.
  - Let $u_3 = v_3 - \text{Proj}_{u_1} v_3 - \text{Proj}_{u_2} v_3$. Then normalize it to get $e_3 = \frac{u_3}{\|u_3\|}$.

* And this process continues for $v_4, v_5, \dots$ subtracting projections onto all previous vectors to ensure orthogonality, and normalizing to get unit vectors. The result is an orthonormal set $e_1, e_2, \dots, e_n$.

Ex Continuing from the previous example, find the set of orthonormal vectors.

Normalizing the vectors to get orthonormal vectors;

$$e_1 = \frac{u}{\|u\|} = \frac{1}{1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$e_2 = \frac{v_1}{\|v_1\|} = \frac{1}{1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

final orthonormal set; $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

Singular Value Decomposition

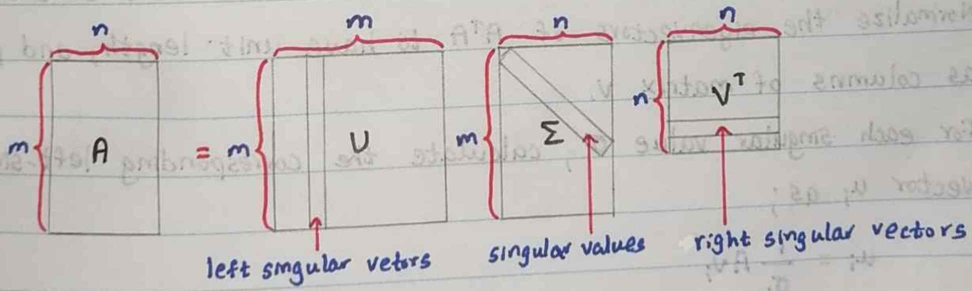
* The singular value decomposition of an $m \times n$ real matrix A is a factorization of the form;
    - $A = U \Sigma V^T$

* U is an $m \times m$ orthogonal matrix (its columns and rows are orthonormal vectors). The columns of U are called the left-singular vectors of A.

* Σ is an $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal. The diagonal entries $\sigma_i = \Sigma_{ii}$ are known as the singular values of A. and are typically arranged in descending order ($\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$). The number of the non-zero singular values is equal to the rank of A.

* V is an $n \times n$ orthogonal matrix. The columns of V are called the right-singular vectors of A.

* When constructing the matrices U and V in a singular value decomposition, the columns should correspond to the singular values in the Σ.
    - The i^{th} column of U corresponds to the i^{th} singular value in Σ.
    - The i^{th} column of V corresponds to the i^{th} singular value in Σ.



* Every matrix has a singular value decomposition.

* Singular value decomposition relation (SVD equation);
    - $Av_i = \sigma_i u_i$ for $i \leq r$; where $r = \text{rank}(A)$

$v_i = i^{th}$ column of V

$u_i = i^{th}$ column of U

* The singular value decomposition of a matrix A can be computed using the following observations.
  - The left-singular vectors of A are a set of orthonormal eigenvectors of AA^T.
  - The right-singular vectors of A are a set of orthonormal eigenvectors of A^TA.
  - The non-zero singular values of A are the square roots of the non-zero eigenvalues of both A^TA and AA^T.
  - If $U\Sigma V^T$ is the SVD of A, then for each singular value σ_i;

$$u_i = \frac{1}{\sigma_i} A v_i$$

where u_i is the i^{th} column of U and v_i is the i^{th} column of V.

* Based on the above observations, we can compute the SVD of an mxn matrix A using the following steps.
  - Construct the matrix A^TA.
  - Compute the eigenvalues and eigenvectors of A^TA. The eigenvalues will be the squares of the singular values of A, and the eigenvectors will form the columns of the matrix V in the SVD.
  - Arrange the singular values of A in descending order. Create an mxn diagonal matrix Σ with the singular values on the diagonal, padding with zeros if necessary so that the matrix has the same dimensions as A.
  - Normalize the eigenvectors of A^TA to have unit length, and place them as columns of matrix V.
  - For each singular value σ_i, calculate the corresponding left-singular vector u_i as;

$$u_i = \frac{1}{\sigma_i} A v_i$$

where v_i is the i^{th} column of V. Place these vectors as columns in the matrix U.

* Note that it is also possible to start the SVD computation by finding the left-singular vectors (the eigenvectors of AA^T) and then use the following relationship to find the right singular vectors.

  - $v_i = \frac{1}{\sigma_i} A^T u_i$

* The choice of using either AA^T or A^TA depends on which matrix is smaller.

* Another way to find the left-singular vectors of A (the columns of U) compute the eigenvectors of AA^T (If started with right-singular vectors), but this process is usually more time consuming than using the relationship between the left and right singular vectors (observation 4) and computing the null space of A^T (if necessary).

Compute the SVD of the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix}$.

Let $U\Sigma V^T$ be the SVD of A. The dimensions of A are 3×2. Therefore, the size of U is 3×3, the size of Σ is 3×2, and the size of V is 2×2.

Since the size of A^TA (2×2) is smaller than the size of AA^T (3×3), it makes sense to start with the right-singular vectors of A.

$$A^TA = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

$$A^TA - \lambda I = \begin{bmatrix} 5-\lambda & 4 \\ 4 & 5-\lambda \end{bmatrix}$$

$$\begin{aligned} \det(A^TA - \lambda I) &= (5-\lambda)^2 - 16 \\ &= [(5-\lambda)-4][(5-\lambda)+4] \\ &= (-\lambda+1)(-\lambda+9) \\ &= (\lambda-1)(\lambda-9) \end{aligned}$$

$\det(A^TA - \lambda I) = 0$

$(\lambda - 1)(\lambda - 9) = 0$

$\lambda = 1, \quad \lambda = 9$

The eigenvalues of A^TA in descending order are $\lambda_1 = 9$ and $\lambda_2 = 1$.
Therefore, the singular values of A are $\sigma_1 = 3$ and $\sigma_2 = 1$.

$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

We now find the right singular vectors (columns of V) by finding an orthonormal set of eigenvectors of A^TA.

When $\lambda = 9$;

$(A^TA - 9I)v_1 = 0$

$$\begin{bmatrix} -4 & 4 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$-4v_{11} + 4v_{12} = 0$

$\therefore v_{12} = v_{11}$

Therefore, the eigenvectors are of the form $v_1 = \begin{bmatrix} v_{11} \\ v_{11} \end{bmatrix}$. For unit length eigenvector we need to have the magnitude of the vector as 1.

$\sqrt{v_{11}^2 + v_{11}^2} = 1$

$\sqrt{2v_{11}^2} = 1$

$v_{11} = \frac{1}{\sqrt{2}}$

$$\therefore v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

When $\lambda = 1$;

$[(A^TA) - I]v_2 = 0$

$$\begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$4v_{21} + 4v_{22} = 0$

$v_{22} = -v_{21}$

Therefore, the eigenvector is of the form $v_2 = \begin{bmatrix} v_{21} \\ -v_{21} \end{bmatrix}$.

$\sqrt{v_{21}^2 + v_{21}^2} = 1$

$v_{21} = \frac{1}{\sqrt{2}}$

$$\therefore v_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

We can now write the matrix V, whose columns are the vectors v_1 and v_2 and combined in the order corresponding to the singular value in Σ.

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Lastly, we find the left-singular vectors of A using observation 4.

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{3}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{18} \\ 1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$$

Since there is only one remaining column vector of U, instead of computing the null space of A^T, we can simply find a unit vector that is perpendicular to both u_1 and u_2.

Let $u_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. To be perpendicular to u_1 and u_2;

$u_3^T \cdot u_1 = 0$ OR $u_1^T \cdot u_3 = 0$ — ①

$u_3^T \cdot u_2 = 0$ OR $u_2^T \cdot u_3 = 0$ — ②

from ②;

$$\begin{bmatrix} a & b & c \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}} = 0$$

$$a = b$$

$a = b$ in ①;

$$\begin{bmatrix} a & a & c \end{bmatrix}\begin{bmatrix} 1/\sqrt{18} \\ 1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{a}{\sqrt{18}} + \frac{a}{\sqrt{18}} + \frac{4c}{\sqrt{18}} = 0$$

$$\frac{2a}{\sqrt{18}} = -\frac{4c}{\sqrt{18}}$$

$$a = -2c$$

$$\therefore u_3 = \begin{bmatrix} -2c \\ -2c \\ c \end{bmatrix}$$

For the vector to be unit-length;

$$\sqrt{4c^2+4c^2+c^2} = 1$$

$$\sqrt{9c^2} = 1$$

$$c = \frac{1}{3}$$

$$\therefore u_3 = \begin{bmatrix} -2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

$$U = \begin{bmatrix} 1/\sqrt{18} & 1/\sqrt{2} & -2/3 \\ 1/\sqrt{18} & -1/\sqrt{2} & -2/3 \\ 4/\sqrt{18} & 0 & 1/3 \end{bmatrix}$$

$$\therefore A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{18} & 1/\sqrt{2} & -2/3 \\ 1/\sqrt{18} & -1/\sqrt{2} & -2/3 \\ 4/\sqrt{18} & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$