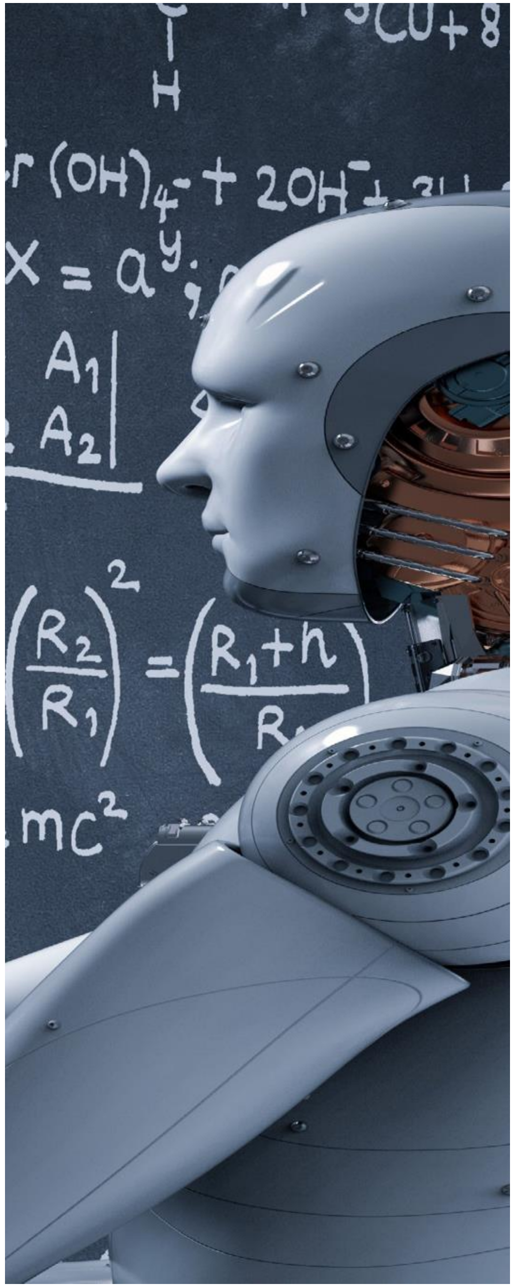
18AIE339T-MATRIX THEORY FOR ARTIFICIAL INTELLIGENCE

Reference Books

1. Xian-DaZhang, "A Matrix Algebra Approach to Artificial Intelligence" , Springer, 2021
2. Xian-DaZhang, "Matrix Analysis and Applications" ,Cambridge University Press, 2017
3. Charu C.Aggarwal, "Linear Algebra and Optimization for Machine Learning" , Springer, 2020.
4. Stephen Boyd,Lieven Vandenberghe, "Introduction to Applied Linear Algebra- Vectors, Matrices, and Least Squares" , Cambridge University Press, 2018
5. "LinearAlgebra", Kenneth Hoffman and RayKunze, Prentice Hall India,2013.
6. "LinearAlgebra", Cheney and Kincaid, Jones and Bartlett learning,2014

UNIT-1 Linear Systems

- Linear Systems – Introduction to Linear Algebra
- Linear Algebra and AI
- Examples of Linear Algebra in AI
- From Fundamental System of Solutions to Linear Space
- System of Linear Equations
- Matrices
- Solving Systems of Linear Equations
- Vector Spaces
- Linear Independence – Basis and Rank
- Linear Mapping



LINEAR ALGEBRA IN AI

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

Why linear algebra?

- Linear algebra plays a vital role in better understanding artificial intelligence and quantum computing.
- Linear algebra, widely considered as the "guru of mathematics" is a computational tool for science, engineering, and data analytics.
- Four pillars of linear algebra are vectors, matrices, tensors, and scalars
- Data is an important factor in the outcome of any event. The four pillars mentioned above will be used as repositories to store data and operate on it as per provided instructions.
- Linear algebra by large is considered as a "storage space".
- In simple terms linear algebra can be widely used as a "data guru" for machine learning and artificial intelligence (deep learning) by exposing itself to data clustering, data classification, data validation, and data fitting.
- Any time-consuming deep analysis of data requires storage, and linear algebra provides the perfect solution for this, by using vectors, matrices, tensors, and scalars as a central data repository.

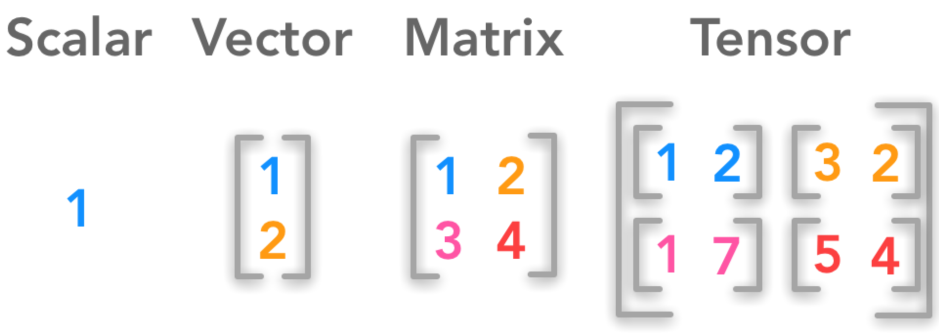
Mathematical Objects in Linear Algebra

Scalar: This is a real or natural number.

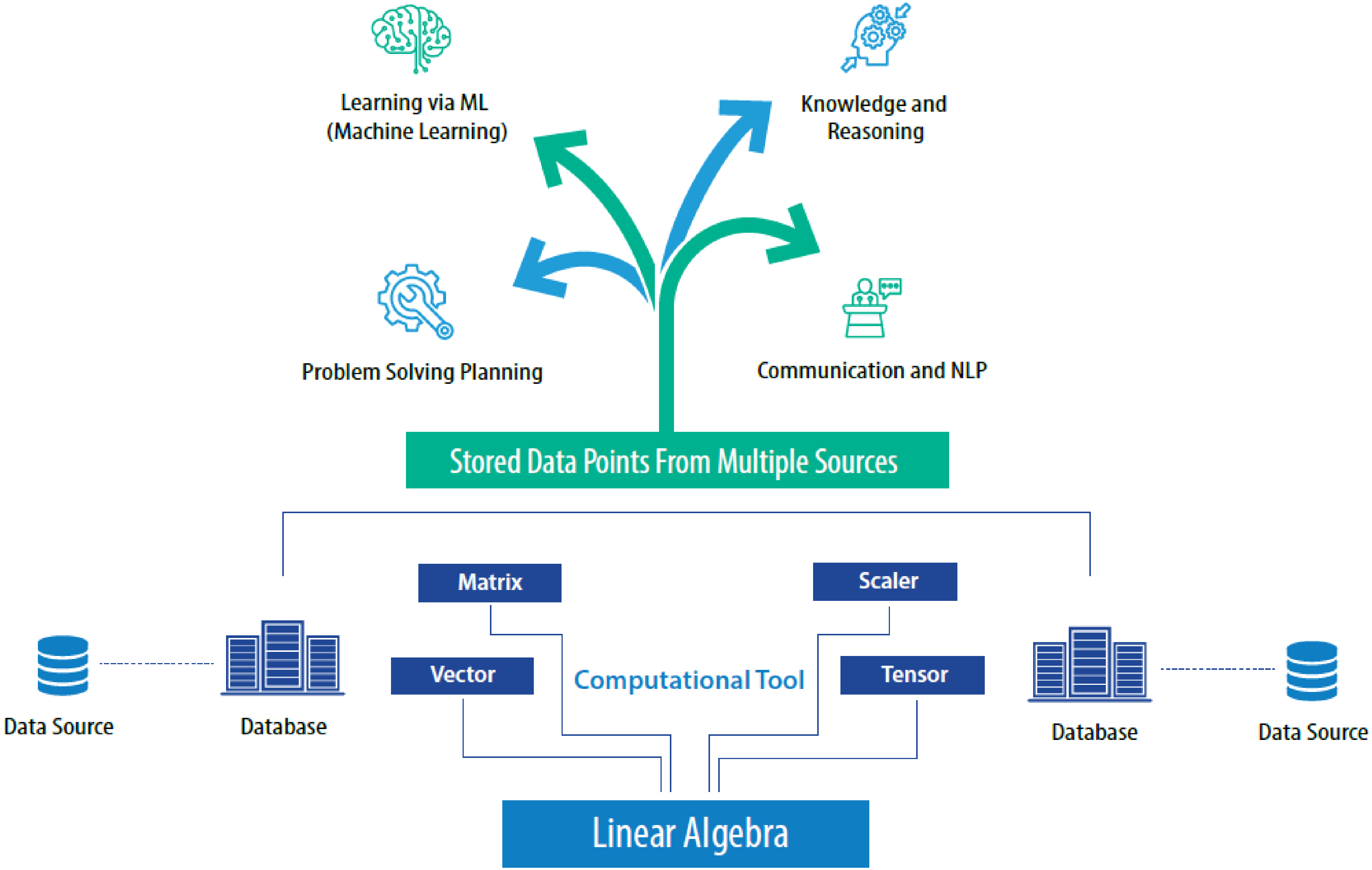
Vector: This represents an axis that requires a coordinate for each element. To identify each element, focus on an index where a list of numbers is arranged in an order. A simple vector is used to identify points in a given space. This is a typical mathematical object which is widely used in current machine learning techniques.

Matrix: A matrix is used to identify each element in a number of arrays in two dimensions; each element in a matrix is ascertained by two indices.

Tensors: Unlike a matrix which is a two-dimensional array, a tensor is an n-dimensional array. The generic nature is carried across all the other objects, such as scalars, vectors, and matrices. An array of elements are arranged on a variable axis. For instance, in a typical software coding language, a tensor is used as a variable to declare multiple data sets and extract the information from a given set of files.



Linear Algebra: A Bird's-eye View



Linear Systems – Introduction to Linear Algebra

- Machines or computers only understand numbers.

- These numbers need to be represented and processed in a way that lets machines solve problems by learning from the data instead of learning from predefined instructions (as in the case of programming).

- All types of programming use mathematics at some level. AI involves programming data to learn the function that best describes the data.

Contd..

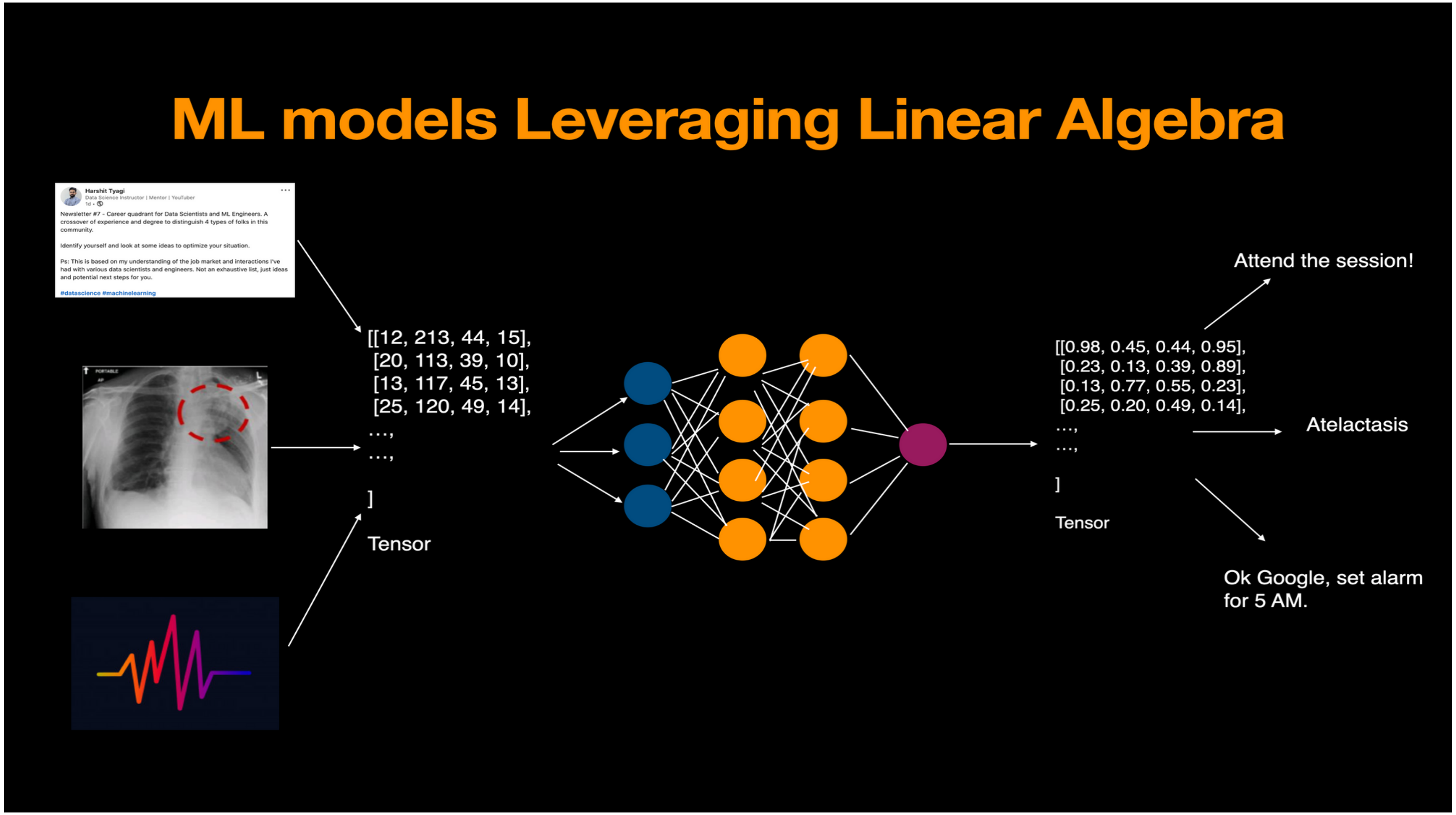
- AI is programming to optimize for the best possible solution – and we need math to understand how that problem is solved.

- The first step towards learning Math for AI is to learn linear algebra.

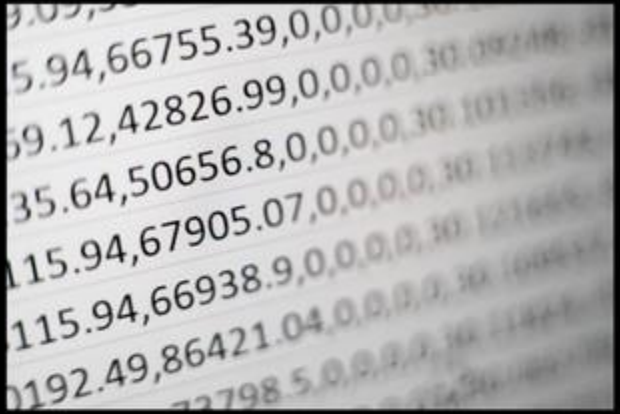
- Linear Algebra is the mathematical foundation that solves the problem of representing data as well as computations in AI models.

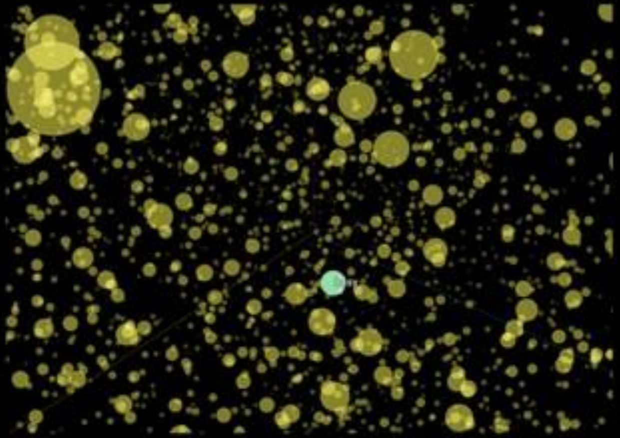
- It is the math of arrays—technically referred to as vectors, matrices and tensors.

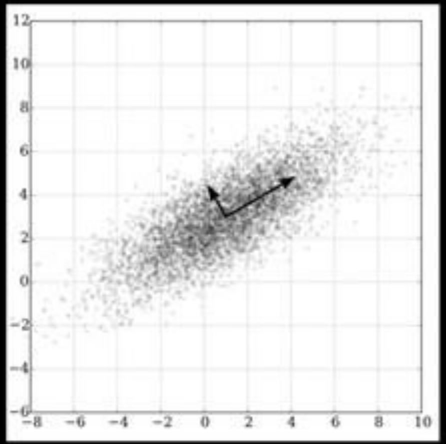
Linear Algebra and AI



Examples of Linear Algebra in AI







Data Representation

Using vectors, matrices and tensors to represent data.

Word Embeddings

Embedding vectors to efficiently represent words for natural language problems

Dimensionality Reduction

Using eigenvectors and eigenvalues to deal with large-dimensional data.

Examples of Linear Algebra in AI

Important areas of application that are enabled by linear algebra are:

- Data and learned model representation
- Word embeddings
- Dimensionality reduction

Data Representation

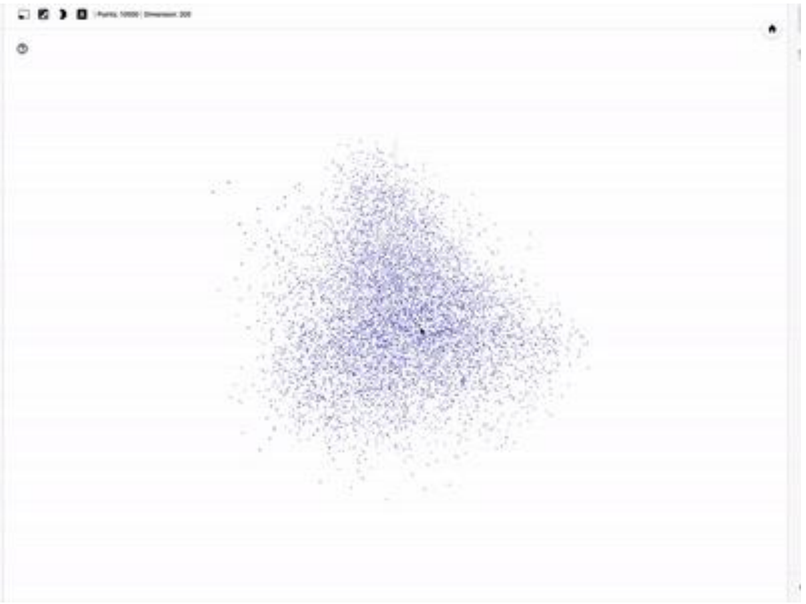
- The fuel of ML models, that is **data**, needs to be converted into arrays before you can feed it into your models.

- The computations performed on these arrays include operations like matrix multiplication (dot product).

- This further returns the output that is also represented as a transformed matrix/tensor of numbers.

Word embeddings

- It is just about representing large-dimensional data (think of a huge number of variables in your data) with a smaller dimensional vector.



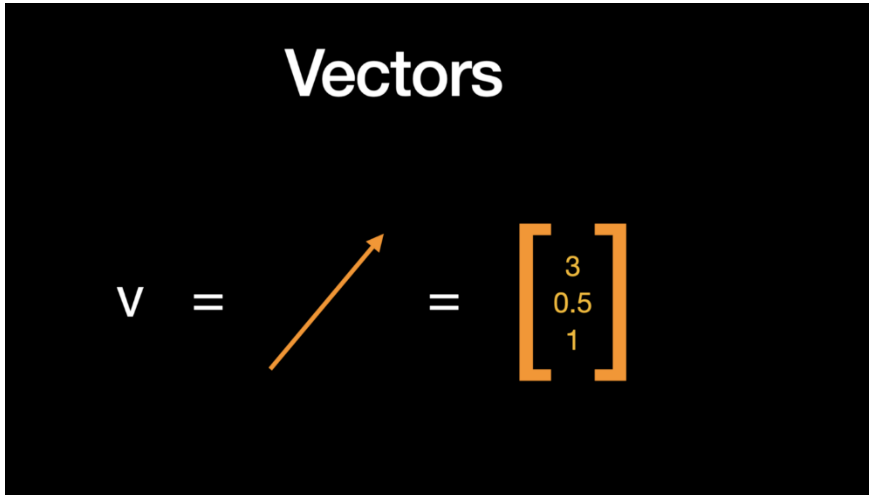
Contd..

- Natural Language Processing (NLP) deals with textual data.
- Dealing with text means comprehending the meaning of a large corpus of words.
- Each word represents a different meaning which might be similar to another word.
- Vector embeddings in linear algebra allow us to represent these words more efficiently.

Eigenvectors (SVD - Singular value decomposition)

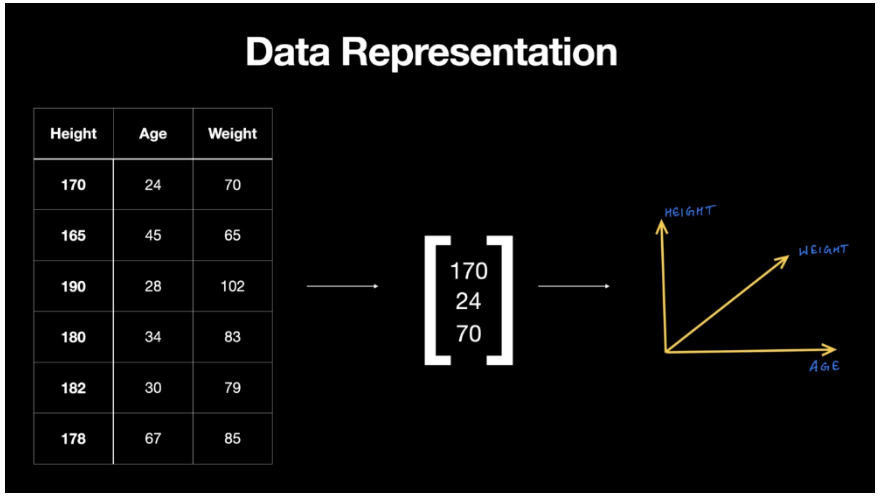
- Finally, concepts like eigenvectors allow us to reduce the number of features or dimensions of the data while keeping the essence of all of them using something called **principal component analysis.**

From Data to Vectors



- Linear algebra basically deals with vectors and matrices (different shapes of arrays) and operations on these arrays.
- In NumPy, vectors are basically a 1-dimensional array of numbers but geometrically, they have both magnitude and direction.

Data Representation



Our data can be represented using a vector. In the figure above, one row in this data is represented by a feature vector which has 3 elements or components representing 3 different dimensions. N-entries in a vector makes it n-dimensional vector space and in this case, we can see 3-dimensions.

Applications

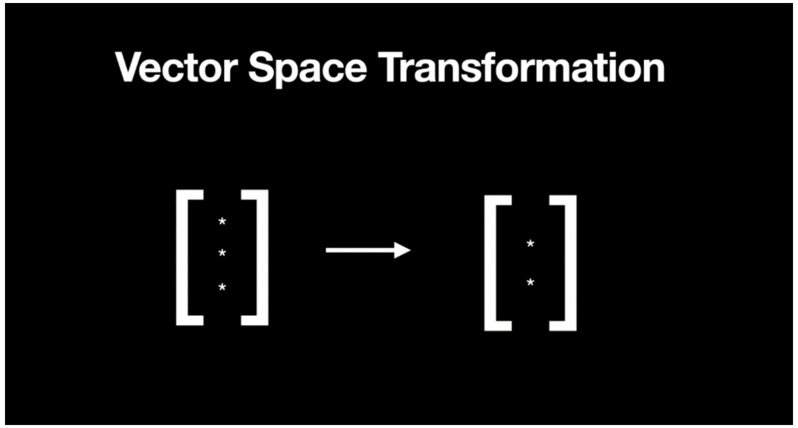
We can see linear algebra in action across all the major applications today:

- Examples include sentiment analysis on a LinkedIn or a Twitter post (embeddings), detecting a type of lung infection from X-ray images (computer vision), or any speech to text bot (NLP).

- All of these data types are represented by numbers in tensors. We run vectorized operations to learn patterns from them using a neural network. It then outputs a processed tensor which in turn is decoded to produce the final inference of the model.

- Each phase performs mathematical operations on those data arrays.

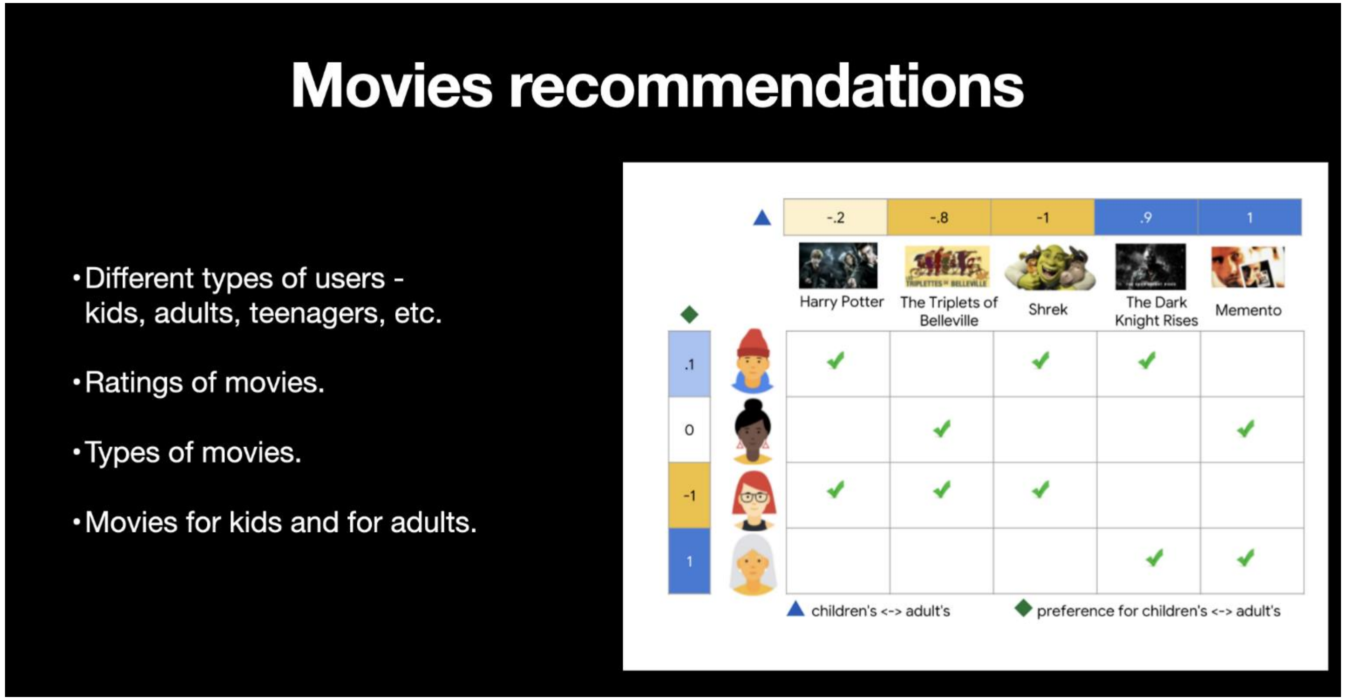
Dimensionality Reduction—Vector Space Transformation



- For example, here is a 3-dimensional vector that is replaced by a 2-dimensional space.
- Reducing dimensions doesn’t mean dropping features from the data.
- Instead, it's about finding new features that are linear functions of the original features and preserving the variance of the original features.
- Finding these new variables (features) translates to finding the principal components (PCs).
- This then converges to solving eigenvectors and eigenvalues problems.

Recommendation Engines—Making use of embeddings

- Just to give you a real-world use case to relate to all of this discussion on vector embeddings, all applications that are giving you personalized recommendations are using vector embedding in some form.



Linear System

- In AI, linear systems are particularly relevant in the context of linear regression and certain types of machine learning models, such as linear classifiers. Let's explore these two main aspects:

1. **Linear Regression**: Linear regression is a supervised learning algorithm used for predicting a continuous output variable based on input features. It assumes a linear relationship between the input features and the output variable. The linear regression model can be represented by the following equation:

$$y = b + w_1 * x_1 + w_2 * x_2 + \ldots + w_n * x_n$$

where:

1. y is the predicted output variable,
2. b is the bias term (the y-intercept of the regression line),
3. $w_1, w_2, \ldots, w_n$ are the weights (coefficients) associated with each input feature $x_1, x_2, \ldots, x_n$,
4. $x_1, x_2, \ldots, x_n$ are the input features.

The goal of linear regression is to find the best values for the weights and bias that minimize the difference between the predicted output and the actual output (given the training data).

Contd..

2.Linear Classifiers: In some machine learning tasks, the objective is to classify data into different categories or classes. Linear classifiers are algorithms that use linear equations to draw decision boundaries in the input space to separate different classes. One of the simplest linear classifiers is the linear perceptron, which can be represented by the following equation:

$$y = w_1 * x_1 + w_2 * x_2 + \ldots + w_n * x_n$$

where:

1. y is the output representing the class label (e.g., 0 or 1),
2. $w_1, w_2, \ldots, w_n$ are the weights associated with each input feature $x_1, x_2, \ldots, x_n$,
3. $x_1, x_2, \ldots, x_n$ are the input features.

The sign of the output y determines the predicted class label. For example, if $y > 0$, the sample is classified into one class, and if $y < 0$, it is classified into the other class.

Contd..

- Linear systems and linear equations play a crucial role in AI, especially in simple models like linear regression and linear classifiers.

- While these models have limitations in handling complex relationships in data, they serve as building blocks for more sophisticated models used in AI, such as neural networks, which can learn non-linear mappings and patterns in data.

Question 1

Multiply: $\begin{bmatrix} 2 & 0 & -5 \\ -1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

Possible Answers:

$\begin{bmatrix} 9 \\ 5 \end{bmatrix}$

$\begin{bmatrix} -1 \\ 8 \end{bmatrix}$

$\begin{bmatrix} 9 \\ 8 \end{bmatrix}$

$\begin{bmatrix} -1 \\ 5 \end{bmatrix}$

$\begin{bmatrix} -1 \\ -3 \end{bmatrix}$

Correct answer:

$\begin{bmatrix} -1 \\ 5 \end{bmatrix}$

Explanation:

To multiply, add:

$$2\begin{bmatrix} 2 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 3 \end{bmatrix} + 1\begin{bmatrix} -5 \\ 1 \end{bmatrix}$$

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

Question no. 2

Find the product $A \times B$

Where $A = \begin{bmatrix} 3 & 10 \\ 3 & -10 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -8 \\ 11 & -19 \end{bmatrix}$

Possible Answers:

The multiplication cannot be performed.

$\begin{bmatrix} 182 & -148 \\ -38 & 232 \end{bmatrix}$

$\begin{bmatrix} 85 & -245 \\ -135 & 135 \end{bmatrix}$

$\begin{bmatrix} 116 & -214 \\ -104 & 166 \end{bmatrix}$

Correct answer:

$\begin{bmatrix} 116 & -214 \\ -104 & 166 \end{bmatrix}$

Explanation:

In order to multiply two matrices, $A \times b$, the respective dimensions of each must be of the form $m \times n$ and $n \times p$ to create an $m \times p$ matrix. Note that order does matter:

$(A \times b \neq b \times A)$

Since A has dimensions: 2×2

and B has dimensions: 2×2

The resultant matrix has dimensions: 2×2

$$\begin{bmatrix} 3 & 10 \\ 3 & -10 \end{bmatrix} \times \begin{bmatrix} 2 & -8 \\ 11 & -19 \end{bmatrix}$$

$$\begin{bmatrix} (3)(2) + (10)(11) & (3)(-8) + (10)(-19) \\ (3)(2) + (-10)(11) & (3)(-8) + (-10)(-19) \end{bmatrix}$$

$$\begin{bmatrix} 116 & -214 \\ -104 & 166 \end{bmatrix}$$

Question no. 3

Compute AB

$A = [\,5\,]$

$B = [\,-5 \quad 8 \quad 12 \quad 7\,]$

Possible Answers:

None of the other answers

$AB = \begin{bmatrix} -25 \\ 40 \\ 60 \\ 35 \end{bmatrix}$

$AB = \begin{bmatrix} 0 \\ 13 \\ 17 \\ 12 \end{bmatrix}$

$AB = [\,-25 \quad 40 \quad 60 \quad 35\,]$

$AB = [\,0 \quad 13 \quad 17 \quad 12\,]$

Correct answer:

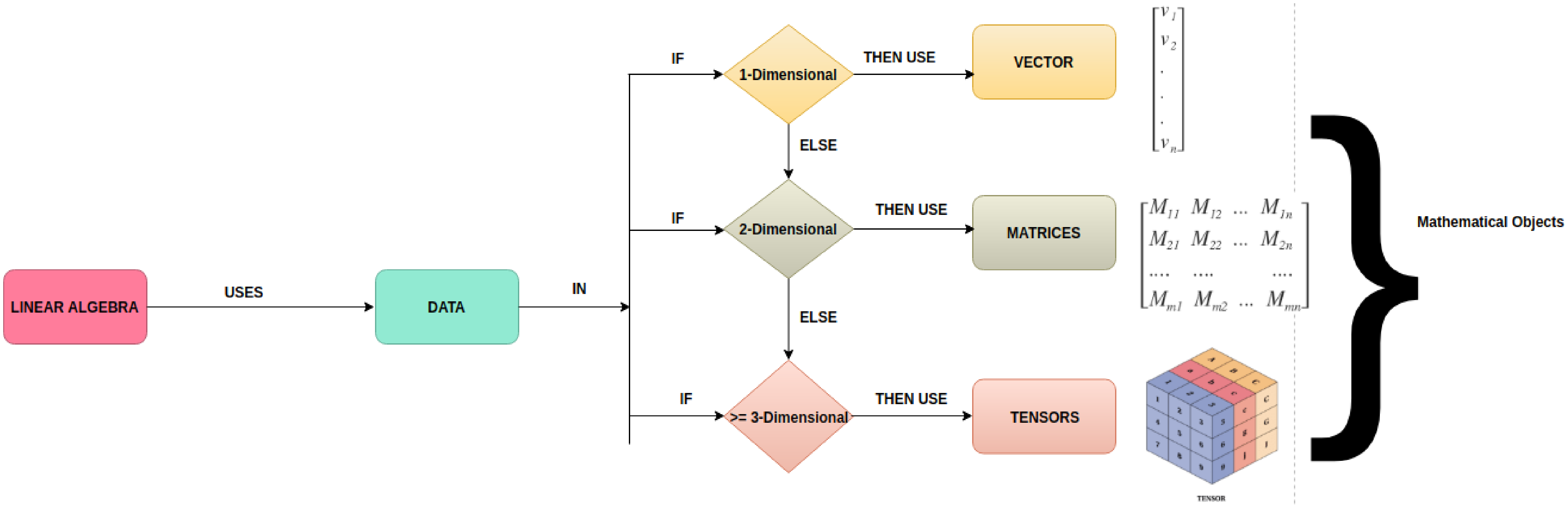
$AB = [\,-25 \quad 40 \quad 60 \quad 35\,]$

Explanation:

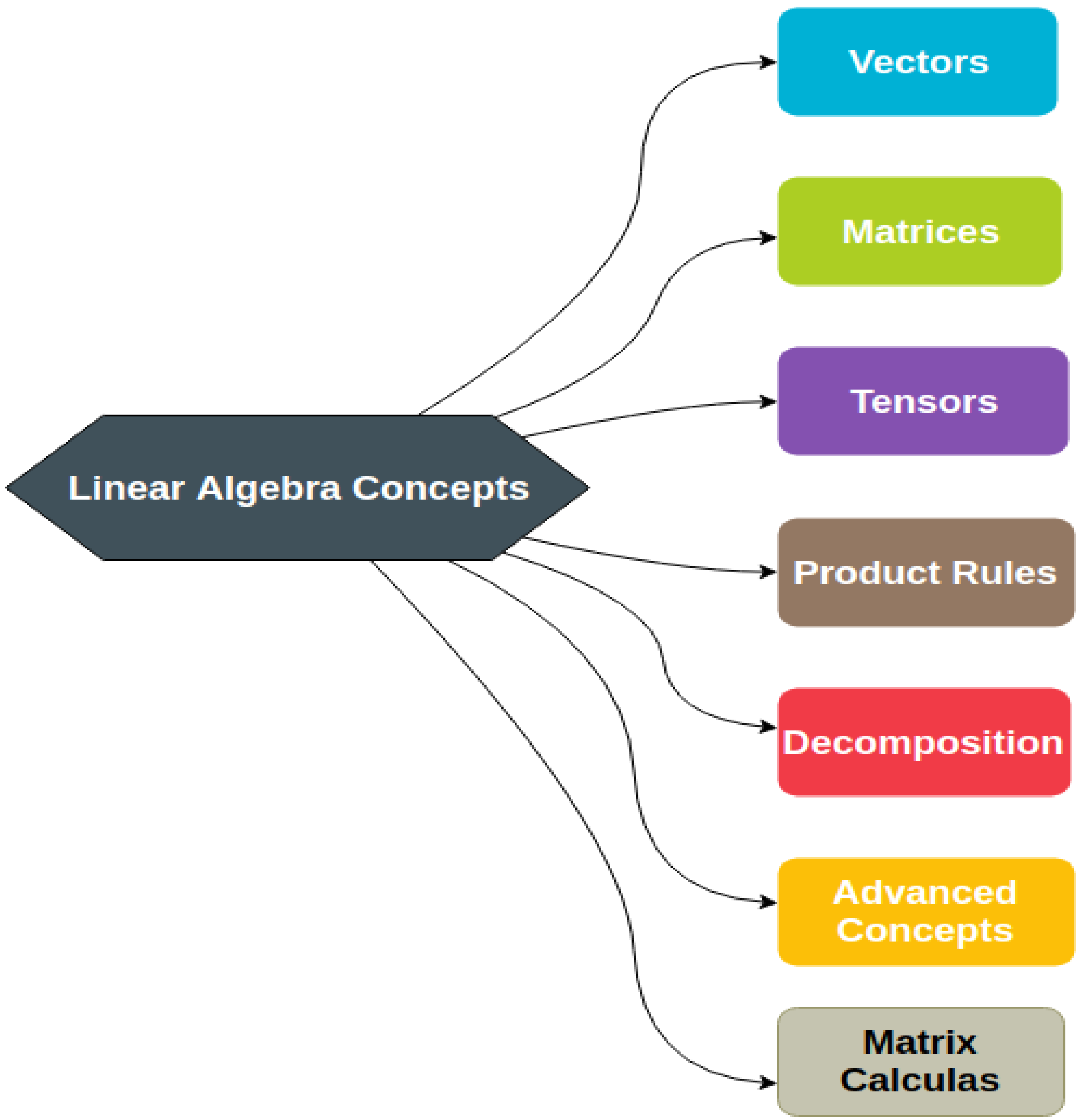
Because the number of columns in matrix A and the number of rows in matrix B are equal, we know that product AB does in fact exist. Matrix AB should have the same number of rows as A and the same number of columns as B. In this case, AB is a 1x4 matrix:

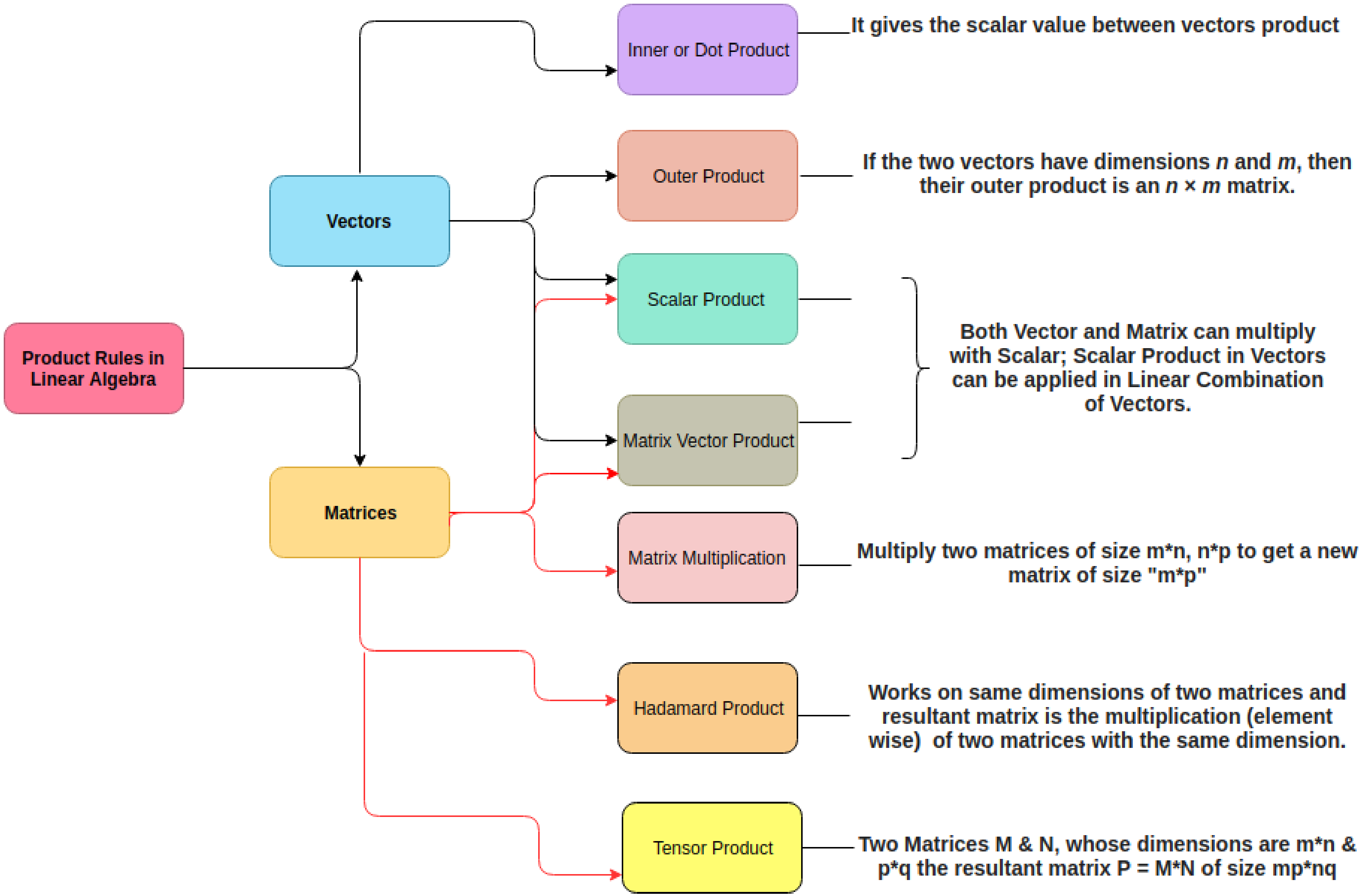
$AB = [\,5 \cdot -5 \quad 5 \cdot 8 \quad 5 \cdot 12 \quad 5 \cdot 7\,]$

Linear Algebra



Linear Algebra Concepts




Product Rules in Linear Algebra
Vectors
Matrices
Inner or Dot Product
It gives the scalar value between vectors product
Outer Product
If the two vectors have dimensions n and m, then their outer product is an n × m matrix.
Scalar Product
Matrix Vector Product
Both Vector and Matrix can multiply with Scalar; Scalar Product in Vectors can be applied in Linear Combination of Vectors.
Matrix Multiplication
Multiply two matrices of size m*n, n*p to get a new matrix of size "m*p"
Hadamard Product
Works on same dimensions of two matrices and resultant matrix is the multiplication (element wise) of two matrices with the same dimension.
Tensor Product
Two Matrices M & N, whose dimensions are m*n & p*q the resultant matrix P = M*N of size mp*nq

Vector Products

Two vectors:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Inner product = scalar

Inner product X^TY is a scalar

(1xn) (nx1)

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1y_1 + x_2y_2 + x_3y_3 = \sum_{i=1}^{3} x_i y_i$$

Outer product = matrix

$$xy^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \\ x_3y_1 & x_3y_2 & x_3y_3 \end{bmatrix}$$

Outer product XY^T is a matrix

(nx1) (1xn)

Scalar Product (Dot Product) of Two Vectors

- The scalar product (dot product) of two vectors measures the degree of similarity or alignment between them.
- It is a scalar value obtained by summing the element-wise products of the corresponding components of the two vectors.
- **Notation:** The scalar product of two vectors **u** and **v** is denoted as **u · v** or **u • v**.
- **Formula:** If **u** = $[u_1, u_2, ..., u_n]$ and **v** = $[v_1, v_2, ..., v_n]$, then the scalar product **u · v** is given by:
- $\mathbf{u} \cdot \mathbf{v} = u_1 * v_1 + u_2 * v_2 + ... + u_n * v_n$

Example: Scalar Product of Two Vectors

Consider the following two vectors:

u = [1, 2, 3] **v** = [4, 5, 6]

Step 1: Element-wise multiplication

u · v = 1 * 4 + 2 * 5 + 3 * 6

Step 2: Sum up the result

= 4 + 10 + 18

=32

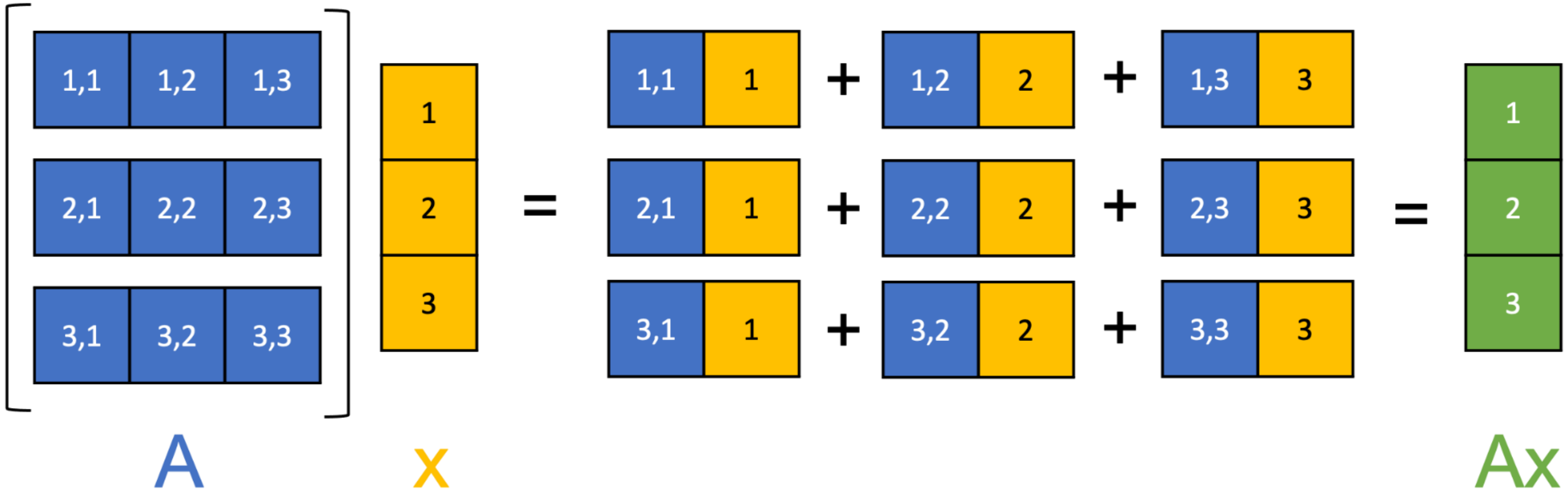
The result, 32, represents the degree of alignment or similarity between the two vectors.

Matrix Vector Product

Definition 1 (Matrix-vector multiplication): Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and vector $\boldsymbol{x} \in \mathbb{R}^n$ the **matrix-vector multiplication** of $\boldsymbol{A}$ and $\boldsymbol{x}$ is defined as

$$\boldsymbol{A}\boldsymbol{x} := x_1 \boldsymbol{a}_{*,1} + x_2 \boldsymbol{a}_{*,2} + \cdots + x_n \boldsymbol{a}_{*,n}$$

where $\boldsymbol{a}_{*,i}$ is the ith column vector of $\boldsymbol{A}$.



Matrix multiplication

- Multiplication method:

Sum over product of respective rows and columns

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \times \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

A **B**

Define output matrix

$$= \begin{bmatrix} (1\times 2)+(0\times 3) & (1\times 1)+(0\times 1) \\ (2\times 2)+(3\times 3) & (2\times 1)+(3\times 1) \end{bmatrix}$$

$$= \begin{pmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{13} & \mathbf{5} \end{pmatrix}$$

Hadamard Product (Element-wise Multiplication)

- The order of matrices/vectors to be multiplied should be same and the resulting matrix will also be of same order.

$$\overset{G}{\begin{bmatrix} 3 & 5 & 7 \\ 4 & 9 & 8 \end{bmatrix}} \circ \overset{H}{\begin{bmatrix} 1 & 6 & 3 \\ 0 & 2 & 9 \end{bmatrix}} = \overset{N}{\begin{bmatrix} 3\times 1 & 5\times 6 & 7\times 3 \\ 4\times 0 & 9\times 2 & 8\times 9 \end{bmatrix}}$$

- Hadamard product is used in image compression techniques such as JPEG.
- Hadamard Product is used in LSTM (Long Short-Term Memory) cells of Recurrent Neural Networks (RNNs).

Tensor Product

- The tensor product, also known as the Kronecker product, is an operation that combines two matrices to create a larger matrix.
- It is an important concept in linear algebra and finds applications in various fields, including quantum mechanics, signal processing, and image processing.

Tensor Product (Kronecker Product):

- Given two matrices A of size m × n and B of size p × q, the tensor product (Kronecker product) is denoted as A ⊗ B.

Formula:

- The resulting tensor product A ⊗ B is a new matrix of size (m * p) × (n * q), obtained by multiplying each element of A with the entire matrix B.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a\begin{bmatrix} e & f \\ g & h \end{bmatrix} & b\begin{bmatrix} e & f \\ g & h \end{bmatrix} \\ c\begin{bmatrix} e & f \\ g & h \end{bmatrix} & d\begin{bmatrix} e & f \\ g & h \end{bmatrix} \end{bmatrix} = \begin{bmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{bmatrix}$$

From Fundamental System of Solutions to Linear Space

- A linear space (or vector space) is a mathematical structure that consists of a set of elements called vectors, along with two operations: vector addition and scalar multiplication.

Properties of Linear Space:

Property	**Explanation**	**Example**
Closure under Vector Addition	For any vectors u and v in the space, their sum u + v is also in the space.	u + v = [1, 3] + [-2, 5] = [-1, 8] which is in R^2.
Closure under Scalar Multiplication	For any vector u in the space and any scalar c, their product c * u is also in the space.	c * u = 2 * [1, 3] = [2, 6] which is in R^2.
Associativity of Vector Addition	For any vectors u, v, and w in the space, **(u + v) + w = u + (v + w).**	([1, 3] + [-2, 5]) + [3, 4] = [-1, 8] + [3, 4] = [2, 12].
Existence of Additive Identity	There exists a special vector 0 in the space such that for any vector u in the space, u + 0 = u.	[1, 3] + [0, 0] = [1, 3].
Existence of Additive Inverse	For any vector u in the space, there exists a vector -u in the space such that u + (-u) = 0.	[1, 3] + [-1, -3] = [0, 0].

From Fundamental System of Solutions to Linear Space (Contd.)

Properties of Linear Space:

Property	**Explanation**	**Example**
Distributivity of Scalar Multiplication over Vector Addition	For any vectors u and v in the space and any scalar c, c * (u + v) = c * u + c * v.	2 * ([1, 3] + [-2, 5]) = 2 * [-1, 8] = [-2, 16] and 2 * [1, 3] + 2 * [-2, 5] = [2, 6] + [-4, 10] = [-2, 16].
Distributivity of Scalar Multiplication over Field Addition	For any vector u in the space and any scalars c and d, (c + d) * u = c * u + d * u.	(2 + 3) * [1, 3] = 5 * [1, 3] = [5, 15] and 2 * [1, 3] + 3 * [1, 3] = [2, 6] + [3, 9] = [5, 15].
Associativity of Scalar Multiplication	For any vector u in the space and any scalars c and d, (c * d) * u = c * (d * u).	(2 * 3) * [1, 3] = 6 * [1, 3] = [6, 18] and 2 * (3 * [1, 3]) = 2 * [3, 9] = [6, 18].
Identity Element for Scalar Multiplication	For any vector u in the space, 1 * u = u.	1 * [1, 3] = [1, 3].

System of Linear Equations

- In mathematics, a system of linear equations (or linear system) is a collection of one or more linear equations involving the same variables
- For example,

$$\begin{cases} 3x + 2y - z = 1 \\ 2x - 2y + 4z = -2 \\ -x + \frac{1}{2}y - z = 0 \end{cases}$$

Definition 1. *A* linear equation *in the variables $x_1, \ldots, x_n$ is an equation in the form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where b and the coefficients $a_1, \ldots, a_n$ are real or complex constants.

Example 1. *The equation $x_2 = 2(\sqrt{6} - x_1) + x_3$ is linear, whereas the equation $4x_1 - 5x_2 = x_1x_2$ is not linear, because of the term x_1x_2.*

System of Linear Equations (Contd.)

Definition 2. *A* system of linear equations *is a collection of one or more linear equations. A* solution *of the system is a list of values that makes each equation a true statement when the values are substituted for the variables. The set of all possible solutions is called the* solution set *of the linear system. Linear systems that have the same solution set are* equivalent.

Fact. *A system of linear equations has exactly one of the following:*

1. *no solution*
2. *exactly one solution*
3. *infinitely many solutions*

Solving Systems of Linear Equations

Example 2. *Given the system*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

the matrix

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

is called the coefficient matrix *and*

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

is called the augmented matrix *of the system. An augmented matrix of a system consists of the coefficient matrix with an added column containing the constants from the right sides of the equations.*

Solving Systems of Linear Equations(Contd.)

Definition 3. *A system of linear equations is said to be* consistent *if it has either one solution or infinitely many solutions. A system is* inconsistent *if it has no solution.*

The essential information of a linear system can be recorded compactly in a rectangular array called a *matrix*.

Solution.

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 10 & -10 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 30 & -30 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Thus the solution is $(x_1, x_2, x_3) = (1, 0, -1)$. It is good practice to check the solution.

There are three elementary row operations on the matrix that result in an equivalent system.

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.

2. (Interchange) Interchange two rows.

3. (Scaling) Multiply all entries in a row by a nonzero constant.

System of Linear Equations (Contd.)

Definition 4. *Two matrices are called* row equivalent *if there is a sequence of elementary row operations that transforms one matrix into the other.*

Fact. *It is important to note that row operations are reversible.*

If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set. With practice, the calculations in row operations become easier and faster. We must learn to perform row operations accurately because we will use them throughout the course.

Two fundamental questions about a linear system are as follows:

1. Is the system consistent; that is, does at least one solution exist?

2. If a solution exists, is it the only one; that is, is the solution unique?

Solving Systems of Linear Equations(Contd.)

Example 4. *Determine if the following system is consistent:*

$$
\begin{aligned}
x_2 - 4x_3 &= 8 \\
2x_1 - 3x_2 + 2x_3 &= 1 \\
5x_1 - 8x_2 + 7x_3 &= 1
\end{aligned}
$$

Solution. The augmented matrix is

$$
\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}
$$

We use elementary row operations to arrive at an equivalent triangular form of the system:

$$
\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix}
$$

The augmented matrix is now in triangular form. To interpret it correctly, go back to equation notation.

$$
\begin{aligned}
2x_1 - 3x_2 + 2x_3 &= 1 \\
x_2 - 4x_3 &= 8 \\
0 &= 5/2
\end{aligned}
$$

The equation $0 = 5/2$ is a short form of $0x_1 + 0x_2 + 0x_3 = 5/2$. There are no values of x_1, x_2, x_3 that satisfy the system, because the equation $0 = 5/2$ is never true. Thus the original system is inconsistent, that is, has no solution.

Matrices

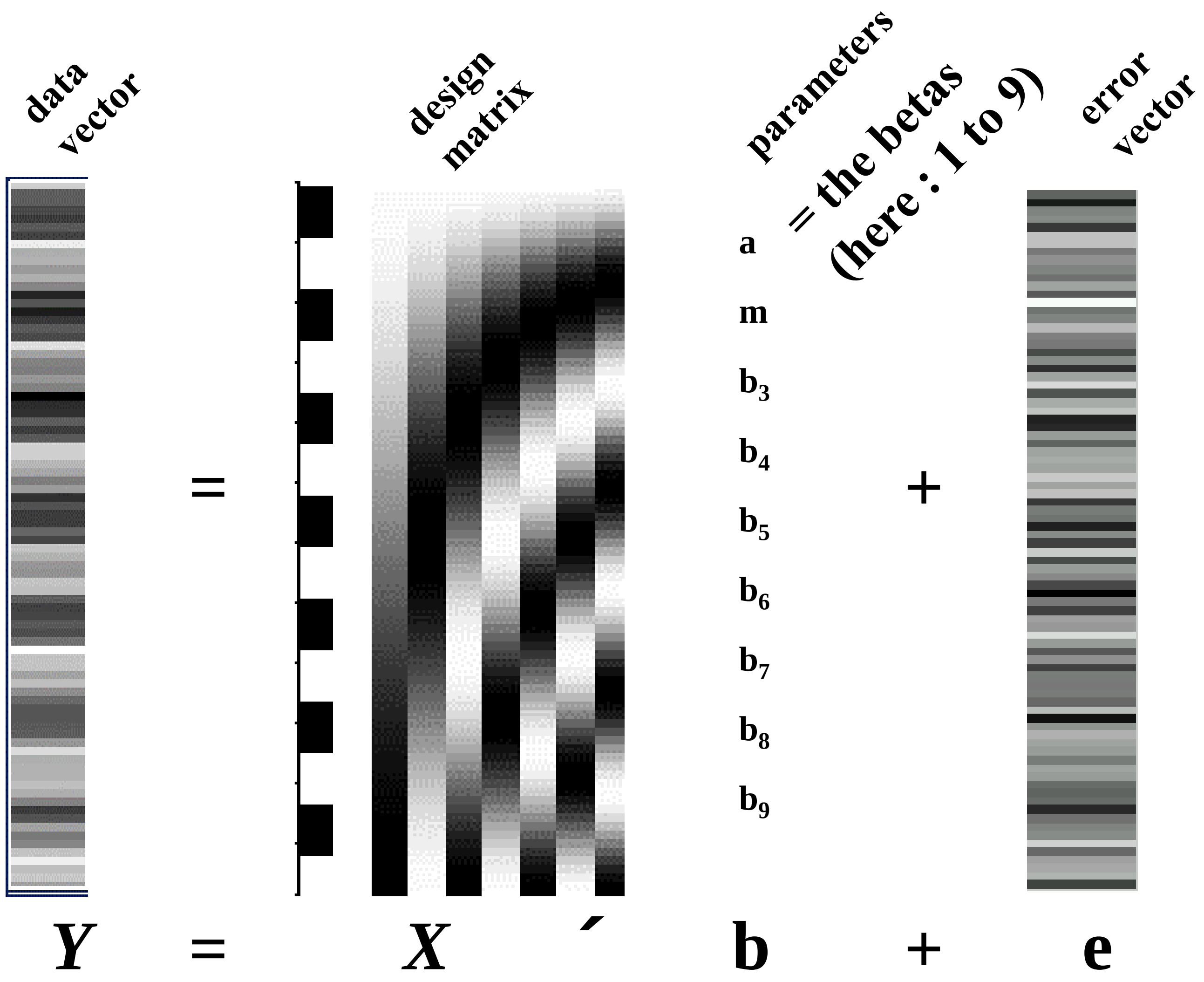
- Rectangular display of vectors in rows and columns
- Can inform about the same vector intensity at different times or different voxels at the same time
- Vector is just a n x 1 matrix

$$\mathbf{A}=\begin{bmatrix}1 & 2 & 3\\ 5 & 4 & 1\\ 6 & 7 & 4\end{bmatrix} \qquad \mathbf{C}=\begin{bmatrix}1 & 4\\ 2 & 7\\ 3 & 8\end{bmatrix} \qquad \mathbf{D}=\begin{bmatrix}d_{11} & d_{12} & d_{13}\\ d_{21} & d_{22} & d_{23}\\ d_{31} & d_{32} & d_{33}\end{bmatrix}$$

Square (3 x 3) Rectangular (3 x 2) d_{ij} : i^{th} row, j^{th} column

Defined as rows x columns (R x C)

Design matrix



Transposition

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \qquad \mathbf{b}^T = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \qquad \mathbf{d} = \begin{bmatrix} 3 & 4 & 9 \end{bmatrix} \qquad \mathbf{d}^T = \begin{bmatrix} 3 \\ 4 \\ 9 \end{bmatrix}$$

column $\longrightarrow$ row row $\longrightarrow$ column

$$\mathbf{A} = \begin{bmatrix} \boxed{1 \quad 2 \quad 3} \\ \boxed{5 \quad 4 \quad 1} \\ \boxed{6 \quad 7 \quad 4} \end{bmatrix} \qquad \mathbf{A}^T = \begin{bmatrix} 1 & 5 & 6 \\ 2 & 4 & 7 \\ 3 & 1 & 4 \end{bmatrix}$$

Matrix Calculations

Addition

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

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}2 & 4\\ 2 & 5\end{bmatrix}+\begin{bmatrix}1 & 0\\ 3 & 1\end{bmatrix}=\begin{bmatrix}2+1 & 4+0\\ 2+3 & 5+1\end{bmatrix}\begin{bmatrix}3 & 4\\ 5 & 6\end{bmatrix}$$

Subtraction

- By adding a negative matrix

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

Scalar multiplication

- Scalar x matrix = scalar multiplication

$$\lambda \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b & \lambda c \\ \lambda d & \lambda e & \lambda f \end{pmatrix}$$

Matrix Multiplication

*"When **A** is a **m**x**n** matrix & **B** is a **k**x**l** matrix, **AB** is only possible if **n=k**. The result will be an **mxl** matrix"*

$$\mathbf{m}\downarrow\ \overset{\mathbf{n}\ \longrightarrow}{\begin{matrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \\ A_7 & A_8 & A_9 \\ A_{10} & A_{11} & A_{12} \end{matrix}} \quad \text{x} \quad \overset{\mathbf{l}\ \longrightarrow}{\begin{matrix} B_{13} & B_{14} \\ B_{15} & B_{16} \\ B_{17} & B_{18} \end{matrix}}\ \downarrow\mathbf{k} \qquad = \mathbf{m} \text{ x } \mathbf{l} \text{ matrix}$$

Number of columns in A = Number of rows in B

Matrix multiplication

- Matrix multiplication is NOT commutative

$$AB \neq BA$$

- Matrix multiplication IS associative

$$A(BC)=(AB)C$$

- Matrix multiplication IS distributive

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

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

Identity matrix

Is there a matrix which plays a similar role as the number 1 in number multiplication?

Consider the nxn matrix:

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

For any nxn matrix $\boldsymbol{A}$, we have $\boldsymbol{A}\,\boldsymbol{I}_n = \boldsymbol{I}_n\,\boldsymbol{A} = \boldsymbol{A}$

For any nxm matrix $\boldsymbol{A}$, we have $\boldsymbol{I}_n\,\boldsymbol{A} = \boldsymbol{A}$, and $\boldsymbol{A}\,\boldsymbol{I}_m = \boldsymbol{A}$ (so 2 possible matrices)

Identity matrix (Contd.)

Worked example
$A\,I_3 = A$
for a 3x3 matrix:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+0+0 & 0+2+0 & 0+0+3 \\ 4+0+0 & 0+5+0 & 0+0+6 \\ 7+0+0 & 0+8+0 & 0+0+9 \end{bmatrix}$$

Matrix inverse(Contd.)

- **Definition.** A matrix ***A*** is called **nonsingular** or **invertible** if there exists a matrix ***B*** such that:

$$A\,B = B\,A = I_n$$

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

- **Notation.** A common notation for the inverse of a matrix ***A*** is $\boldsymbol{A}^{-1}$. So:

$$A\,A^{-1} = A^{-1}\,A = I_n\,.$$

- The inverse matrix is unique when it exists. So if ***A*** is invertible, then $\boldsymbol{A}^{-1}$ is also invertible and then $(A^T)^{-1} = (A^{-1})^T$

•Matrix division: A/B= $\boldsymbol{A}*\boldsymbol{B}^{-1}$

Matrix inverse (Contd.)

- For a XxX square matrix: $$A = \begin{pmatrix} x_{1,1} & \dots & x_{1,j} \\ \vdots & \ddots & \vdots \\ x_{i,1} & \dots & x_{i,j} \end{pmatrix}$$

- The inverse matrix is: $$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \mathrm{cof}(A, x_{1,1}) & \dots & \mathrm{cof}(A, x_{1,j}) \\ \vdots & \ddots & \vdots \\ \mathrm{cof}(A, x_{i,1}) & \dots & \mathrm{cof}(A, x_{i,j}) \end{pmatrix}^T$$

- E.g.: 2x2 matrix $$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

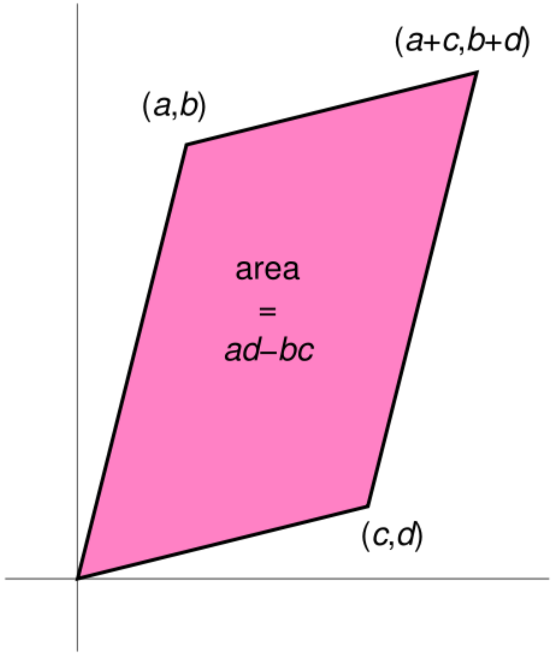
Determinants

- Determinants are mathematical objects that are very useful in the analysis and solution of systems of linear equations.
- The **determinant** is a function that associates a scalar det(A) to every square matrix A.
  - Input is nxn matrix
  - Output is a single number (real or complex) called the determinant

Determinants (Contd.)

- Determinants can only be found for square matrices.
- For a 2x2 matrix A, **det(A) = ad-bc**. Lets have at closer look at that:

$$\det(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$



- A matrix A has an inverse matrix A^{-1} if and only if **det(A)≠0**.

Solving simultaneous equations

For one linear equation ***ax=b*** where the unknown is *x* and *a* and *b* are constants,

3 possibilities:

- If $\boxed{a \neq 0}$ then $x = \frac{b}{a} \equiv a^{-1}b$ thus there is single solution
- If $\boxed{a = 0,\ b = 0}$ then the equation $ax = b$ becomes $0 = 0$ and any value of *x* will do
- If $\boxed{a = 0,\ b \neq 0}$ then $ax = b$ becomes $0 = b$ which is a contradiction

With >1 equation and >1 unknown

- Can use solution $x = a^{-1}b$ from the single equation to solve
- For example

$$2x_1 + 3x_2 = 1$$
$$x_1 - 2x_2 = 4$$

- In matrix form

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

A X = B

$$X = A^{-1}B$$

Contd.

- X = A^{-1}B
- To find A^{-1}

$$A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

- Need to find determinant of matrix A

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

- From earlier

$$\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix} \qquad (2 \cdot -2) - (3 \cdot 1) = -4 - 3 = -7$$

- So determinant is -7

Contd.

$$A^{-1} = \frac{1}{(-7)}\begin{bmatrix} -2 & -3 \\ -1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}$$

if B is $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

$$x = a^{-1}b$$

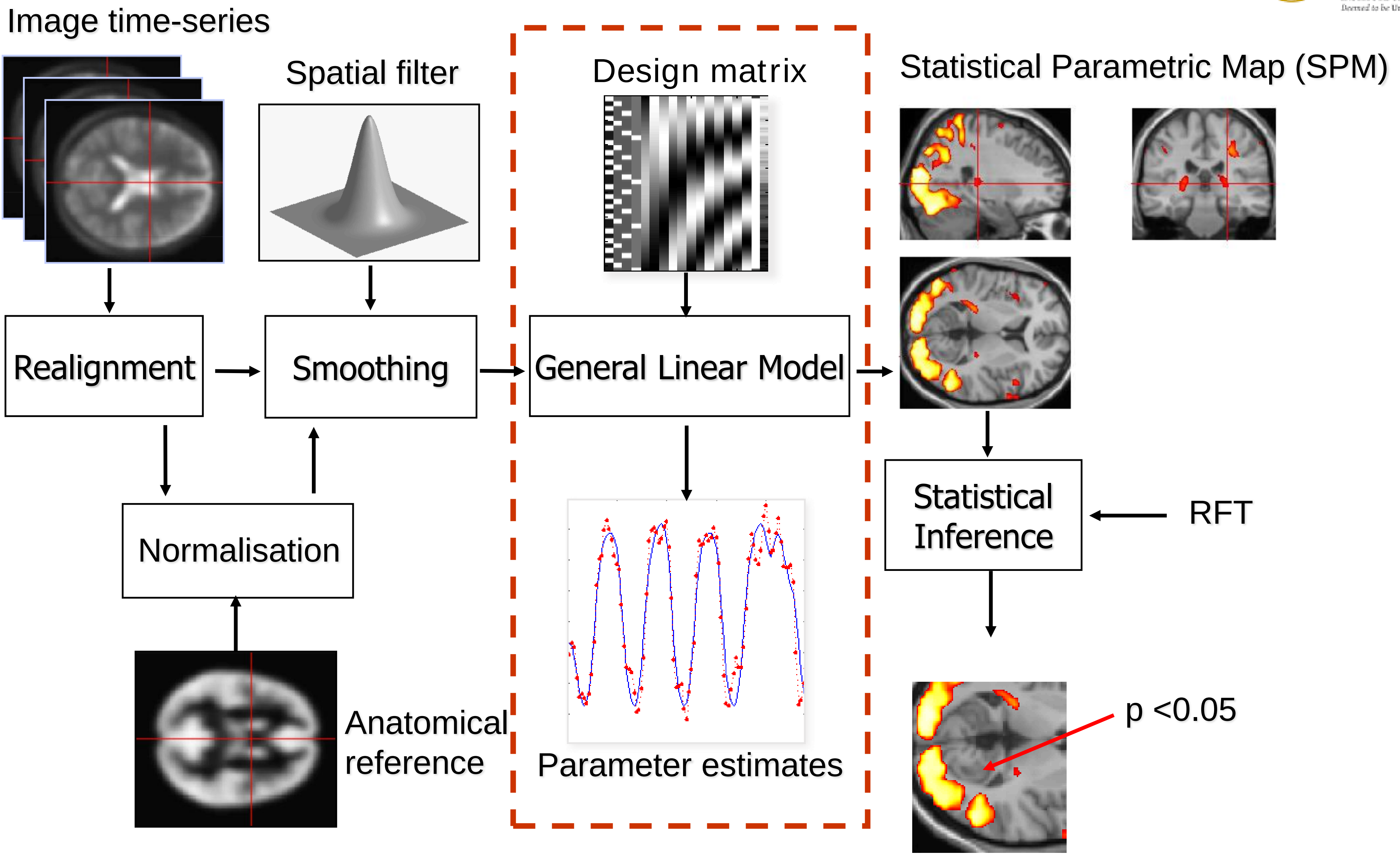
$$X = \frac{1}{7}\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 14 \\ -7 \end{bmatrix} = \frac{2}{-1}$$

So $\boxed{\begin{matrix} x_1 = 2 \\ x_2 = -1 \end{matrix}}$

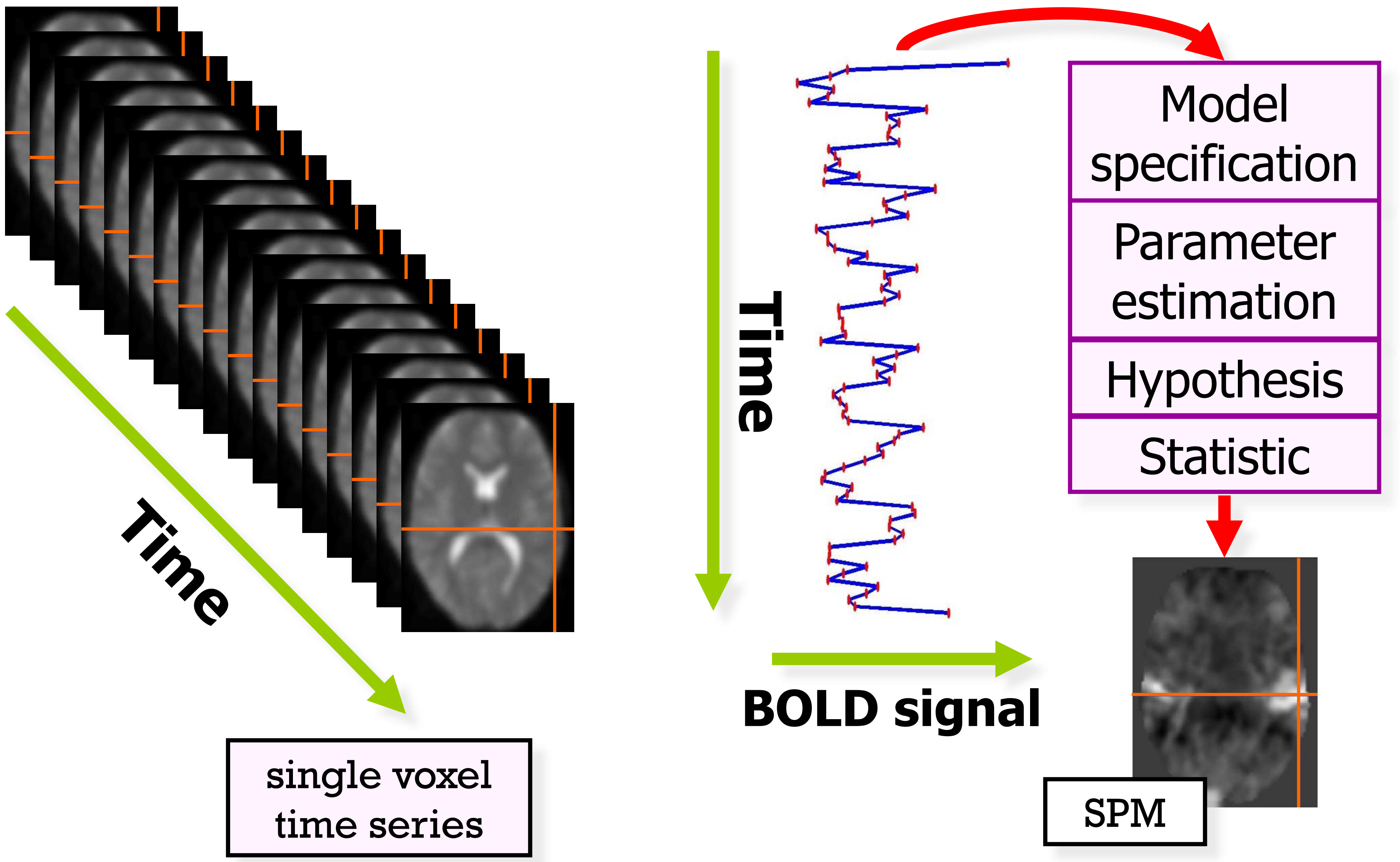
How are matrices relevant to fMRI data?

How are matrices relevant to fMRI data?

SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

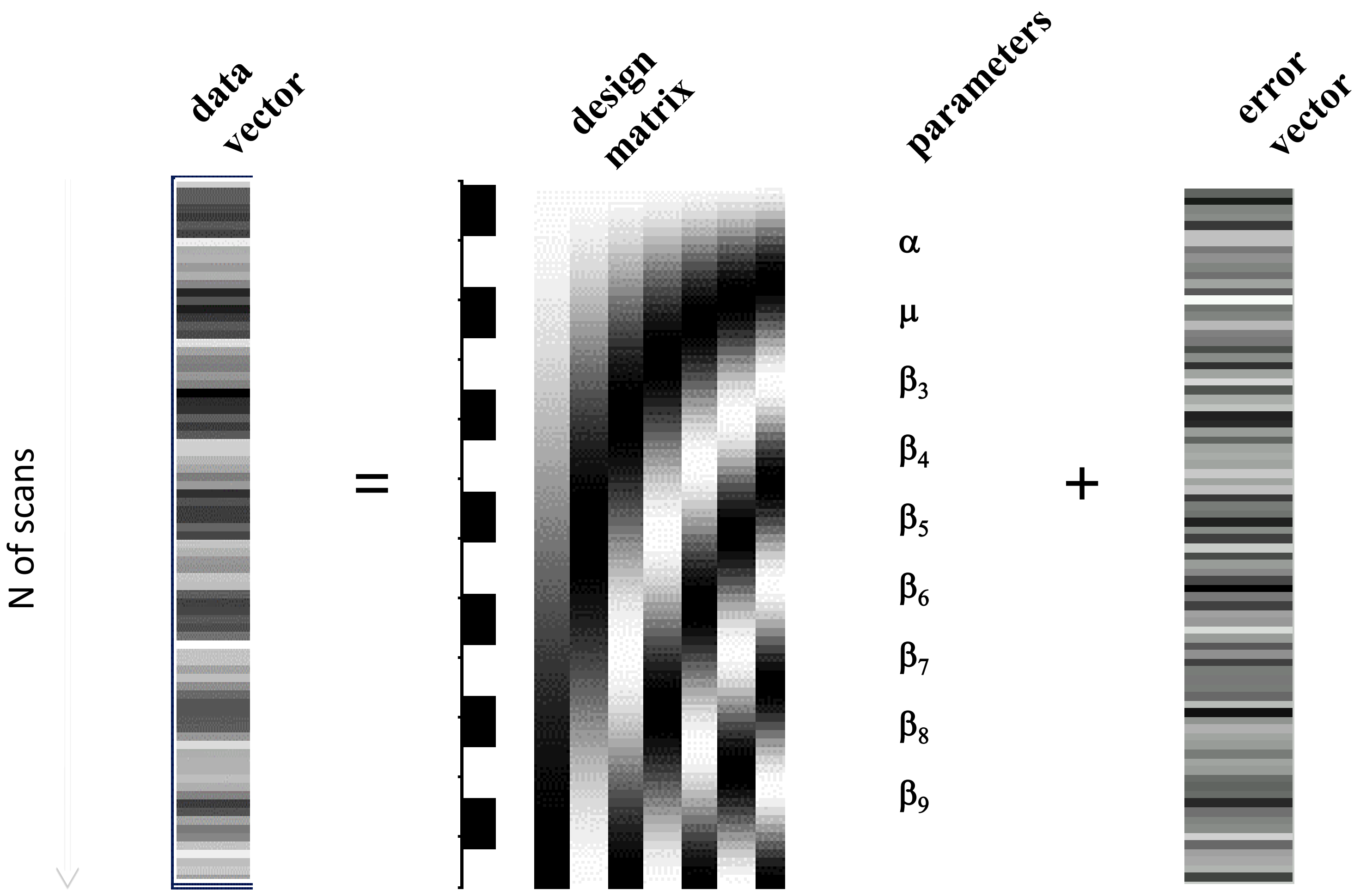


Voxel-wise time series analysis



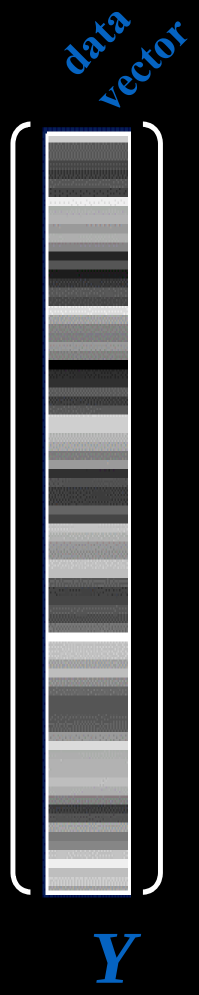
How are matrices relevant to fMRI data?

(Generalized Linear Model (GLM) equation



$$Y = X \times \beta + \varepsilon$$

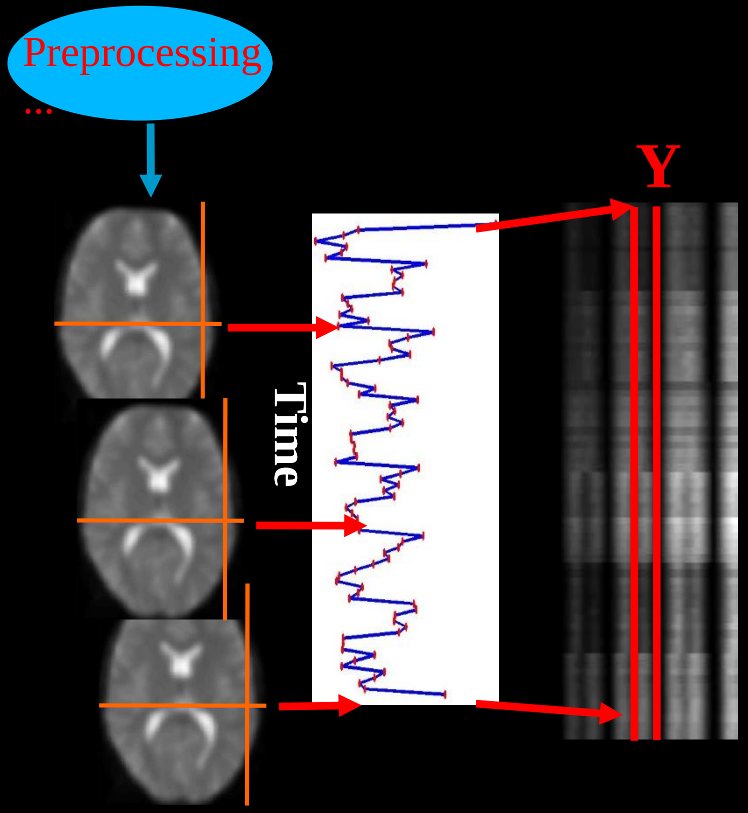
How are matrices relevant to fMRI data?



[Response variable]

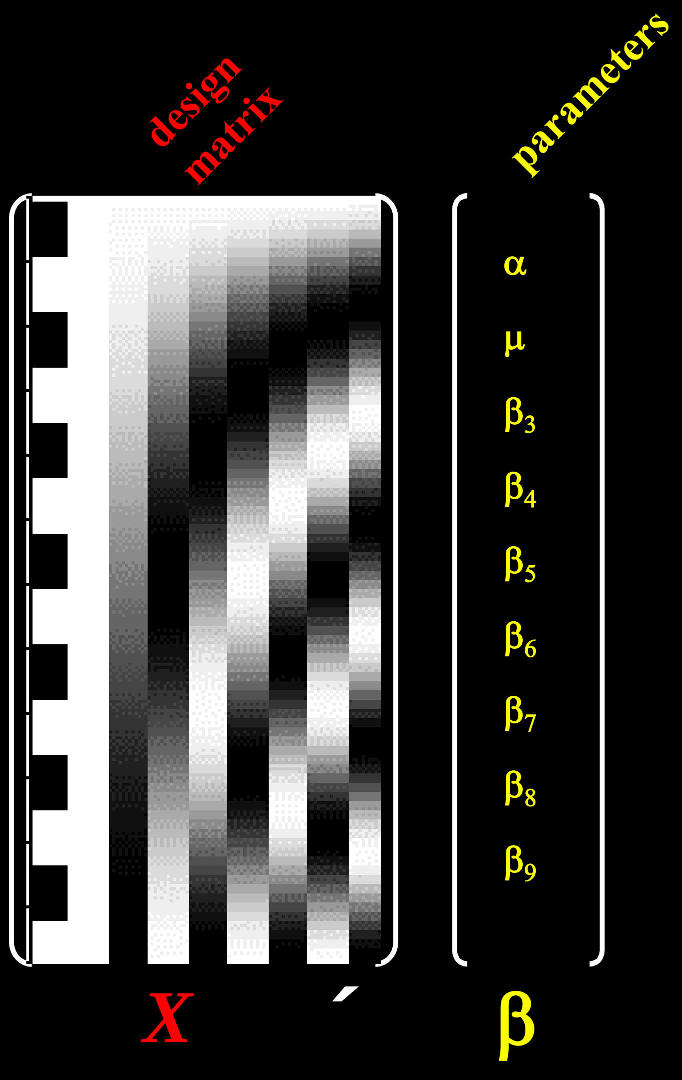
e.g BOLD signal at a particular voxel

A single voxel sampled at successive time points.
Each voxel is considered as independent observation.



$$Y = X \, . \, \beta + \varepsilon$$

How are matrices relevant to fMRI data?



Explanatory variables

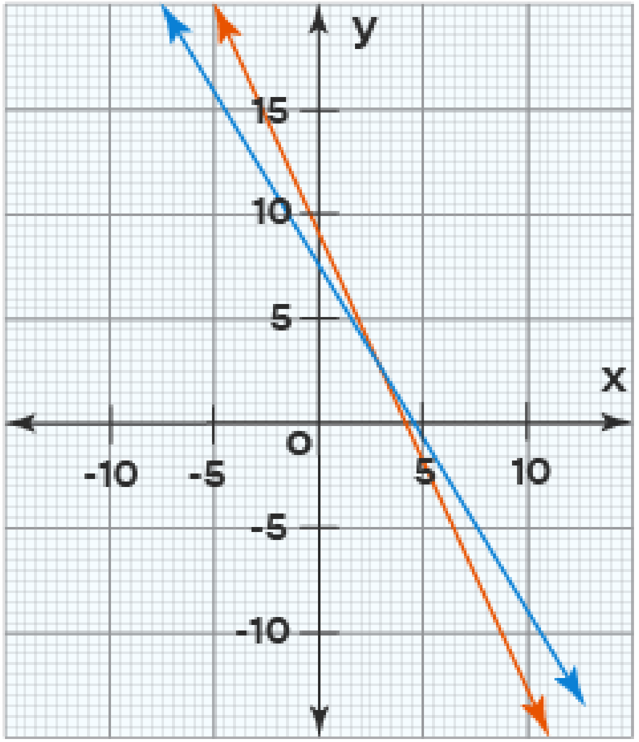
- These are assumed to be measured without error.
- May be continuous;
- May be dummy, indicating levels of an experimental factor.

Solve equation for β – tells us how much of the BOLD signal is explained by X

$$Y = X \,.\, \beta + \varepsilon$$

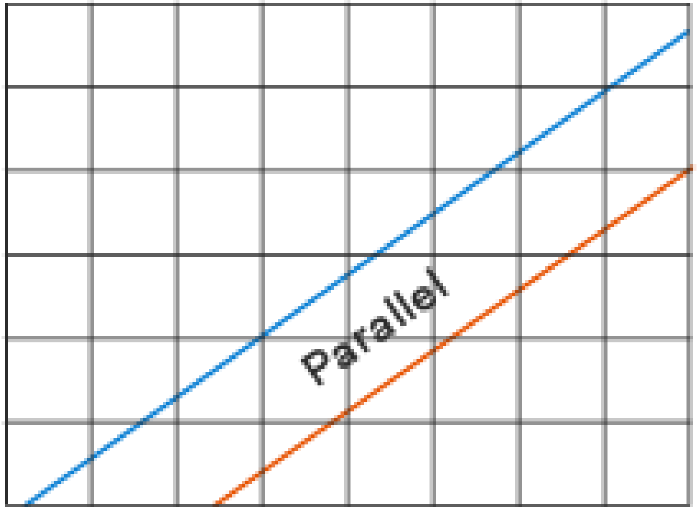
What are the Solutions of a Linear Equation?

- The solutions of linear equations are the points at which the lines or planes representing the linear equations intersect or meet each other.
- A solution set of a system of linear equations is the set of values to the variables of all possible solutions.
- For example, while solving linear equations one can visualize the solution of a system of simultaneous linear equations by drawing 2 linear graphs and finding out their intersection point.

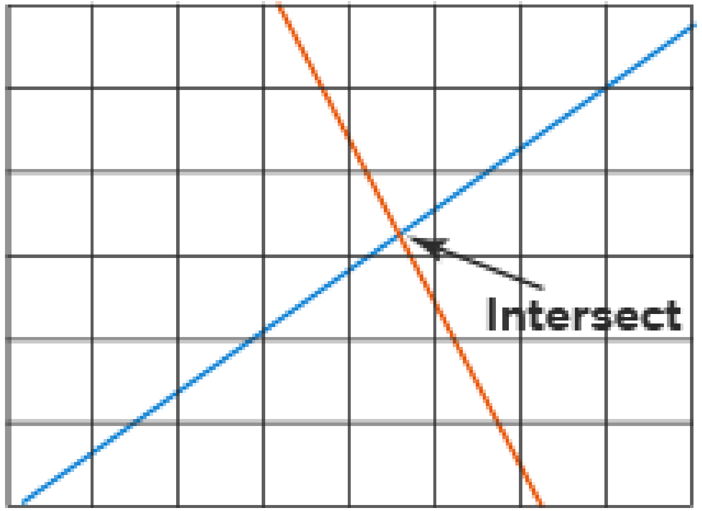


Types of Solutions for Linear Equations

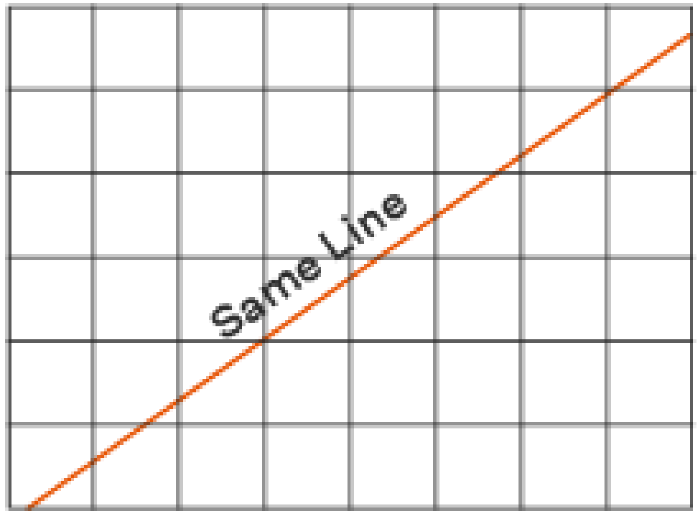
- A system of linear equations can have 3 types of solutions.



No Solution



One Solution



∞ Solution

Types of Solutions for Linear Equations

Unique Solution of a system of linear equations

The unique solution of a linear equation means that there exists only one point, on substituting which, L.H.S and R.H.S of an equation become equal. The linear equation in one variable has always a unique solution.

- For example, 3m =6 has a unique solution m = 2 for which L.H.S = R.H.S. Similarly, for simultaneous linear equations in two variables, the unique solution is an ordered pair (x,y) which will satisfy both the equations.

No Solution

A system of linear equations has no solution when there exists no point where lines intersect each other or the graphs of linear equations are parallel.

Infinite Many Solutions

A system of linear equations has infinitely many solutions when there exists a solution set of infinite points for which L.H.S and R.H.S of an equation become equal, or in the graph straight lines overlap each other.

How to find the Solution of a Linear Equation?

Solutions for Linear Equations in One Variable

Consider the equation, $2x + 4 = 8$

- To find the value of x, first, we remove 4 from L.H.S, so we subtract 4 from both sides of the equation. $2x + 4 - 4 = 8 - 4$
- Simply. Now we get, $2x = 4$
- Now we have to remove 2 from L.H.S in order to get x, therefore we divide the equation by 2. $2x/2 = 4/2$, $x=2$

Hence, the solution of the equation $2x + 4 = 8$ is $x=2$.

Types of solutions using the following conditions:

- Unique solution (consistent and independent) $a1/a2 \neq b1/b2$
- No solution (inconsistent and independent) $a1/a2 = b1/b2 \neq c1/c2$
- Infinite many solutions (consistent and dependent) $a1/a2 = b1/b2 = c1/c2$

Vector Space

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

Definition of Vector Space

Let V be a set on which two operations (**vector addition** and **scalar multiplication**) are defined. If the listed axioms are satisfied for every $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and every scalar (real number) c and d, then V is called a **vector space.**

Addition:

1. $\mathbf{u} + \mathbf{v}$ is in V.	Closure under addition
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Commutative property
3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	Associative property
4. V has a **zero vector** $\mathbf{0}$ such that for every $\mathbf{u}$ in V, $\mathbf{u} + \mathbf{0} = \mathbf{u}$.	Additive identity
5. For every $\mathbf{u}$ in V, there is a vector in V denoted by $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.	Additive inverse

Scalar Multiplication:

6. $c\mathbf{u}$ is in V.	Closure under scalar multiplication
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	Distributive property
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	Distributive property
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$	Associative property
10. $1(\mathbf{u}) = \mathbf{u}$	Scalar identity

Vector Space (Contd.)

Example: Let $V = \{(x, \frac{1}{2}x) : x \text{ real number}\}$ with standared operations. Is it a vector space. Justify your answer.

Solution:

1. Addition:

(a) For real numbers x, y, We have

$$\left(x, \frac{1}{2}x\right) + \left(y, \frac{1}{2}y\right) = \left(x + y, \frac{1}{2}(x + y)\right).$$

So, V is closed under addition.

(b) Clearly, addition is closed under addition.

(c) Clearly, addition is associative.

(d) The element $\mathbf{0} = (0, 0)$ satisfies the property of the zero element.

(e) We have $-\left(x, \frac{1}{2}x\right) = \left(-x, \frac{1}{2}(-x)\right)$ .So, every element in V has an additive inverse.

2. Scalar multiplication:

(a) For a scalar c, we have

$$c\left(x, \frac{1}{2}x\right) = \left(cx, \frac{1}{2}cx\right).$$

So, V is closed under scalar multiplication.

(b) The distributivity $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ works for $\mathbf{u}, \mathbf{v}$ in V.

(c) The distributivity $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ works, for $\mathbf{u}$ in V and scalars c, d.

(d) The associativity $c(d\mathbf{u}) = (cd)\mathbf{u}$ works.

(e) Also $1\mathbf{u} = \mathbf{u}$.

Linear Independence

Definition 1 A vector $\mathbf{v}$ in a vector space V is called a **linear combination** of vectors $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}$ in V if $\mathbf{v}$ can be written in the form

$$\mathbf{v} = c_1\mathbf{u_1} + c_2\mathbf{u_2} + \cdots + c_k\mathbf{u_k},$$

where $c_1, c_2, \ldots, c_k$ are scalars.

Definition 2 Let V be a vector space over $\mathbb{R}$ and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}\}$ be a subset of V. We say that S is a **spanning set** of V if every vector $\mathbf{v}$ of V can be written as a liner combination of vectors in S. In such cases, we say that S **spans** V.

Linear Independence (Contd.)

Definition 3 Let V be a vector space over $\mathbb{R}$ and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}\}$ be a subset of V. Then the **span of** S is the set of all linear combinations of vectors in S,

$$span(S) = \{c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k} : c_1, c_2, \ldots, c_k \ are \ scalars\}.$$

1. The span of S is denoted by $span(S)$ as above or $span\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}\}$
2. If $V = span(S)$, then say V is spanned by S or S spans V.

Linear Independence (Contd.)

Definition 4 Let V be a vector space. A set of elements (vectors) $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k}\}$ is said to be **linearly independent** if the equation

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k} = \mathbf{0}$$

has only trivial solution

$$c_1 = 0, c_2 = 0, \ldots, c_k = 0.$$

We say S is **linearly dependent**, if S in not linearly independent. (*This means, that S is said to be linearly dependent, if there is at least one nontrivial (i.e. nonzero) solutions to the above equation.*)

Linear Independence (Contd.)

Testing for linear independence

Suppose V is a subspace of the $n-$space $\mathbb{R}^n$. Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k}\}$ be a set of elements (i.e. vectors) in V. To test whether S is linearly independent or not, we do the following:

1. From the equation

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k} = \mathbf{0},$$

   write a homogeneous system of equations in variabled $c_1, c_2, \ldots, c_k$.

2. Use Gaussian elemination (with the help of TI) to determine whether the system has a unique solutions.

3. If the system has only the trivial solution

$$c_1 = 0, c_2 = 0, \cdots, c_k = 0,$$

   then S is linearly independent. Otherwise, S is linearly dependent

Example 1 Let $S = \{(6,2,1), (-1,3,2)\}$. Determine, if S is linearly independent or dependent?

Solution: Let

$$c(6,2,1) + d(-1,3,2) = (0,0,0).$$

If this equation has only trivial solutions, then it is linealry independent. This equaton gives the following system of linear equations:

$$\begin{aligned} 6c \quad -d &= 0 \\ 2c \quad +3d &= 0 \\ c \quad +2d &= 0 \end{aligned}$$

The augmented matrix for this system is

$$\begin{bmatrix} 6 & -1 & 0 \\ 2 & 3 & 0 \\ 1 & 2 & 0 \end{bmatrix}. \quad its\ gauss - Jordan\ form: \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

So, $c = 0, d = 0$. The system has only trivial (i.e. zero) solution. We conclude that S is linearly independent.

Basis

Definition 1 Let V be a vector space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k}\}$ be a set of elements (vectors)in V. We say that S is a **basis** of V if

1. S spans V and
2. S is linearly independent.

Example 1 Most standard example of basis is the **standard basis** of $\mathbb{R}^n$.

1. Consider the vector space $\mathbb{R}^2$. Write

$$\mathbf{e_1} = (1, 0), \mathbf{e_2} = (0, 1).$$

Then, $\mathbf{e_1}, \mathbf{e_2}$ form a basis of $\mathbb{R}^2$.

2. Consider the vector space $\mathbb{R}^3$. Write

$$\mathbf{e_1} = (1, 0, 0), \mathbf{e_2} = (0, 1, 0), \mathbf{e_2} = (0, 0, 1).$$

Then, $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ form a basis of $\mathbb{R}^3$.

Proof. First, for any vector $\mathbf{v} = (x_1, x_2, x_3) \in \mathbb{R}^3$, we have

$$\mathbf{v} = x_1\mathbf{e_1} + x_2\mathbf{e_2} + x_3\mathbf{e_3}.$$

So, $\mathbb{R}^3$ is spanned by $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$.

Now, we prove that $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ are linearly independent. So, suppose

$$c_1\mathbf{e_1} + c_2\mathbf{e_2} + c_3\mathbf{e_3} = \mathbf{0} \quad OR \quad (c_1, c_2, c_3) = (0, 0.0).$$

So, $c_1 = c_2 = c_3 = 0$. Therefore, $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ are linearly independent. Hence $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ forms a basis of $\mathbb{R}^3$. The proof is complete.

Basis (Contd.)

3. More generally, consider vector space $\mathbb{R}^n$. Write

$$\mathbf{e_1} = (1, 0, \ldots, 0), \mathbf{e_2} = (0, 1, \ldots, 0), \ldots, \mathbf{e_n} = (0, 0, \ldots, 1).$$

Then, $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \ldots, \mathbf{e_n}$ form a basis of $\mathbb{R}^n$. The proof will be similar to the above proof. This basis is called the **standard basis** of $\mathbb{R}^n$.

Basis (Contd.)

Example

Consider

$$\mathbf{v_1} = (1,1,1), \mathbf{v_2} = (1,-1,1), \mathbf{v_3} = (1,1,-1) \qquad in \quad \mathbb{R}^3.$$

Then $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ form a basis for $\mathbb{R}^3$.

Basis (Contd.)

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

Proof. First, we prove that $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly independent. Let

$c_1\mathbf{v_1}+c_2\mathbf{v_2}+c_3\mathbf{v_3} = \mathbf{0}$. $\quad OR \quad c_1(1,1,1)+c_2(1,-1,1)+c_3(1,1,-1) = (0,0,0)$.

We have to prove $c_1 = c_2 = c_3 = 0$. The equations give the following system of linear equations:

$$\begin{aligned} c_1 \quad +c_2 \quad +c_3 \quad &= 0 \\ c_1 \quad -c_2 \quad +c_3 \quad &= 0 \\ c_1 \quad +c_2 \quad -c_3 \quad &= 0 \end{aligned}$$

The augmented matrix is

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix} \quad its\ Gauss - Jordan\ form \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

So, $c_1 = c_2 = c_3 = 0$ and this estblishes that $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ are linearly independent.

Basis (Contd.)

Now to show that $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ spans $\mathbb{R}^3$, let $\mathbf{v} = (x_1, x_2, x_3)$ be a vector in $\mathbb{R}^3$. We have to show that, we can find c_1, c_2, c_3 such that

$$(x_1, x_2, x_3) = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3}$$

OR

$$(x_1, x_2, x_3) = c_1(1, 1, 1) + c_2(1, -1, 1) + c_3(1, 1, -1).$$

This gives the system of linear equations:

$$\begin{bmatrix} c_1 & +c_2 & +c_3 \\ c_1 & -c_2 & +c_3 \\ c_1 & +c_2 & -c_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad OR \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Basis (Contd.)

The coefficient matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \quad \textit{has inverse} \quad A^{-1} = \begin{bmatrix} 0 & .5 & .5 \\ .5 & -.5 & 0 \\ .5 & 0 & -.5 \end{bmatrix}.$$

So, the above system has tha solution:

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = A^{-1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & .5 & .5 \\ .5 & -.5 & 0 \\ .5 & 0 & -.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

So, each vector (x_1, x_2, x_3) is in the span of $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$. So, they form a basis of $\mathbb{R}^3$. The proof is complete.

Rank

Definition Let $A = [a_{ij}]$ be an $m \times n$ matrix.

1. The $n-$tuples corresponding to the rows of A are called **row vectors** of A.

2. Similarly, the $m-$tuples corresponding to the columns of A are called **column vectors** of A.

3. The **row space** of A is the subspace of $\mathbb{R}^n$ spanned by row vectors of A.

4. The **column space** of A is the subspace of $\mathbb{R}^m$ spanned by column vectors of A.

Rank (Contd.)

Definition Suppose A is an $m \times n$ matrix. The dimension of the row space (equivalently, of the column space) of A is called the **rank** of A and is denoted by $rank(A)$.

Rank of Matrix: Number of nonzero rows in the matrix after reducing to echelon form.

Given: $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

$R_2 \to R_2 - R_1$

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

$R_3 \to R_3 - R_1$

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

Hence, only one non zero row exists.

$\boxed{\therefore \text{Rank} = 1}$

Linear Mapping

Definition *linear map*

A *linear map* from V to W is a function $T : V \to W$ with the following properties:

additivity

$$T(u + v) = Tu + Tv \text{ for all } u, v \in V;$$

homogeneity

$$T(\lambda v) = \lambda(Tv) \text{ for all } \lambda \in \mathbf{F} \text{ and all } v \in V.$$

Linear Mapping

1. If $T : U \to V$ is a linear map over a field F, then for any scalars $\alpha_i \in F$ and vectors $u_i \in U$ such that $T(\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n) = \alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots + \alpha_n T(u_n)$.

2. A linear map $T : U \to V$ is said to be one to one when $T(u_1) = T(u_2)$ where $u_1, u_2 \in U$, then $u_1 = u_2$.

3. A linear map $T : U \to V$ is said to be onto if the range of T is equal to V. That is, for every v in V there exists a vector $u \in U$ such that $v = T(u)$.

4. Let U and V be the two vector spaces over the field F. Suppose $T : U \to V$ is a linear map, then

   (i) $T(0) = 0$

   (ii) $T(-u) = -T(u)$ for all $u \in U$

   (iii) $T(u_1 - u_2) = T(u_1) - T(u_2)$ for all $u_1, u_2 \in U$

Linear Transformations

1. Introduction to Linear Transformations
2. Matrices for Linear Transformations
3. Similarity

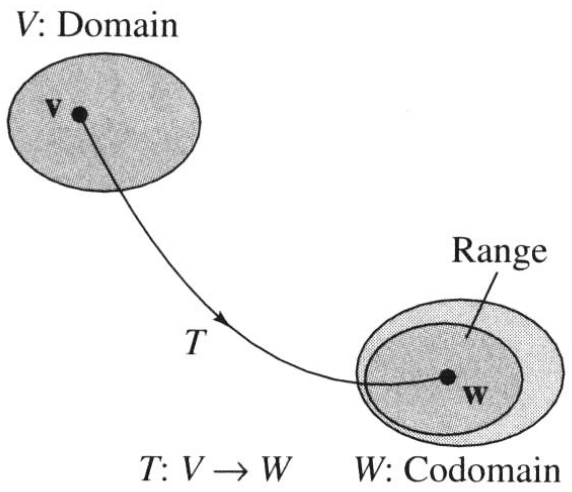
1. Introduction to Linear Transformations

- Function T that maps a vector space V into a vector space W:

$$T: V \xrightarrow{\text{mapping}} W, \quad V, W: \text{vector space}$$

V: the domain of T

W: the codomain of T



- Image of $\mathbf{v}$ under T:

If $\mathbf{v}$ is in V and $\mathbf{w}$ is in W such that

$$T(\mathbf{v}) = \mathbf{w}$$

Then $\mathbf{w}$ is called the image of $\mathbf{v}$ under T.

- the range of T:

The set of all images of vectors in V.

- the preimage of $\mathbf{w}$:

The set of all $\mathbf{v}$ in V such that $T(\mathbf{v})=\mathbf{w}$.

■Ex 1: (A function from R^2 into R^2)

$$T: R^2 \rightarrow R^2 \qquad \mathbf{v} = (v_1, v_2) \in R^2$$

$$T(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2)$$

(a) Find the image of **v**=(-1,2). (b) Find the preimage of **w**=(-1,11)

Sol:

$$(a) \ \ \mathbf{v} = (-1, 2)$$

$$\Rightarrow T(\mathbf{v}) = T(-1, 2) = (-1-2, -1+2(2)) = (-3, 3)$$

$$(b) \ \ T(\mathbf{v}) = \mathbf{w} = (-1, 11)$$

$$T(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2) = (-1, 11)$$

$$\Rightarrow v_1 - v_2 = -1$$

$$v_1 + 2v_2 = 11$$

$\Rightarrow v_1 = 3, \ v_2 = 4$ Thus {(3, 4)} is the preimage of **w**=(-1, 11).

■Linear Transformation (L.T.):

V, W : vector space

$T: V \rightarrow W$: V to W linear transformation

$$(1)\ T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in V$$

$$(2)\ T(c\mathbf{u}) = cT(\mathbf{u}), \quad \forall c \in R$$

- Notes:

(1) A linear transformation is said to be operation preserving.

$$T(\mathbf{u} \underbrace{+}_{\text{Addition in } V} \mathbf{v}) = T(\mathbf{u}) \underbrace{+}_{\text{Addition in } W} T(\mathbf{v}) \qquad T(\underbrace{c\mathbf{u}}_{\text{Scalar multiplication in } V}) = \underbrace{cT}_{\text{Scalar multiplication in } W}(\mathbf{u})$$

(2) A linear transformation $T : V \to V$ from a vector space into itself is called a **linear operator**.

■Ex 2: (Verifying a linear transformation *T* from R^2 into R^2)

$$T(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2)$$

Pf:

$\mathbf{u} = (u_1, u_2),\ \mathbf{v} = (v_1, v_2)$: vector in R^2, c : any real number

(1)Vector addition :

$$\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$$

$$\begin{aligned}
T(\mathbf{u} + \mathbf{v}) &= T(u_1 + v_1, u_2 + v_2) \\
&= ((u_1 + v_1) - (u_2 + v_2), (u_1 + v_1) + 2(u_2 + v_2)) \\
&= ((u_1 - u_2) + (v_1 - v_2), (u_1 + 2u_2) + (v_1 + 2v_2)) \\
&= (u_1 - u_2, u_1 + 2u_2) + (v_1 - v_2, v_1 + 2v_2) \\
&= T(\mathbf{u}) + T(\mathbf{v})
\end{aligned}$$

(2) Scalar multiplication

$$c\mathbf{u} = c(u_1, u_2) = (cu_1, cu_2)$$

$$\begin{aligned} T(c\mathbf{u}) &= T(cu_1, cu_2) = (cu_1 - cu_2, cu_1 + 2cu_2) \\ &= c(u_1 - u_2, u_1 + 2u_2) \\ &= cT(\mathbf{u}) \end{aligned}$$

Therefore, T is a linear transformation.

- Ex 3: (Functions that are not linear transformations)

$$(a)\ f(x) = \sin x$$

$$\sin(x_1 + x_2) \neq \sin(x_1) + \sin(x_2)$$

$$\sin(\tfrac{\pi}{2} + \tfrac{\pi}{3}) \neq \sin(\tfrac{\pi}{2}) + \sin(\tfrac{\pi}{3})$$

$\Leftarrow f(x) = \sin x$ is not a linear transformation

$$(b)\ f(x) = x^2$$

$$(x_1 + x_2)^2 \neq x_1^2 + x_2^2$$

$\Leftarrow f(x) = x^2$ is not a linear transformation

$$(1+2)^2 \neq 1^2 + 2^2$$

$$(c)\ f(x) = x + 1$$

$$f(x_1 + x_2) = x_1 + x_2 + 1$$

$$f(x_1) + f(x_2) = (x_1 + 1) + (x_2 + 1) = x_1 + x_2 + 2$$

$$f(x_1 + x_2) \neq f(x_1) + f(x_2)$$

$\Leftarrow f(x) = x + 1$ is not a linear transformation

- Notes: Two uses of the term “linear”.

(1) $f(x) = x + 1$ is called a linear function because its graph is a line.

(2) $f(x) = x + 1$ is not a linear transformation from a vector space R into R because it preserves neither vector addition nor scalar multiplication.

- Zero transformation:

$$T: V \to W \qquad T(\mathbf{v}) = 0, \ \forall \mathbf{v} \in V$$

- Identity transformation:

$$T: V \to V \qquad T(\mathbf{v}) = \mathbf{v}, \ \forall \mathbf{v} \in V$$

- Thm 6.1: (Properties of linear transformations)

$$T: V \to W, \quad \mathbf{u}, \mathbf{v} \in V$$

$$(1)\ T(\mathbf{0}) = \mathbf{0}$$

$$(2)\ T(-\mathbf{v}) = -T(\mathbf{v})$$

$$(3)\ T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$$

(4) If $\mathbf{v} = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$

Then $T(\mathbf{v}) = T(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)$

$$= c_1 T(v_1) + c_2 T(v_2) + \cdots + c_n T(v_n)$$

- Ex 4: (Linear transformations and bases)

Let $T: R^3 \rightarrow R^3$ be a linear transformation such that

$$T(1,0,0) = (2,-1,4)$$

$$T(0,1,0) = (1,5,-2)$$

$$T(0,0,1) = (0,3,1)$$

Find $T(2, 3, -2)$.

Sol:

$$(2,3,-2) = 2(1,0,0) + 3(0,1,0) - 2(0,0,1)$$

$$T(2,3,-2) = 2T(1,0,0) + 3T(0,1,0) - 2T(0,0,1) \qquad (T \text{ is a L.T.})$$

$$= 2(2,-1,4) + 3(1,5,-2) - 2T(0,3,1)$$

$$= (7,7,0)$$

- Ex 5: (A linear transformation defined by a matrix)

The function $T: R^2 \rightarrow R^3$ is defined as $T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

(a) Find $T(\mathbf{v})$, where $\mathbf{v} = (2, -1)$

(b) Show that T is a linear transformation form R^2 into R^3

Sol: $(a)\ \mathbf{v} = (2, -1)$

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix} \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{R^2 \text{ vector}} = \underbrace{\begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix}}_{R^3 \text{ vector}}$$

$$\therefore T(2, -1) = (6, 3, 0)$$

$(b)\ T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v})$ (vector addition)

$T(c\mathbf{u}) = A(c\mathbf{u}) = c(A\mathbf{u}) = cT(\mathbf{u})$ (scalar multiplication)

- Thm 6.2: (The linear transformation given by a matrix)

  Let A be an $m \times n$ matrix. The function T defined by

  $$T(\mathbf{v}) = A\mathbf{v}$$

  is a linear transformation from R^n into R^m.

- Note:

$$A\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}}_{R^n \text{ vector}} = \underbrace{\begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}}_{R^m \text{ vector}}$$

$$T(\mathbf{v}) = A\mathbf{v}$$

$$T : R^n \longrightarrow R^m$$

- Ex 7: (Rotation in the plane)

Show that the L.T. $T: R^2 \to R^2$ given by the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

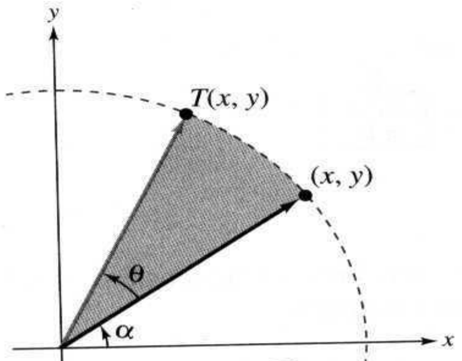
has the property that it rotates every vector in R^2 counterclockwise about the origin through the angle θ.

Sol:

$$v = (x, y) = (r\cos\alpha, r\sin\alpha) \qquad \text{(polar coordinates)}$$

r : the length of v

α : the angle from the positive x-axis counterclockwise to the vector v



Rotation in the Plane

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

$$= \begin{bmatrix} r\cos\theta\cos\alpha - r\sin\theta\sin\alpha \\ r\sin\theta\cos\alpha + r\cos\theta\sin\alpha \end{bmatrix}$$

$$= \begin{bmatrix} r\cos(\theta+\alpha) \\ r\sin(\theta+\alpha) \end{bmatrix}$$

r : the length of $T(\mathbf{v})$

$\theta+\alpha$: the angle from the positive x-axis counterclockwise to the vector $T(\mathbf{v})$

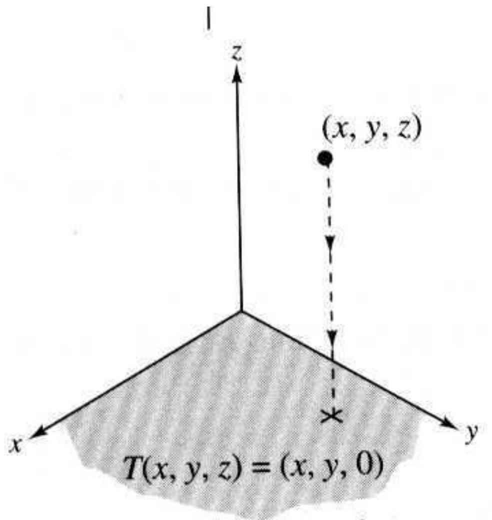
Thus, $T(\mathbf{v})$ is the vector that results from rotating the vector $\mathbf{v}$ counterclockwise through the angle θ.

- Ex 8: (A projection in R^3)

The linear transformation $T: R^3 \rightarrow R^3$ is given by

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

is called a projection in R^3.



Projection onto xy-plane

■Ex 9: (A linear transformation from $M_{m\times n}$ into $M_{n\times m}$)

$$T(A) = A^T \quad (T : M_{m\times n} \to M_{n\times m})$$

Show that T is a linear transformation.

Sol:

$$A, B \in M_{m\times n}$$

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

$$T(cA) = (cA)^T = cA^T = cT(A)$$

Therefore, T is a linear transformation from $M_{m\times n}$ into $M_{n\times m}$.

2. Matrices for Linear Transformations

- Two representations of the linear transformation $T: R^3 \to R^3$:

$$(1)\,T(x_1, x_2, x_3) = (2x_1 + x_2 - x_3, -x_1 + 3x_2 - 2x_3, 3x_2 + 4x_3)$$

$$(2)\,T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 3 & -2 \\ 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- Three reasons for matrix representation of a linear transformation:
  - It is simpler to write.
  - It is simpler to read.
  - It is more easily adapted for computer use.

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

■Thm 6.10: (Standard matrix for a linear transformation)

Let $T : R^n \rightarrow R^m$ be a linear trtansformation such that

$$T(e_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \cdots, \quad T(e_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix},$$

Then the $m \times n$ matrix whose n columns correspond to $T(e_i)$

$$A = \left[T(e_1) \mid T(e_2) \mid \cdots \mid T(e_n) \right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is such that $T(\mathbf{v}) = A\mathbf{v}$ for every $\mathbf{v}$ in R^n.

A is called the standard matrix for T.

Pf:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 e_1 + v_2 e_2 + \cdots + v_n e_n$$

$$T \text{ is a L.T.} \Rightarrow T(\mathbf{v}) = T(v_1 e_1 + v_2 e_2 + \cdots + v_n e_n)$$

$$= T(v_1 e_1) + T(v_2 e_2) + \cdots + T(v_n e_n)$$

$$= v_1 T(e_1) + v_2 T(e_2) + \cdots + v_n T(e_n)$$

$$A\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}$$

$$= v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= v_1 T(e_1) + v_2 T(e_2) + \cdots + v_n T(e_n)$$

Therefore, $T(\mathbf{v}) = A\mathbf{v}$ for each $\mathbf{v}$ in R^n

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

■ Ex 1: (Finding the standard matrix of a linear transformation)

Find the standard matrix for the L.T. $T : R^3 \to R^2$ define by

$$T(x, y, z) = (x - 2y, 2x + y)$$

Sol:

Vector Notation	Matrix Notation
$T(e_1) = T(1, 0, 0) = (1, 2)$	$T(e_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
$T(e_2) = T(0, 1, 0) = (-2, 1)$	$T(e_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$
$T(e_3) = T(0, 0, 1) = (0, 0)$	$T(e_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$A = [T(e_1) \mid T(e_2) \mid T(e_3)]$$
$$= \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

- Check:

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

$$\text{i.e.}\, T(x,y,z) = (x-2y, 2x+y)$$

- Note:

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{array}{l} \leftarrow \;\; 1x-2y+0z \\ \leftarrow \;\; 2x+1y+0z \end{array}$$

■Ex 2: (Finding the standard matrix of a linear transformation)

The linear transformation $T: R^2 \rightarrow R^2$ is given by projecting each point in R^2 onto the x-axis. Find the standard matrix for T.

Sol:

$$T(x, y) = (x, 0)$$

$$A = \left[T(e_1) \mid T(e_2)\right] = \left[T(1,0) \mid T(0,1)\right] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

- Notes:

(1) The standard matrix for the zero transformation from R^n into R^m is the $m \times n$ zero matrix.

(2) The standard matrix for the zero transformation from R^n into R^n is the $n \times n$ identity matrix I_n

■Ex 3: (The standard matrix of a composition)

Let T_1 and T_2 be L.T. from R^3 into R^3 s.t.

$$T_1(x, y, z) = (2x + y, 0, x + z)$$

$$T_2(x, y, z) = (x - y, z, y)$$

Find the standard matrices for the compositions

$T = T_2 \circ T_1$ and $T' = T_1 \circ T_2$,

Sol:

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad (\text{standard matrix for } T_1)$$

$$A_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (\text{standard matrix for } T_2)$$

The standard matrix for $T = T_2 \circ T_1$

$$A = A_2 A_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The standard matrix for $T' = T_1 \circ T_2$

$$A' = A_1 A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Inverse linear transformation:

If $T_1 : R^n \to R^n$ and $T_2 : R^n \to R^n$ are L.T. s.t. for every $\mathbf{v}$ in R^n

$$T_2(T_1(\mathbf{v})) = \mathbf{v} \quad \text{and} \quad T_1(T_2(\mathbf{v})) = \mathbf{v}$$

Then T_2 is called the inverse of T_1 and T_1 is said to be invertible

- Note:

If the transformation T is invertible, then the inverse is unique and denoted by T^{-1}.

■Ex 5: (Finding a matrix relative to nonstandard basis)

Let $T : R^2 \to R^2$ be a L.T. defined by

$$T(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2)$$

Find the matrix of T relative to the basis

$B = \{(1, 2), (-1, 1)\}$ and $B' = \{(1, 0), (0, 1)\}$

Sol:

$$T(1, 2) = (3, 0) = 3(1, 0) + 0(0, 1)$$

$$T(-1, 1) = (0, -3) = 0(1, 0) - 3(0, 1)$$

$$[T(1, 2)]_{B'} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad [T(-1, 1)]_{B'} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

the matrix for T relative to B and B'

$$A = \left[[T(1, 2)]_{B'} \quad [T(1, 2)]_{B'} \right] = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$$

■ Ex 6:

For the L.T. $T: R^2 \rightarrow R^2$ given in Example 5, use the matrix A to find $T(\mathbf{v})$, where $\mathbf{v} = (2,1)$

Sol:

$$\mathbf{v} = (2,1) = 1(1,2) - 1(-1,1) \qquad B = \{(1,2), (-1,1)\}$$

$$\Rightarrow [\mathbf{v}]_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\Rightarrow [T(\mathbf{v})]_{B'} = A[\mathbf{v}]_B = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$\Rightarrow T(\mathbf{v}) = 3(1,0) + 3(0,1) = (3,3) \qquad B' = \{(1,0), (0,1)\}$$

- Check:

$$T(2,1) = (2+1, 2(2)-1) = (3,3)$$

3. Similarity

■Ex 4: (Similar matrices)

$$(a)A=\begin{bmatrix}2 & -2\\ -1 & 3\end{bmatrix} \text{ and } A'=\begin{bmatrix}3 & -2\\ -1 & 2\end{bmatrix} \text{ are similar}$$

$$\text{because } A'=P^{-1}AP, \text{ where } P=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$$

$$(b)A=\begin{bmatrix}-2 & 7\\ -3 & 7\end{bmatrix} \text{ and } A'=\begin{bmatrix}2 & 1\\ -1 & 3\end{bmatrix} \text{ are similar}$$

$$\text{because } A'=P^{-1}AP, \text{ where } P=\begin{bmatrix}3 & -2\\ 2 & -1\end{bmatrix}$$

■Ex 5: (A comparison of two matrices for a linear transformation)

Suppose $A = \begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ is the matrix for $T: R^3 \rightarrow R^3$ relative to the standard basis. Find the matrix for T relative to the basis

$$B' = \{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$$

Sol:

The transition matrix from B' to the standard matrix

$$P = \left[[(1, 1, 0)]_B \quad [(1, -1, 0)]_B \quad [(0, 0, 1)]_B \right] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

matrix of T relative to B':

$$A' = P^{-1}AP = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

■Notes: Computational advantages of diagonal matrices:

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

$$(1) D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

$$(2) D^T = D$$

$$(3) D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}, \quad d_i \neq 0$$

SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL

Example : Consider two product reviews for a smartphone:

- Review 1: "The battery life is excellent, and the camera quality is fantastic."
- Review 2: "The battery life is terrible, and the camera quality is poor.“

Step 1: Preprocessing

- Tokenization: ["The", "battery", "life", "is", "excellent", "and", "the", "camera", "quality", "is", "fantastic"]
- Stop word Removal: ["battery", "life", "excellent", "camera", "quality", "fantastic"]
- Stemming/Lemmatization: ["batteri", "life", "excel", "camera", "qualiti", "fantast"]

SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL

Step 2: Building the Vocabulary

- Vocabulary: {"batteri": 0, "life": 1, "excel": 2, "camera": 3, "qualiti": 4, "fantast": 5}

Step 3: Vectorization (Bag-of-Words Representation)

- Review 1 vector: [1, 1, 1, 1, 1, 1] (One occurrence of each word in the vocabulary)
- Review 2 vector: [1, 1, 0, 1, 1, 0] (One occurrence of "battery," "life," "camera," and "quality")

SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL

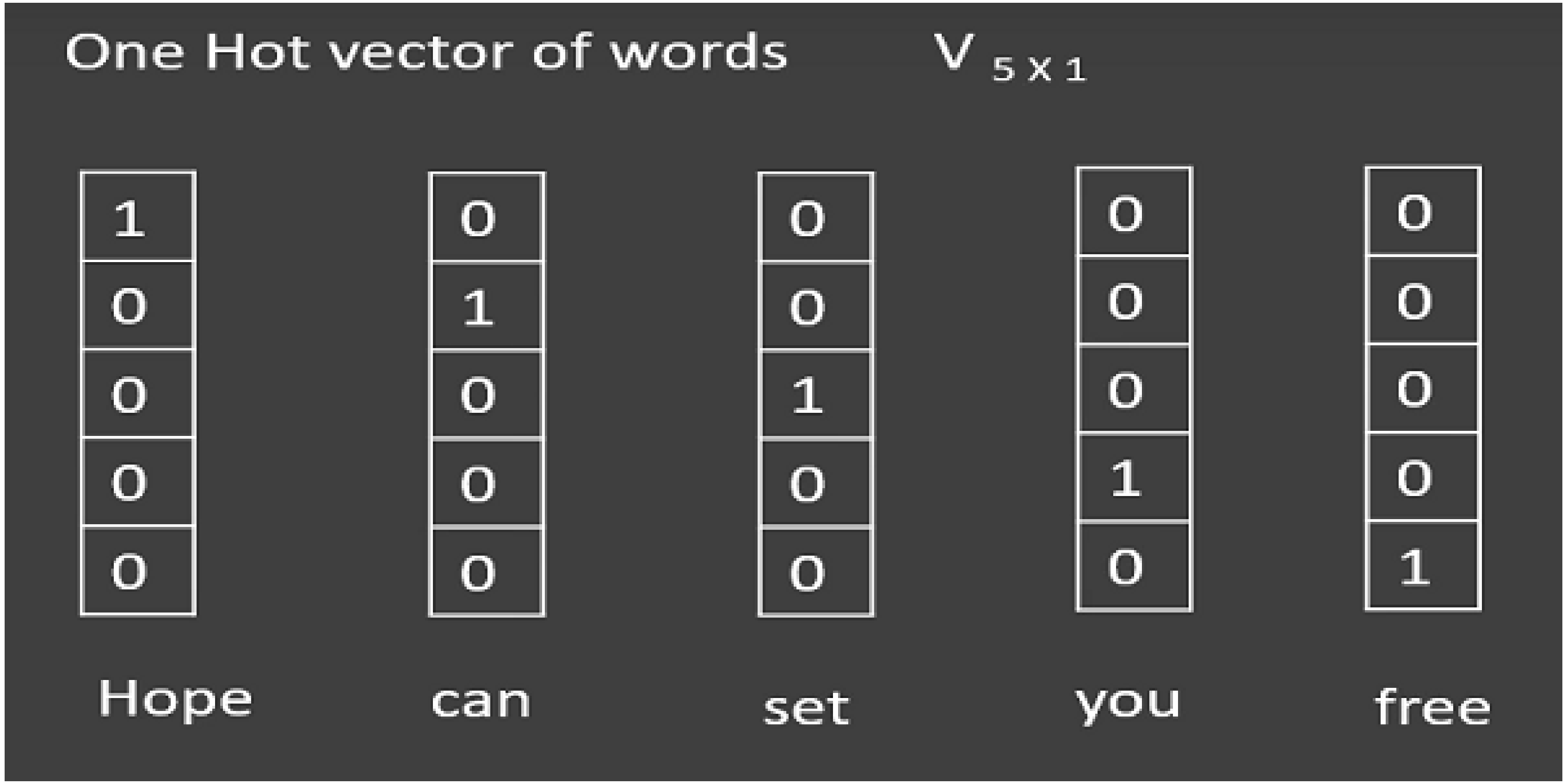
Step 2: Building the Vocabulary

- Vocabulary: {"batteri": 0, "life": 1, "excel": 2, "camera": 3, "qualiti": 4, "fantast": 5}

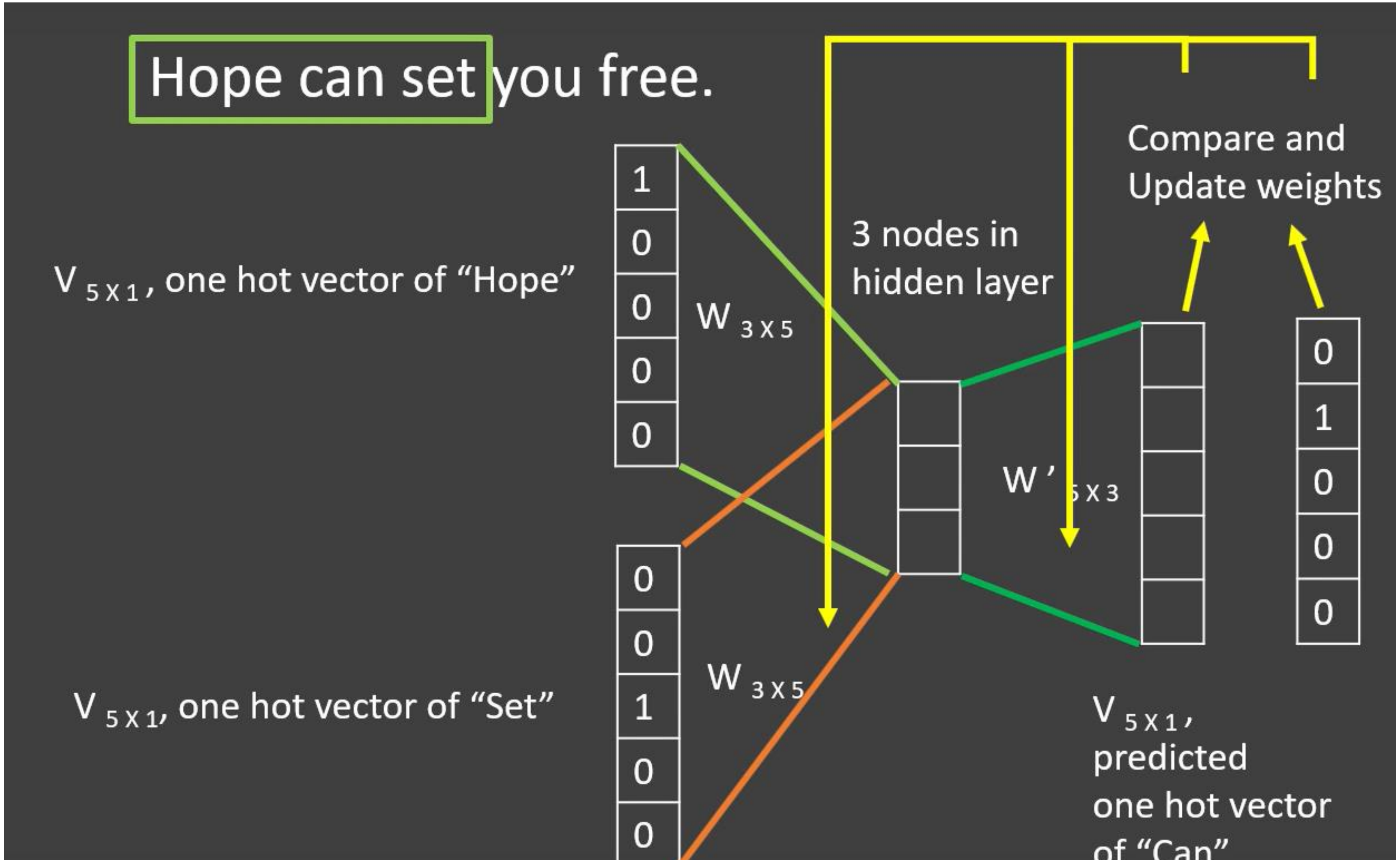
Step 3: Vectorization (Bag-of-Words Representation)

- Review 1 vector: [1, 1, 1, 1, 1, 1] (One occurrence of each word in the vocabulary)
- Review 2 vector: [1, 1, 0, 1, 1, 0] (One occurrence of "battery," "life," "camera," and "quality")

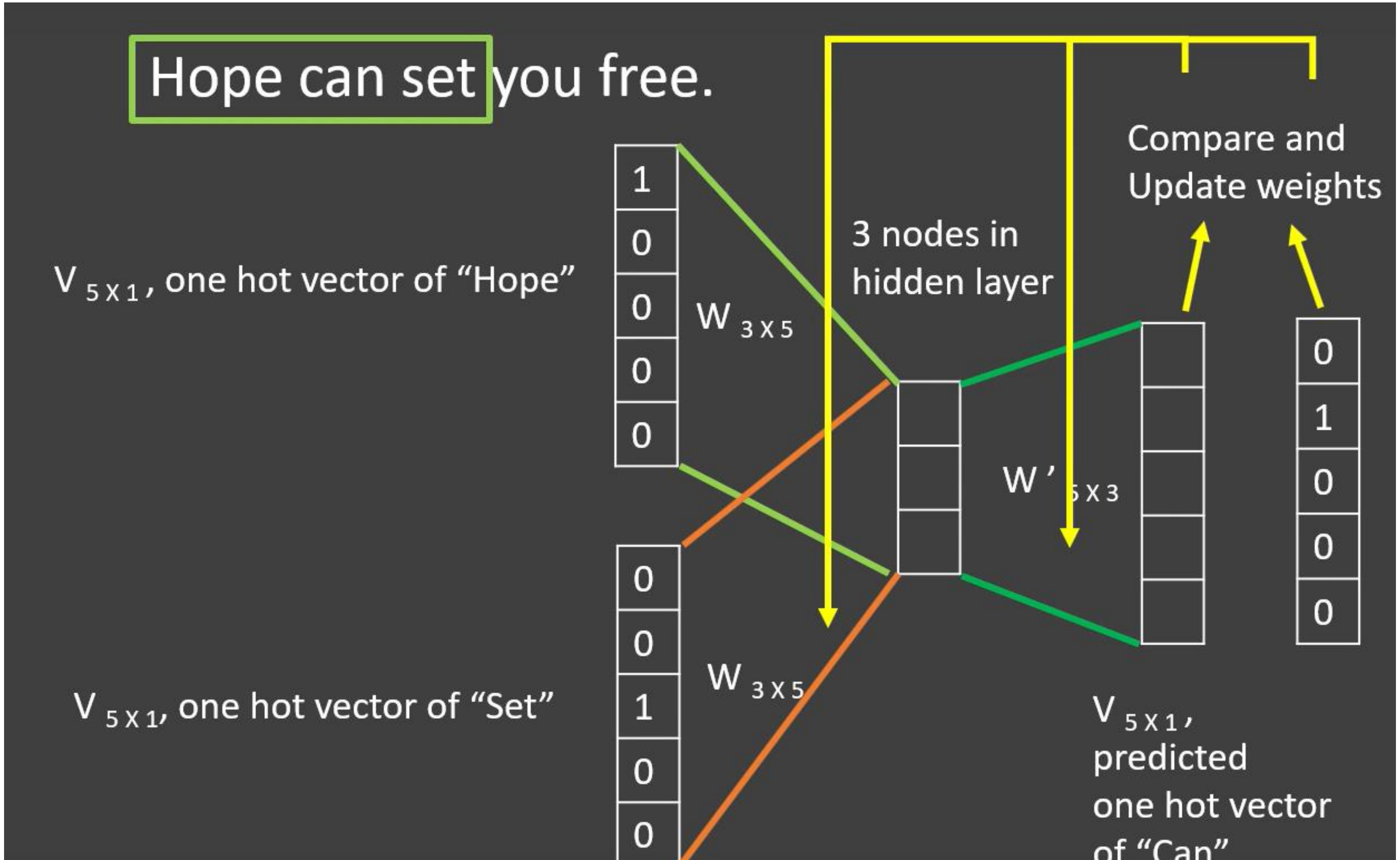
SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL



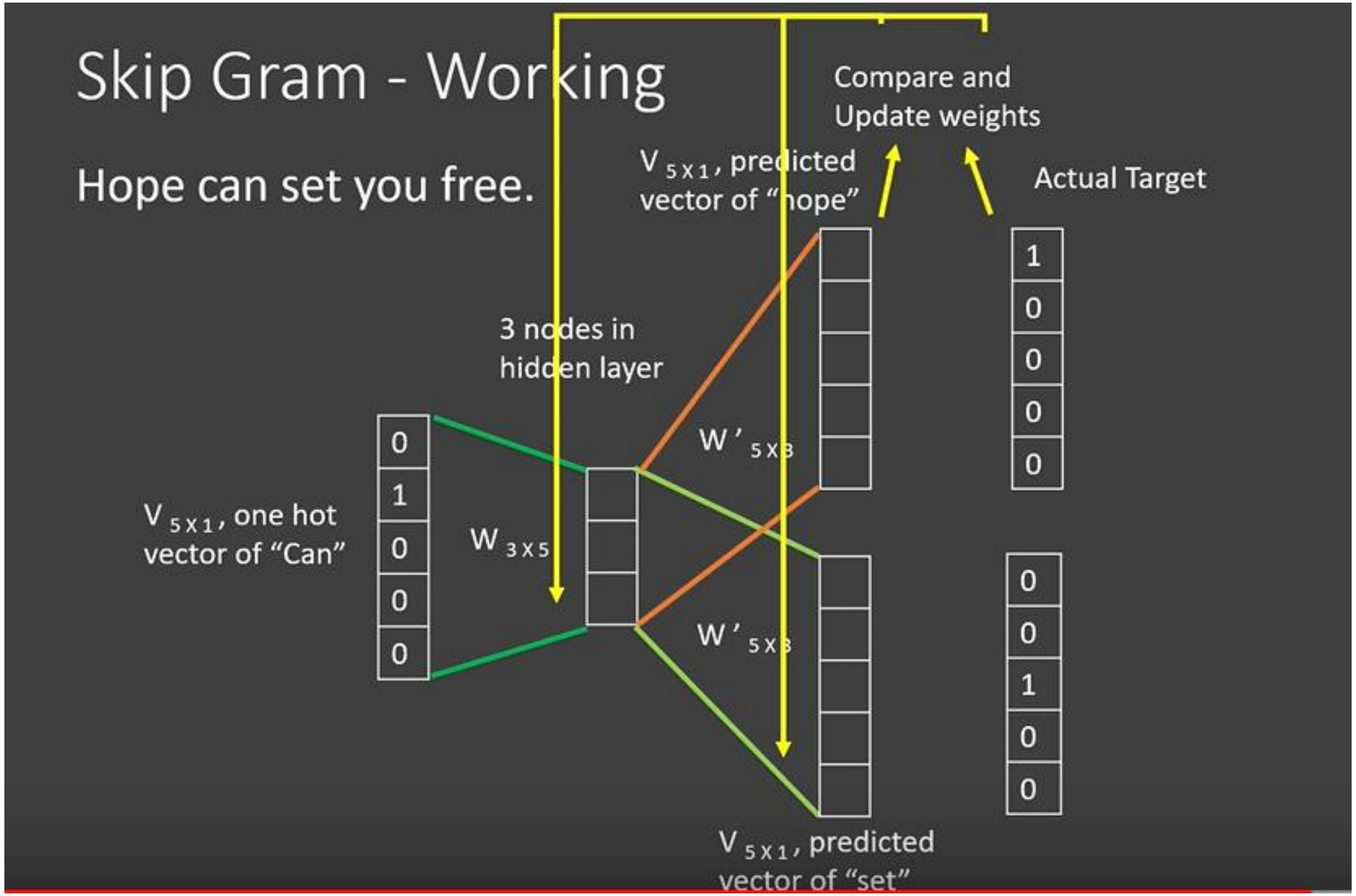
SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL



SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL



SENTIMENT ANALYSIS USING THE VECTOR SPACE MODEL



INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

THANK YOU