MTH205

LINEAR ALGEBRA II

Series 1 Systems of Linear Equations and Change of Basis

Series Outline

Introduction

Linear algebra provides powerful tools for solving problems that arise in mathematics, engineering, and the sciences. One of the most fundamental topics is the study of **systems of linear equations**, which appear in diverse contexts such as modelling networks, balancing chemical equations, and analysing economic systems. Closely related to this is the concept of **change of basis**, which allows us to view vectors and transformations from different perspectives, simplifying calculations and revealing deeper structure. This Series therefore sets the stage for much of what follows in the course.

The aim of this Series is to equip you with the ability to recognise, represent, and solve systems of linear equations using various methods, and to explain and apply the concept of change of basis in vector spaces.

We shall commence this Series by examining systems of linear equations, their representation, and methods of solution. Next, we will consider matrix representation and solution techniques, including augmented matrices, row operations, and the role of rank. Finally, we will wrap up by exploring the concept of change of basis, focusing on transition matrices, coordinate transformations, and their applications in simplifying problems.

- **SLO 1.1** define systems of linear equations and describe their representation in algebraic and matrix form,
- **SLO 1.2** demonstrate methods of solving systems of linear equations, including substitution, elimination, and matrix approaches,
- **SLO 1.3** analyse the consistency of systems and classify solutions as unique, infinite, or non-existent,
- **SLO 1.4** apply matrix representation techniques such as augmented matrices and row operations to solve linear systems,
- **SLO 1.5** explain the concept of rank and evaluate its role in determining solutions of systems of equations,
- **SLO 1.6** describe the concept of basis in vector spaces and explain how change of basis affects coordinate representation,
- **SLO 1.7** calculate transition matrices and apply them to perform coordinate transformations, and
- **SLO 1.8** evaluate practical applications of change of basis in simplifying linear algebra problems

1.1 Systems Of Linear Equations

1.1.1 Definition and representation of systems

A **system of linear equations** is a collection of two or more linear equations involving the same set of variables. Each equation represents a straight line in two dimensions or a plane in higher dimensions. The solution to the system is the set of values for the variables that satisfies all equations simultaneously.

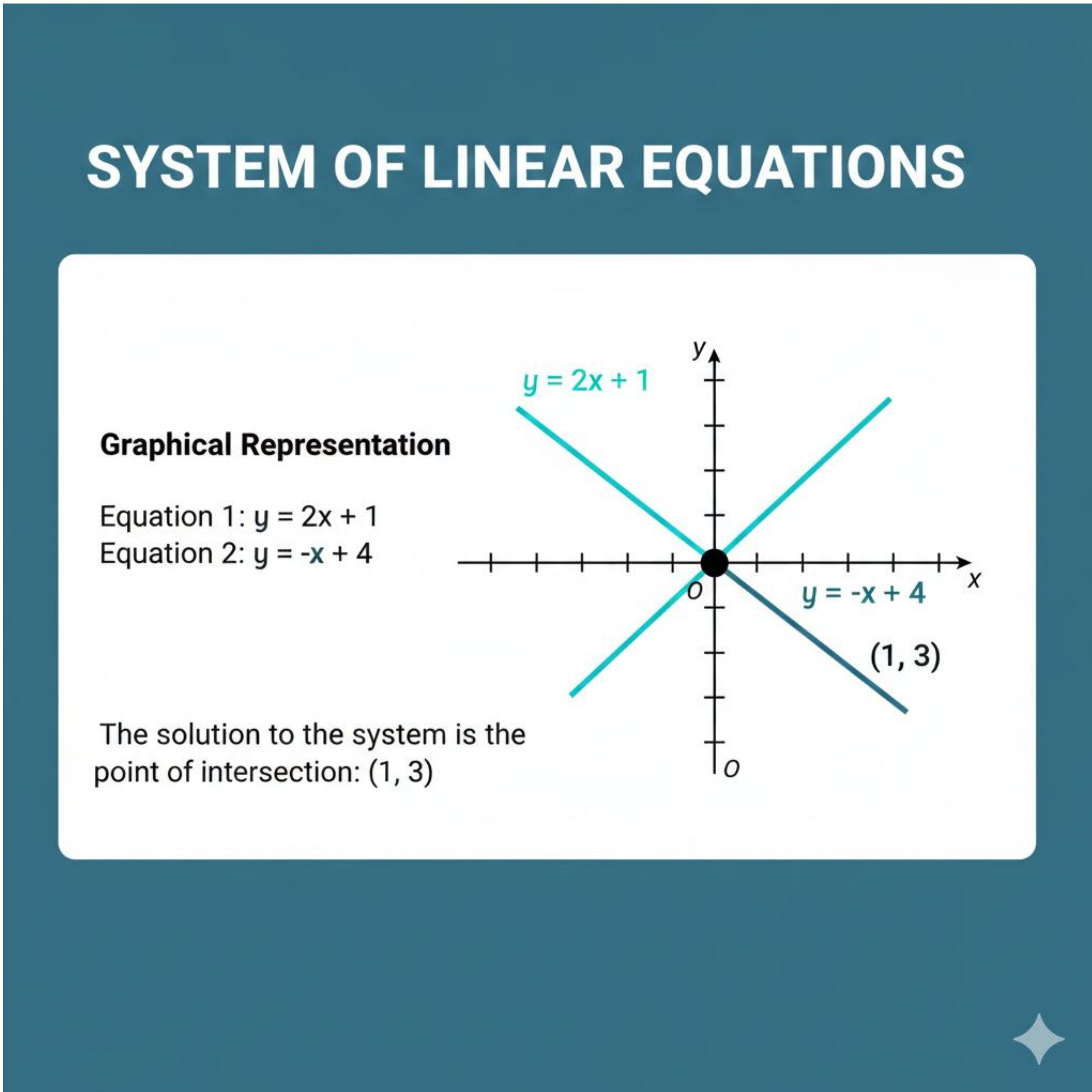
For example, the system:

$$x+y=5,2x-y=1$$

represents two lines in the plane. The point where they intersect is the solution to the system.

Matrix form of representation: Systems of equations can be expressed compactly using matrices. The coefficients form a **coefficient matrix**, the variables form a vector, and the constants form another vector. For the system above:

$$1\ 1\ 2\ -1\ x\ y = 5\ 1\ .$$



representation of a system of linear equations

Fill in the blank: A system of linear equations is solved by finding values of variables that satisfy ______ equations simultaneously.

1.1.2 Methods of solution

There are several methods for solving systems of linear equations:

Substitution method: Solve one equation for a variable and substitute into the other.

Elimination method: Add or subtract equations to eliminate a variable.

Matrix methods: Use augmented matrices and row operations to systematically solve.

Example (elimination method):

$$x+y=5, 2x-y=1$$

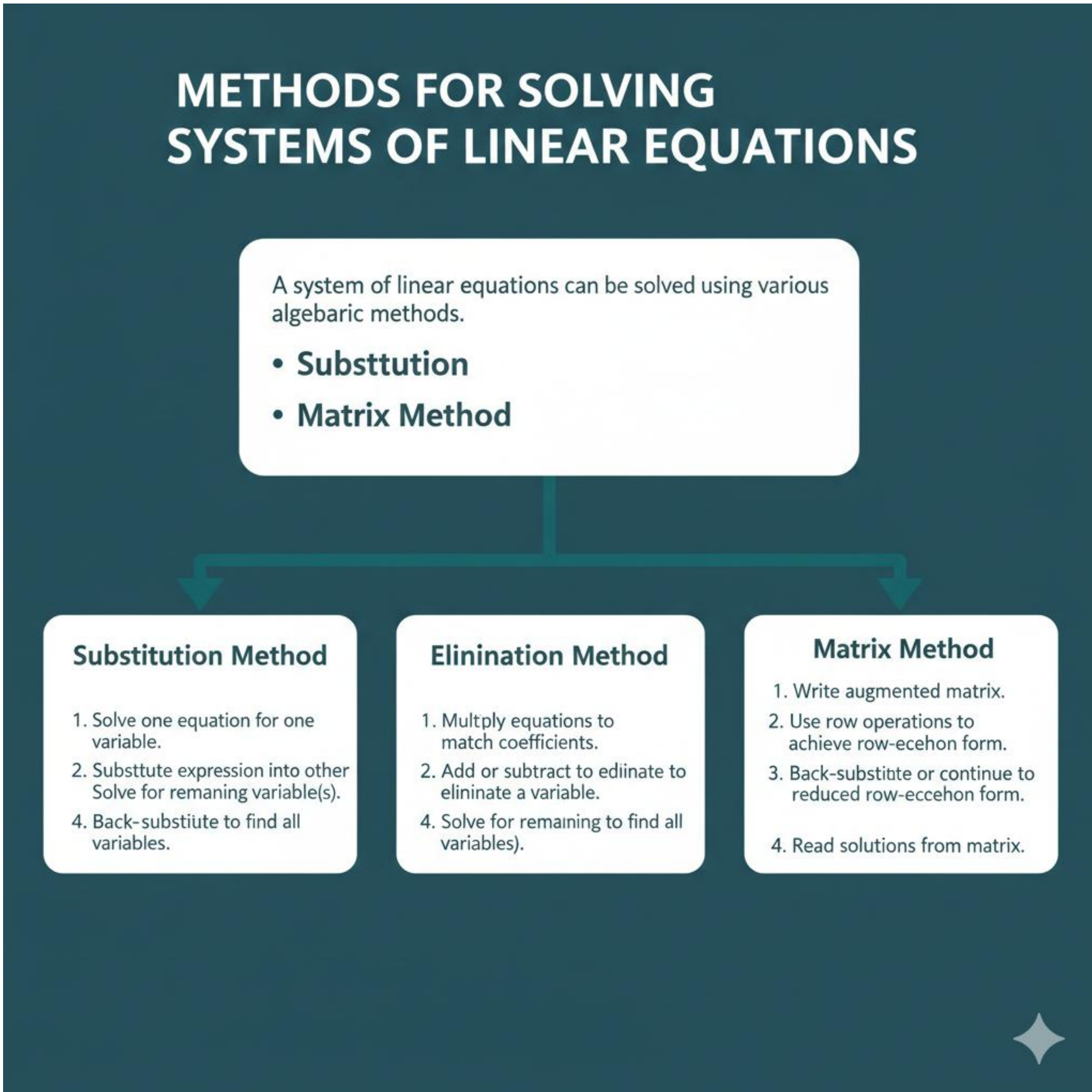
Add the two equations:

$$3x=6 \Rightarrow x=2.$$

Substitute into the first equation:

$$2+y=5 \Rightarrow y=3.$$

So, the solution is 3.



Methods of solving systems of linear equations

Fill in the blank: The elimination method works by adding or subtracting equations to remove a ______.

1.1.3 Consistency and types of solutions

A system may be:

Consistent with a unique solution: The equations intersect at a single point.

Consistent with infinitely many solutions: The equations represent the same line or plane.

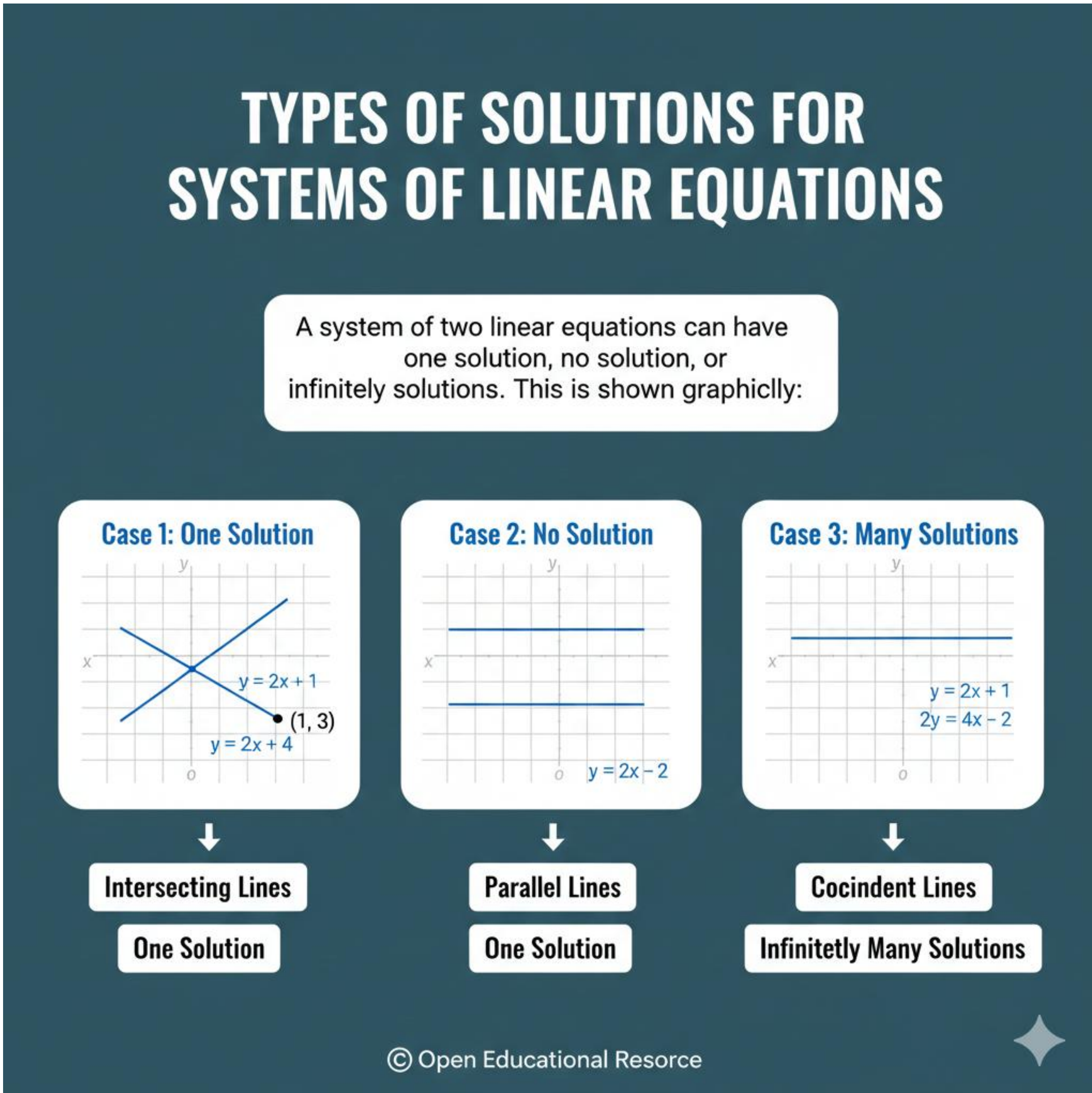
Inconsistent: The equations represent parallel lines or planes that never meet.

Example:

$x+y=5, 2x-y=1 \rightarrow$ unique solution.

$x+y=5, 2x+2y=10 \rightarrow$ infinitely many solutions.

$x+y=5, x+y=6 \rightarrow$ no solution.



Types of solutions in systems of linear equations

Fill in the blank: A system with no solution is called ______.

Pilot Questions 1.1

A system of linear equations can be represented compactly using a: A. Vector space B. Matrix C. Polynomial D. Determinant

Which method involves solving one equation for a variable and substituting into another? A. Elimination B. Substitution C. Matrix method D. Diagonalisation

A system with infinitely many solutions is described as: A. Inconsistent B. Consistent C. Singular D. Non-linear

The elimination method works by: A. Multiplying equations B. Adding or subtracting equations C. Changing basis D. Using determinants

If two equations represent parallel lines, the system is: A. Consistent with a unique solution B. Consistent with infinitely many solutions C. Inconsistent D. Diagonalised

1.2 Matrix Representation And Solution Techniques

1.2.1 Augmented matrices and row operations

An **augmented matrix** is a compact way of representing a system of linear equations by including both the coefficients of the variables and the constants from the equations. This allows us to apply systematic procedures known as **row operations** to simplify and solve the system.

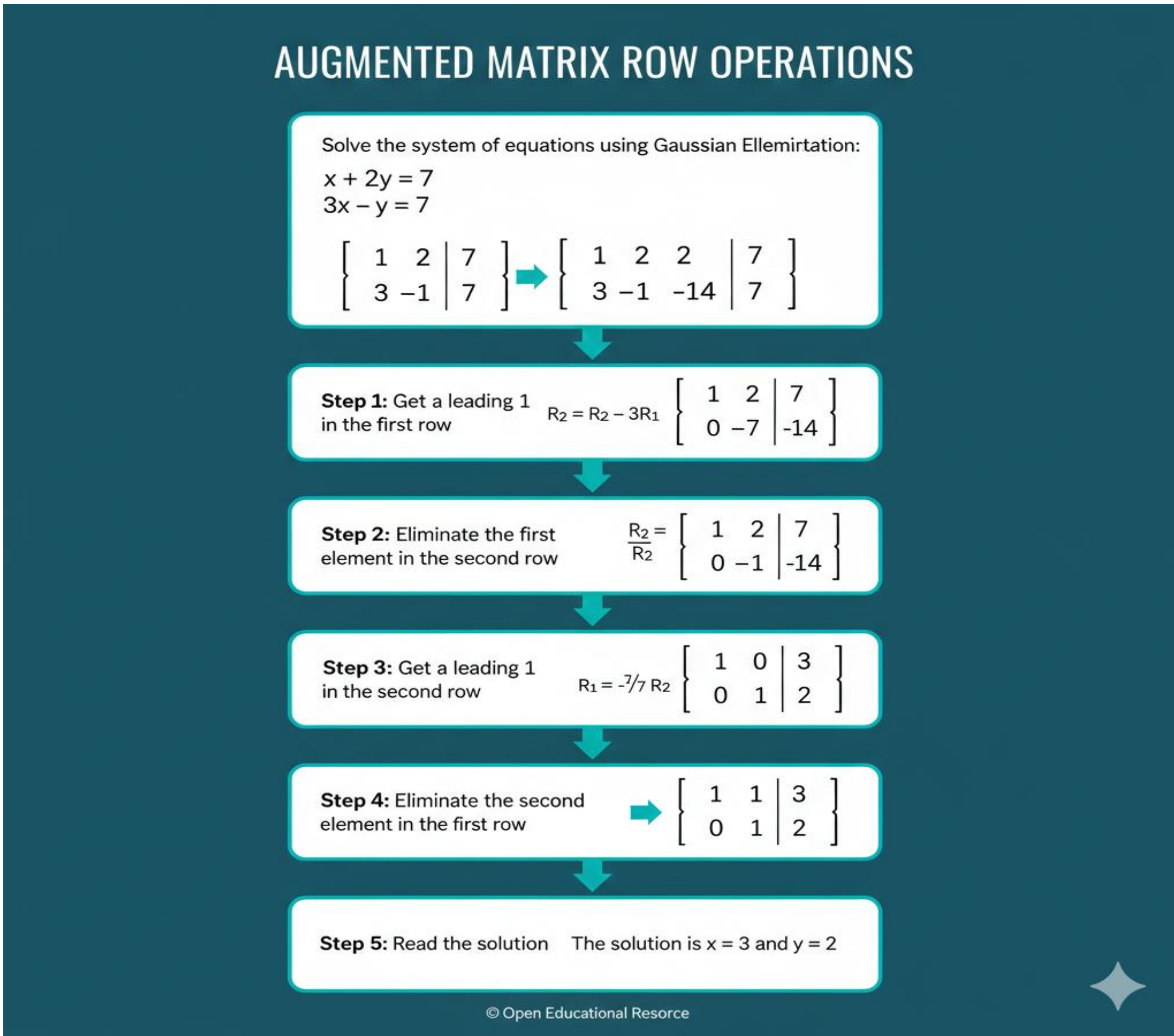
Row operations include:

Swapping two rows,

Multiplying a row by a non-zero constant,

Adding or subtracting a multiple of one row to another.

These operations do not change the solution set of the system but help transform the matrix into a simpler form, often leading to the solution.



Augmented matrix and row operations

Fill in the blank: Multiplying a row by a non-zero constant is an example of a ______ operation.

1.2.2 Rank of a matrix and its role in solving systems

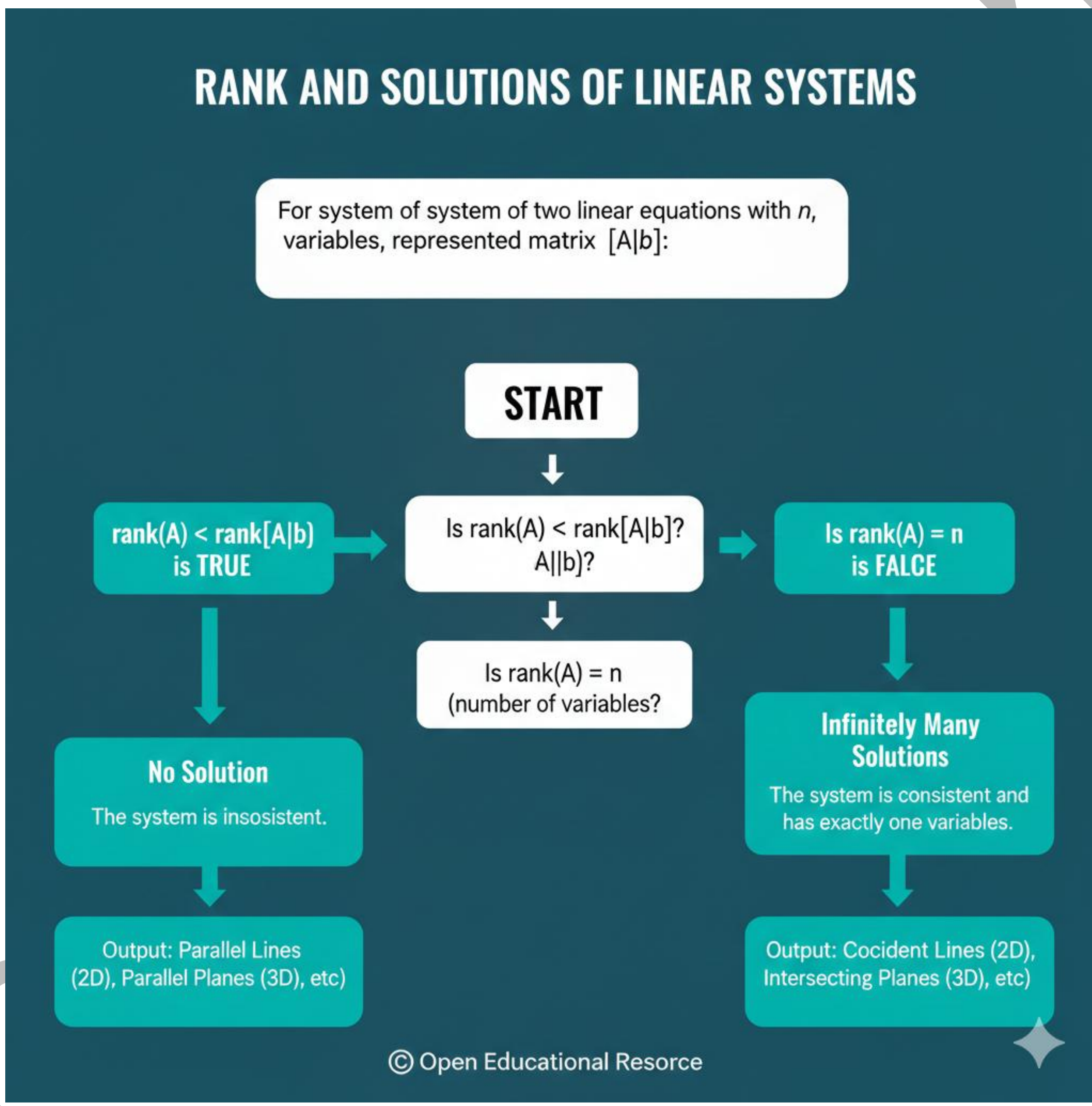
The **rank** of a matrix is the number of linearly independent rows or columns it contains. Rank plays a crucial role in determining whether a system of equations has solutions.

If the rank of the coefficient matrix equals the rank of the augmented matrix, the system is consistent.

If the rank of the coefficient matrix is less than the rank of the augmented matrix, the system is inconsistent.

If the system is consistent and the rank equals the number of variables, there is a unique solution.

If the system is consistent and the rank is less than the number of variables, there are infinitely many solutions.



Rank conditions for solutions of systems of equations

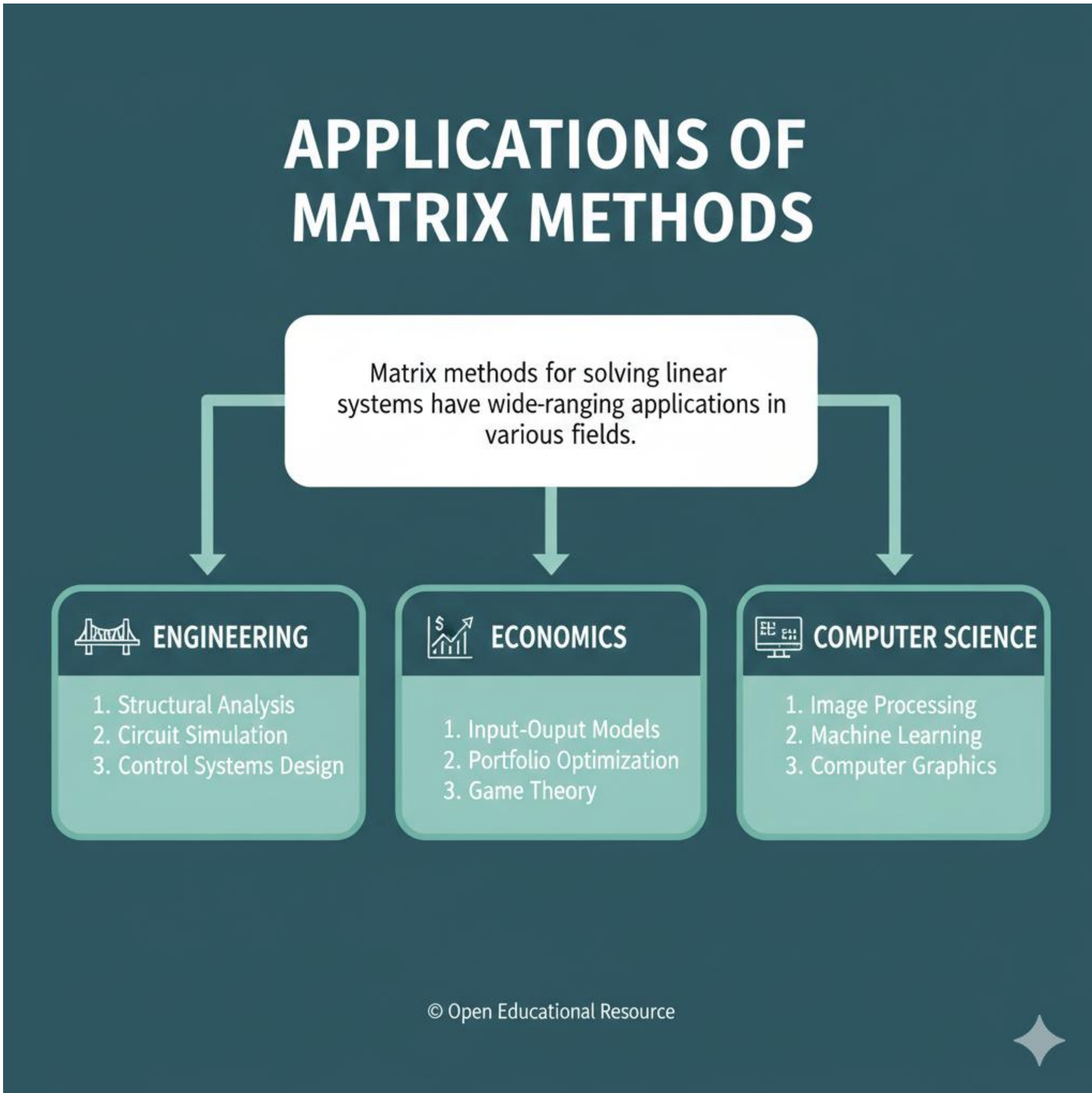
Fill in the blank: A system is consistent if the rank of the coefficient matrix is equal to the rank of the ______ matrix.

1.2.3 Applications of matrix methods in linear systems

Matrix methods are widely used in practical applications because they provide efficient ways to handle large systems. For example:

- In engineering, matrices are used to model electrical networks.
- In economics, they are used to analyse input-output models.
- In computer science, they are applied in graphics transformations and algorithms.

By representing systems in matrix form, we can use computational tools to solve problems that would otherwise be too complex to handle manually.



Applications of matrix methods in solving linear systems

Fill in the blank: Matrix methods are especially useful in solving ______ systems that are too complex for manual calculation.

Pilot Questions 1.2

An augmented matrix includes: A. Only coefficients B. Only constants C. Both coefficients and constants D. Only variables

Which of the following is a valid row operation? A. Multiplying a row by zero B. Swapping two rows C. Adding a variable to a row D. Dividing a row by zero

A system is inconsistent if: A. Rank of coefficient matrix = rank of augmented matrix B. Rank of coefficient matrix < rank of augmented matrix C. Rank of coefficient matrix > number of variables D. Rank of augmented matrix = 0

If the rank equals the number of variables, the system has: A. No solution B. Infinitely many solutions C. A unique solution D. A trivial solution only

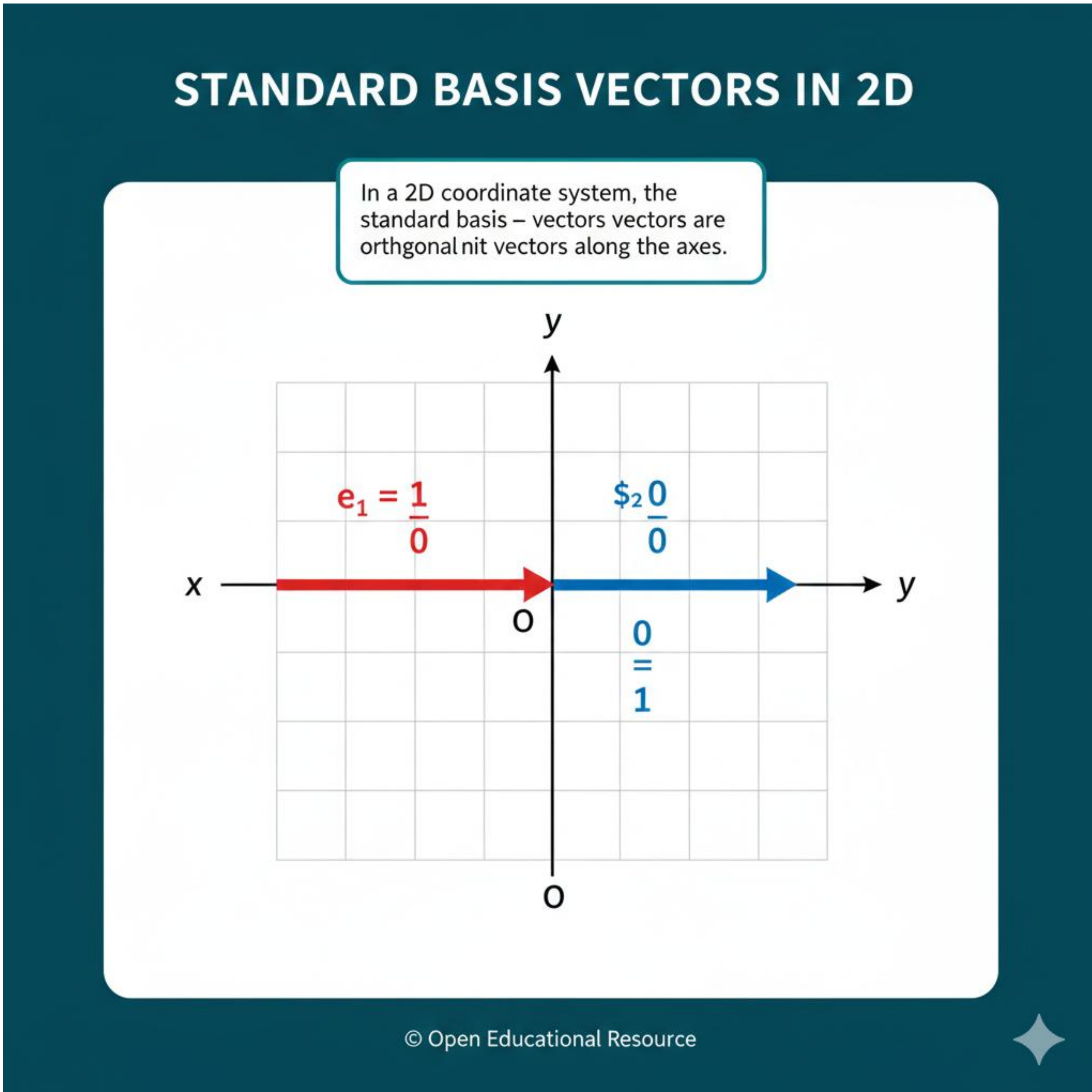
Matrix methods are particularly useful in: A. Solving small systems manually B. Solving large systems computationally C. Avoiding row operations D. Eliminating constants

1.3 Change Of Basis

1.3.1 Concept of basis in vector spaces

A **basis** is a set of linearly independent vectors that span a vector space. This means every vector in the space can be expressed uniquely as a linear combination of the basis vectors. The choice of basis determines how vectors are represented in coordinate form.

For example, in R2, the standard basis is (1,0),(0,1). Any vector ycan be written as x(1,0)+y(0,1). However, other sets of linearly independent vectors can also serve as bases.



Standard basis vectors in R2

Fill in the blank: A basis is a set of linearly independent vectors that _______ a vector space.

1.3.2 Transition matrices and coordinate transformations

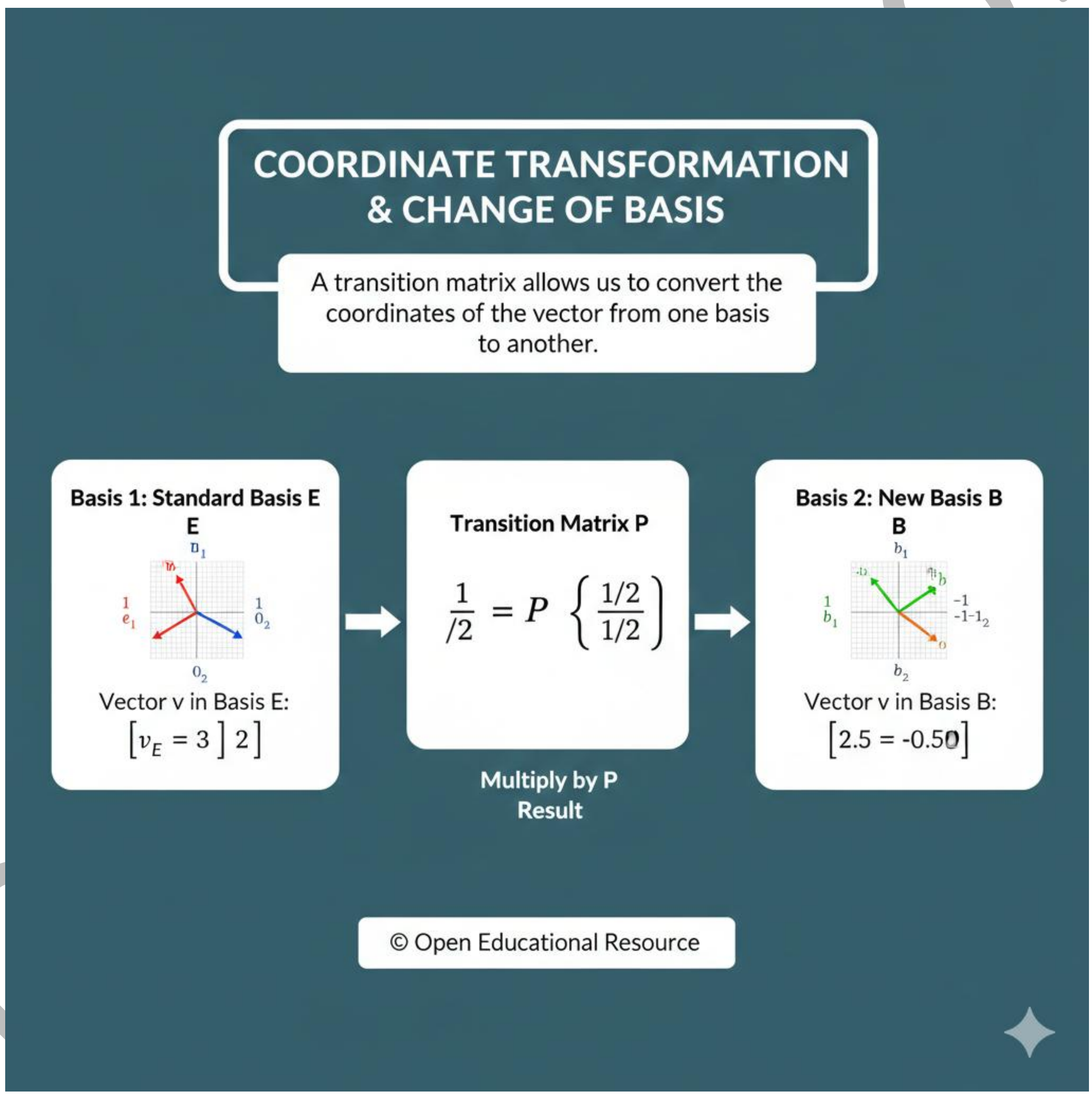
A **transition matrix** is a matrix that converts coordinates of vectors from one basis to another. Suppose we have two bases Band C. The transition matrix from Bto

Callows us to express coordinates relative to Cwhen they are originally given relative to B.

If Pis the transition matrix, then:

$$[x]C=P[x]B,$$

where [x]Band [x]Care coordinate vectors of xin bases Band Crespectively.



Transition matrix converting coordinates between bases

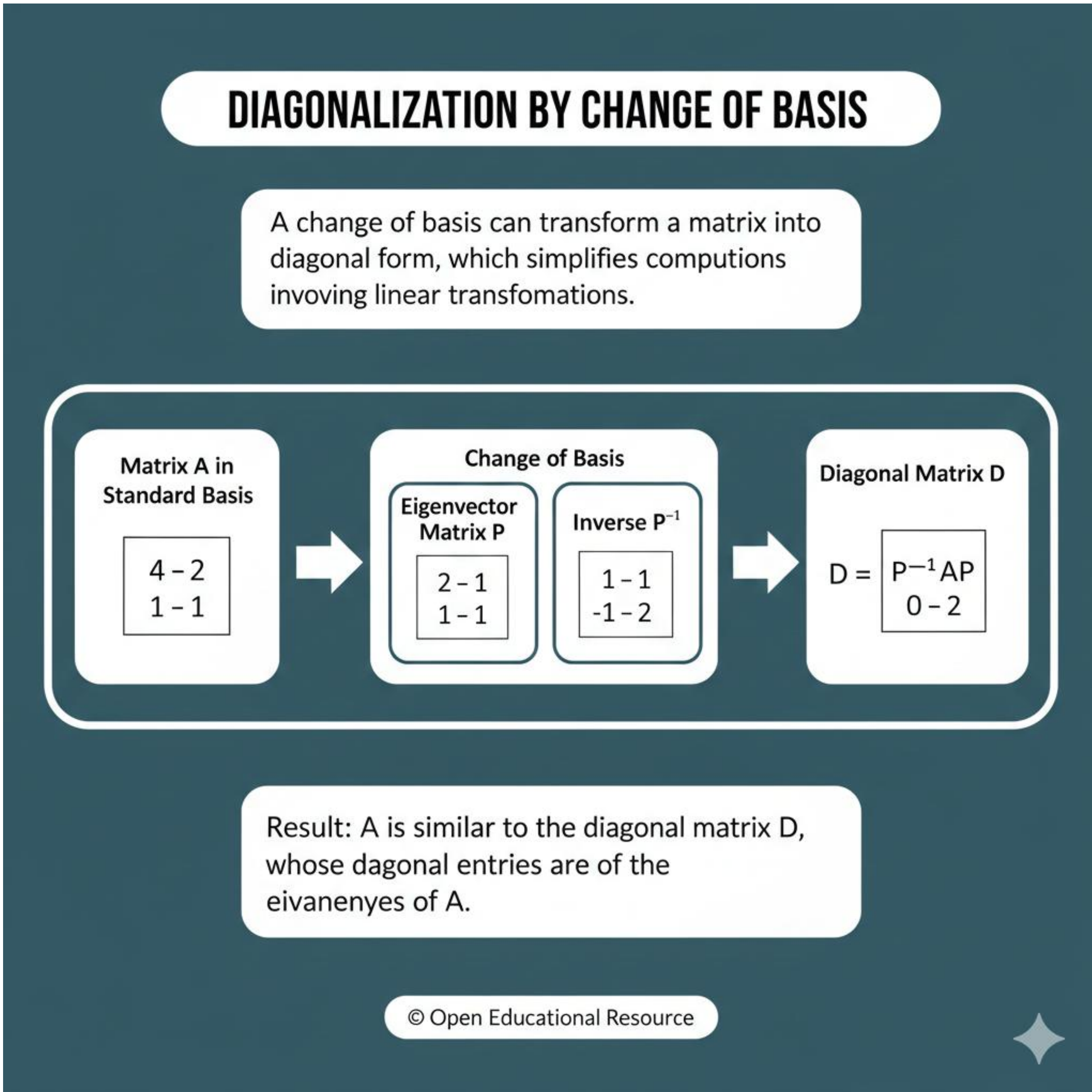
Fill in the blank: A transition matrix converts coordinates of vectors from one ______ to another.

1.3.3 Applications of change of basis in simplifying problems

Change of basis is used to simplify problems in linear algebra and applied mathematics. For example:

- In computer graphics, changing basis allows transformations to be applied more efficiently.
- In physics, it helps express vectors relative to convenient coordinate systems.
- In linear algebra, it is essential for diagonalising matrices and simplifying quadratic forms.

By choosing an appropriate basis, complex problems can often be reduced to simpler forms.



Application of change of basis in diagonalisation

Fill in the blank: Change of basis is often used to simplify problems by choosing a more ______ coordinate system.

Pilot Questions 1.3

A basis is defined as: A. A set of dependent vectors B. A set of linearly independent vectors that span a space C. A set of constants D. A set of equations

The matrix that converts coordinates from one basis to another is called: A. Coefficient matrix B. Transition matrix C. Augmented matrix D. Rank matrix

If a system uses the standard basis 1,0),(0,1, the vector 4can be expressed as: A. 3(0,1)+4(1,0)B. 3(1,0)+4(0,1)C. 3(1,1)+4(0,0)D. None of the above

Change of basis is useful in: A. Making problems more complex B. Simplifying problems and transformations C. Eliminating solutions D. Avoiding diagonalisation

In physics, change of basis allows vectors to be expressed relative to: A. Arbitrary constants B. Convenient coordinate systems C. Augmented matrices D. Polynomial equations

Series Summary

This Series introduced the fundamental concept of **systems of linear equations**, their representation, and methods of solution. We explored how systems can be expressed using matrices, and how augmented matrices and row operations provide systematic ways of solving them. The role of **rank** in determining consistency and the type of solutions was emphasised. Finally, we examined the concept of **change of basis**, transition matrices, and their applications in simplifying problems across mathematics, physics, and computer science.

Glossary of Terms

Augmented matrix: A matrix that represents a system of linear equations by including both coefficients and constants.

Basis: A set of linearly independent vectors that span a vector space, allowing every vector to be expressed uniquely as a linear combination of them.

Change of basis: The process of expressing vectors in terms of a different basis, often to simplify calculations.

Coefficient matrix: A matrix containing only the coefficients of variables in a system of linear equations.

Consistency: A property of a system of equations that indicates whether it has at least one solution.

Rank: The number of linearly independent rows or columns in a matrix, used to determine the solvability of a system.

System of linear equations: A set of two or more linear equations involving the same variables.

Transition matrix: A matrix that converts coordinates of vectors from one basis to another.

Concept Check Questions 1

Objective questions

The augmented matrix of a system of equations includes: A. Only coefficients B. Only constants C. Both coefficients and constants D. Only variables (Linked to SLO 1.4)

The number of linearly independent rows in a matrix is called its: A. Dimension B. Rank C. Basis D. Order (Linked to SLO 1.5)

Short-answer questions

Define a system of linear equations and give one real-life example. (Linked to SLO 1.1)

State the difference between a consistent and an inconsistent system. (Linked to SLO 1.3)

Long-answer questions

Explain the concept of change of basis and demonstrate how a transition matrix is used to transform coordinates. (Linked to SLO 1.6, SLO 1.7)

Basic response: Describe basis, change of basis, and show a simple transformation.

Value-added response: Discuss applications such as simplifying matrix representation or diagonalisation.

Answers to Concept Check Questions 1

C. Both coefficients and constants

B. Rank

A system of linear equations is a set of equations involving the same variables, for example, equations used to balance chemical reactions.

A consistent system has at least one solution, while an inconsistent system has no solution.

Basic response: Change of basis means expressing vectors in terms of a new basis. A transition matrix is used to convert coordinates from the old basis to the new one.

Value-added response: Beyond simple transformations, change of basis is applied in simplifying linear transformations, diagonalising matrices, and making computations more efficient in applied mathematics.

Answers to Pilot Questions 1

From Part 1.1

Matrix

Substitution

Consistent

Adding or subtracting equations

Inconsistent

From Part 1.2

Both coefficients and constants

Swapping two rows

Rank of coefficient matrix < rank of augmented matrix

A unique solution

Solving large systems computationally

From Part 1.3

A set of linearly independent vectors that span a space

Transition matrix

$3(1,0)+4(0,1)$

Simplifying problems and transformations

Convenient coordinate systems

2 Equivalence, Similarity, And Canonical Forms

Series Outline

For this 1-credit course, each Series should contain 2–4 Parts. Here is a balanced outline for Series 2:

- **Part 2.1 Equivalence of matrices and linear transformations**
  - **2.1.1 Definition of equivalence**
  - **2.1.2 Elementary transformations and equivalence classes**
  - **2.1.3 Applications of equivalence**
- **Part 2.2 Similarity of matrices**
  - **2.2.1 Definition of similarity**
  - **2.2.2 Properties of similar matrices**
  - **2.2.3 Examples and applications**
- **Part 2.3 Canonical forms**
  - **2.3.1 Concept of canonical forms**
  - **2.3.2 Diagonal form and Jordan form**
  - **2.3.3 Applications of canonical forms in simplifying problems**

Introduction

Linear algebra often requires us to compare matrices and transformations to determine whether they share essential properties. Two important concepts in this regard are equivalence and similarity, which allow us to classify matrices into groups with common features. These classifications lead naturally to the study of canonical forms, which provide simplified representations of matrices that are easier to analyse and apply.

The aim of this Series is to enable you to define and explain equivalence and similarity of matrices, and to apply canonical forms in solving and simplifying linear algebra problems.

We shall commence this Series by examining the concept of equivalence, including definitions, elementary transformations, and applications. Next, we

will move on to similarity of matrices, exploring its properties and practical uses. Finally, we will conclude with canonical forms, focusing on diagonal and Jordan forms, and their role in simplifying complex problems.

Series Learning Outcomes

Upon completing this Series, you should be able to:

SLO 2.1 define equivalence of matrices and explain how elementary transformations establish equivalence classes,

SLO 2.2 demonstrate how to apply equivalence in simplifying linear transformations,

SLO 2.3 define similarity of matrices and describe its properties,

SLO 2.4 analyse examples of similar matrices and explain their applications,

SLO 2.5 describe the concept of canonical forms and explain their role in matrix classification,

SLO 2.6 calculate diagonal and Jordan canonical forms for given matrices, and

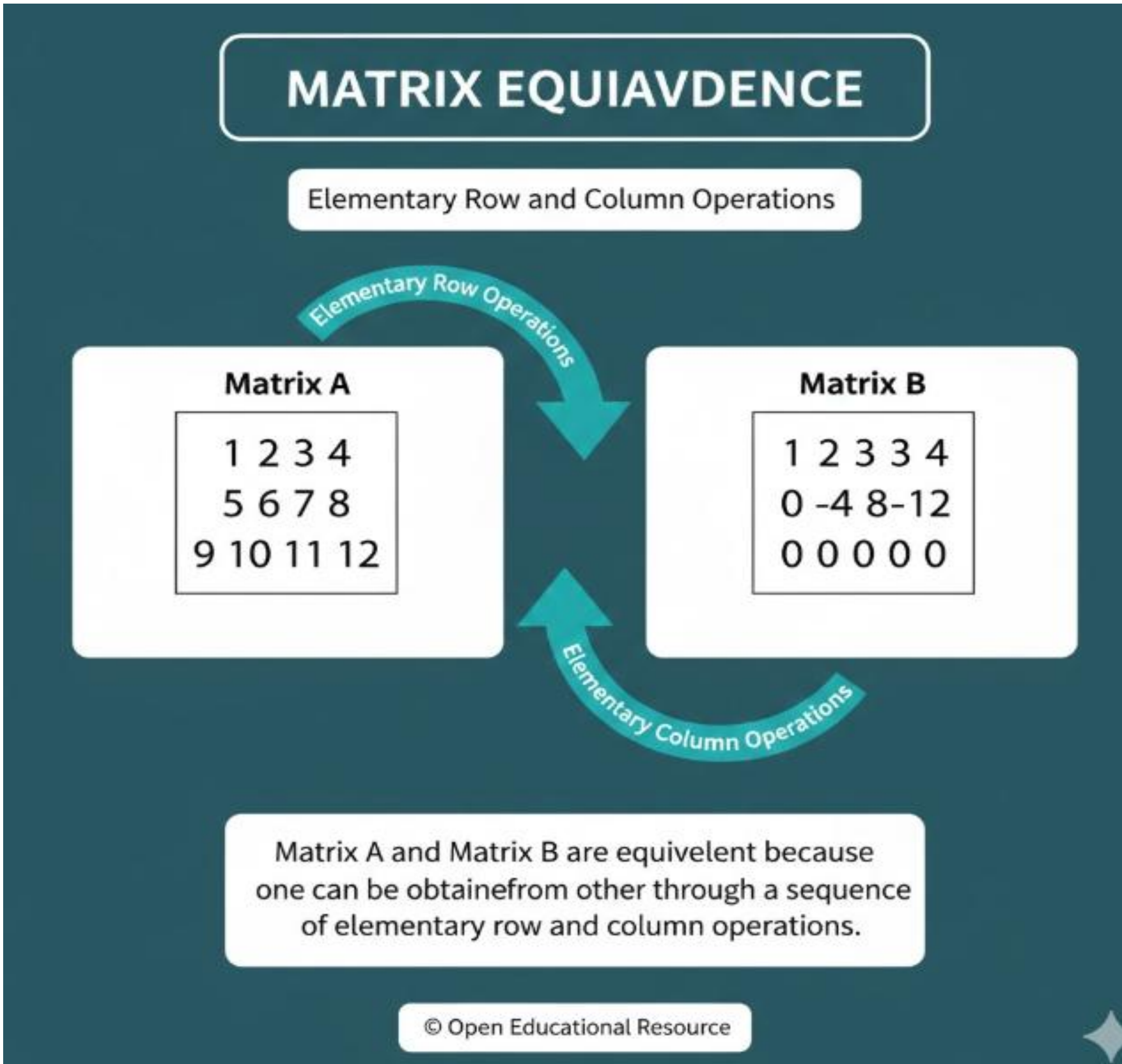
SLO 2.7 evaluate the usefulness of canonical forms in simplifying linear algebra problems.

2.1 Equivalence Of Matrices And Linear Transformations

2.1.1 Definition of equivalence

Two matrices are said to be **equivalent** if one can be transformed into the other using a finite sequence of **elementary row and column operations**. These operations include swapping rows or columns, multiplying a row or column by a non-zero constant, and adding a multiple of one row or column to another.

Equivalence is a way of classifying matrices into groups called **equivalence classes**, where all matrices in the same class share certain structural properties. This concept is particularly useful in simplifying linear transformations and understanding the relationships between different matrix representations.



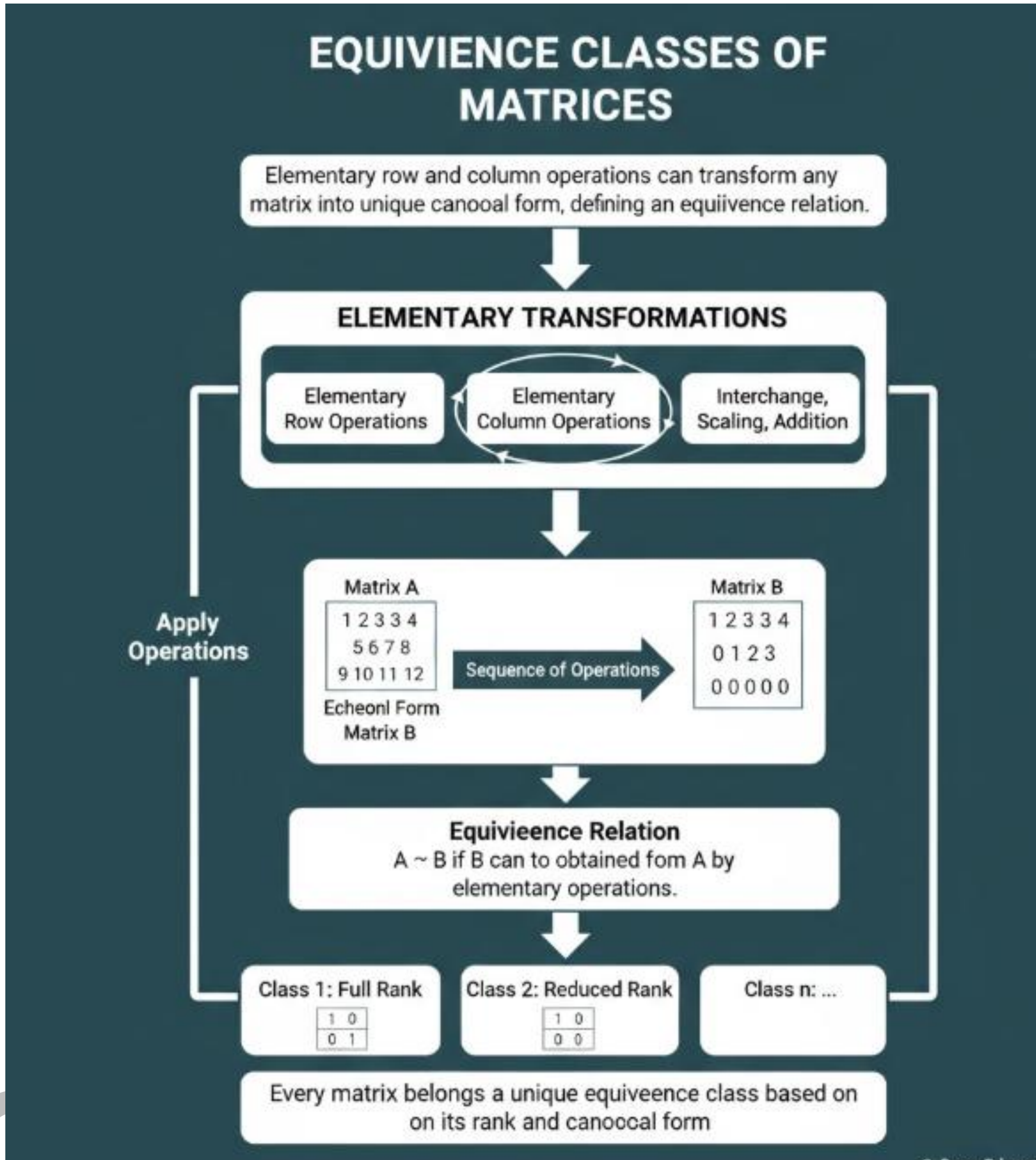
Equivalence of matrices through elementary operations

Fill in the blank: Two matrices are equivalent if one can be transformed into the other using a sequence of ______ operations.

2.1.2 Elementary transformations and equivalence classes

Elementary transformations are the basic operations that preserve equivalence. They allow us to reduce a matrix to a simpler form without changing its equivalence class.

Equivalence classes group matrices that can be transformed into one another. Within each class, matrices share the same rank, which is a key invariant under equivalence. This means that equivalence provides a way to classify matrices based on rank.



transformations and equivalence classes

Fill in the blank: Matrices in the same equivalence class share the same ______.

2.1.3 Applications of equivalence

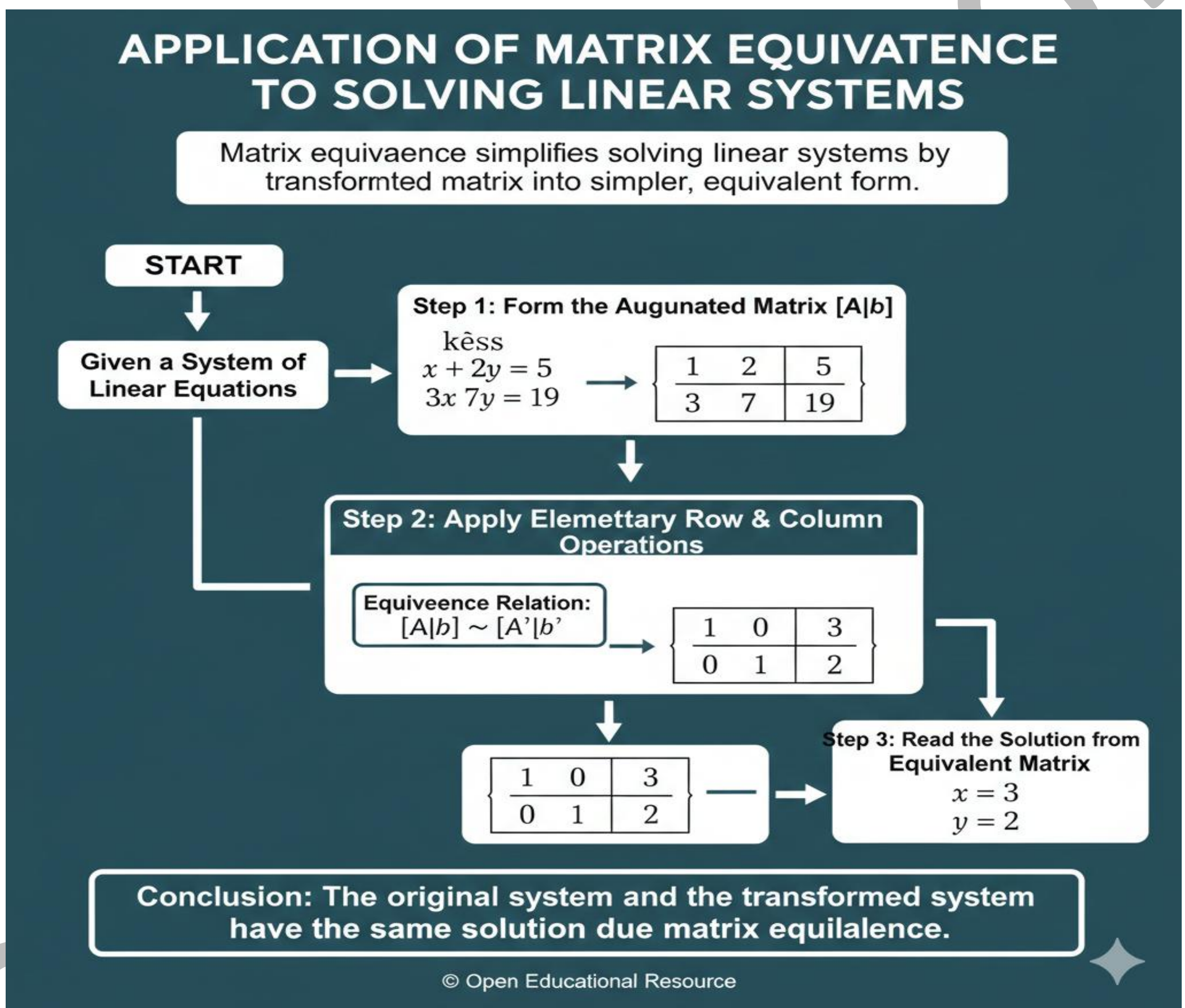
Equivalence is applied in several areas of linear algebra:

Simplifying linear transformations by reducing matrices to simpler forms.

Classifying matrices according to rank.

Providing insight into the structure of vector spaces and their transformations.

For example, in solving systems of equations, equivalence allows us to reduce the augmented matrix to a simpler form that reveals the solution set.



Application of equivalence in solving systems of equations

Fill in the blank: Equivalence is often used to simplify ______ transformations.

Pilot Questions 2.1

Two matrices are equivalent if: A. They have the same determinant B. They can be transformed into one another using elementary operations C. They are both diagonal D. They have the same eigenvalues

Which of the following is an elementary transformation? A. Multiplying a row by zero B. Swapping two rows C. Adding a variable to a row D. Dividing a row by zero

Matrices in the same equivalence class share the same: A. Determinant B. Rank C. Eigenvalues D. Trace

Equivalence is useful in: A. Making matrices more complex B. Simplifying linear transformations C. Eliminating solutions D. Avoiding classification

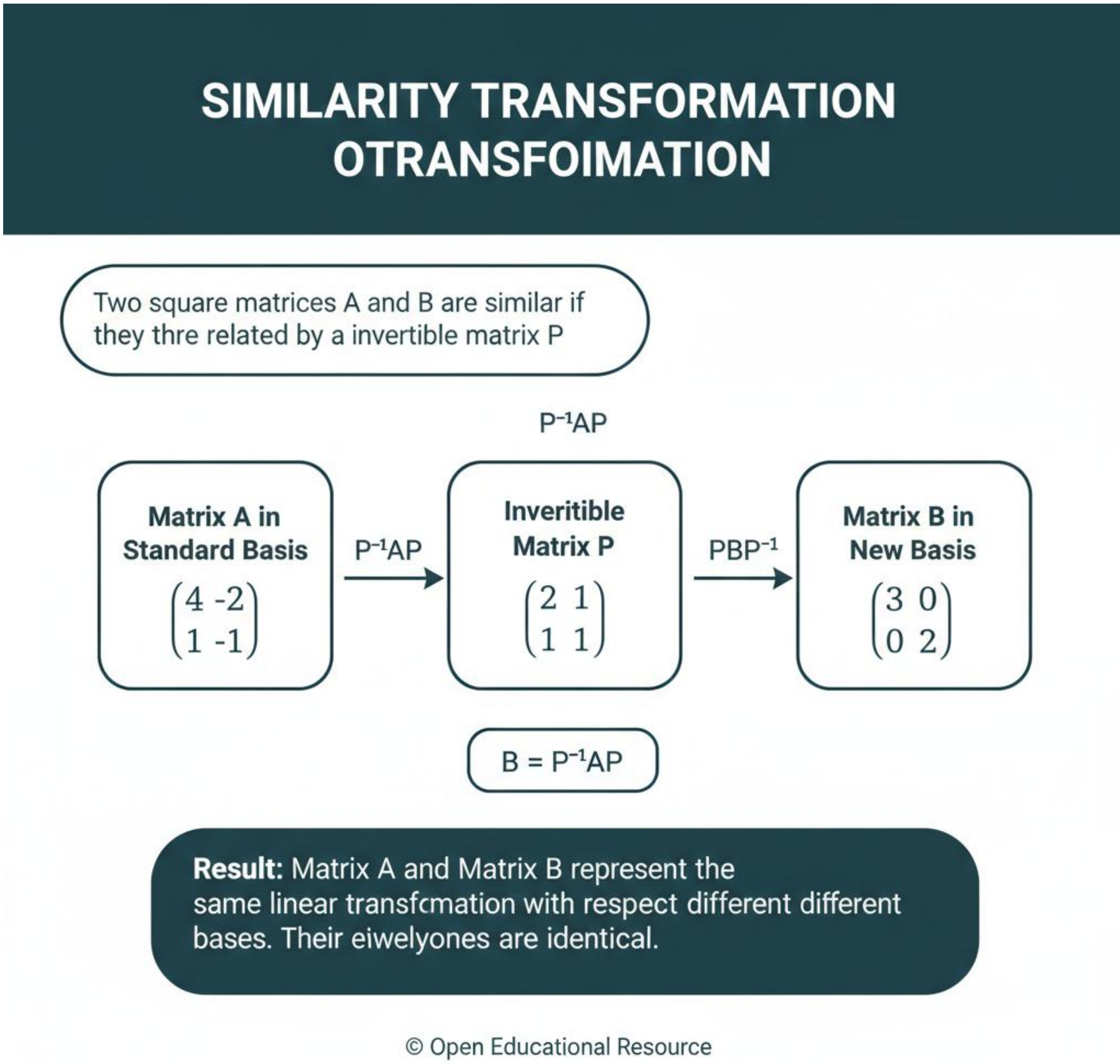
2.2 Similarity Of Matrices

2.2.1 Definition of similarity

Two matrices Aand Bare said to be **similar** if there exists an invertible matrix Psuch that:

$$B=P-1AP.$$

Similarity indicates that the two matrices represent the same linear transformation, but with respect to different bases. This means they share many important properties, such as determinant, trace, and eigenvalues.



Similarity transformation of matrices

Fill in the blank: Two matrices are similar if B=P-1APfor some ______ matrix P.

2.2.2 Properties of similar matrices

Similar matrices share several key properties:

They have the same **determinant**.

They have the same **trace** (sum of diagonal entries).

They have the same **rank**.

They have the same **eigenvalues**.

However, similar matrices may look different in form. The similarity relation allows us to study matrices by transforming them into simpler or more convenient forms without losing essential information.

Properties preserved under similarity of matrices

Property	Preserved?	Description
Determinant	**Yes**	$\det(B) = \det(P^{-1}AP) = \det(A)$. Both matrices have the same "scaling factor."
Trace	**Yes**	The sum of the diagonal elements remains identical ($\text{tr}(A) = \text{tr}(B)$.
Rank	**Yes**	Similar matrices have the same number of linearly independent rows/columns.
Eigenvalues	**Yes**	They share the exact same set of eigenvalues (the spectrum), including multiplicities.
Characteristic Polynomial	**Yes**	Since they share eigenvalues, their characteristic equations $\det(A - \lambda I) = 0$ are the same.
Invertibility	**Yes**	If A is invertible, B must also be invertible, and vice versa.
Eigenvectors	**No**	While eigenvalues are the same, the eigenvectors are typically different, related by the matrix P (if v is an eigenvector of A, then $P^{-1}v$ is an eigenvector of B).

Fill in the blank: Similar matrices always share the same ______ values.

2.2.3 Examples and applications

Example: Let

$$A=2\ 1\ 0\ 2\ ,P=1\ 1\ 0\ 1\ .$$

Then,

$$P\text{-}1AP=2\ 0\ 0\ 2\ =B.$$

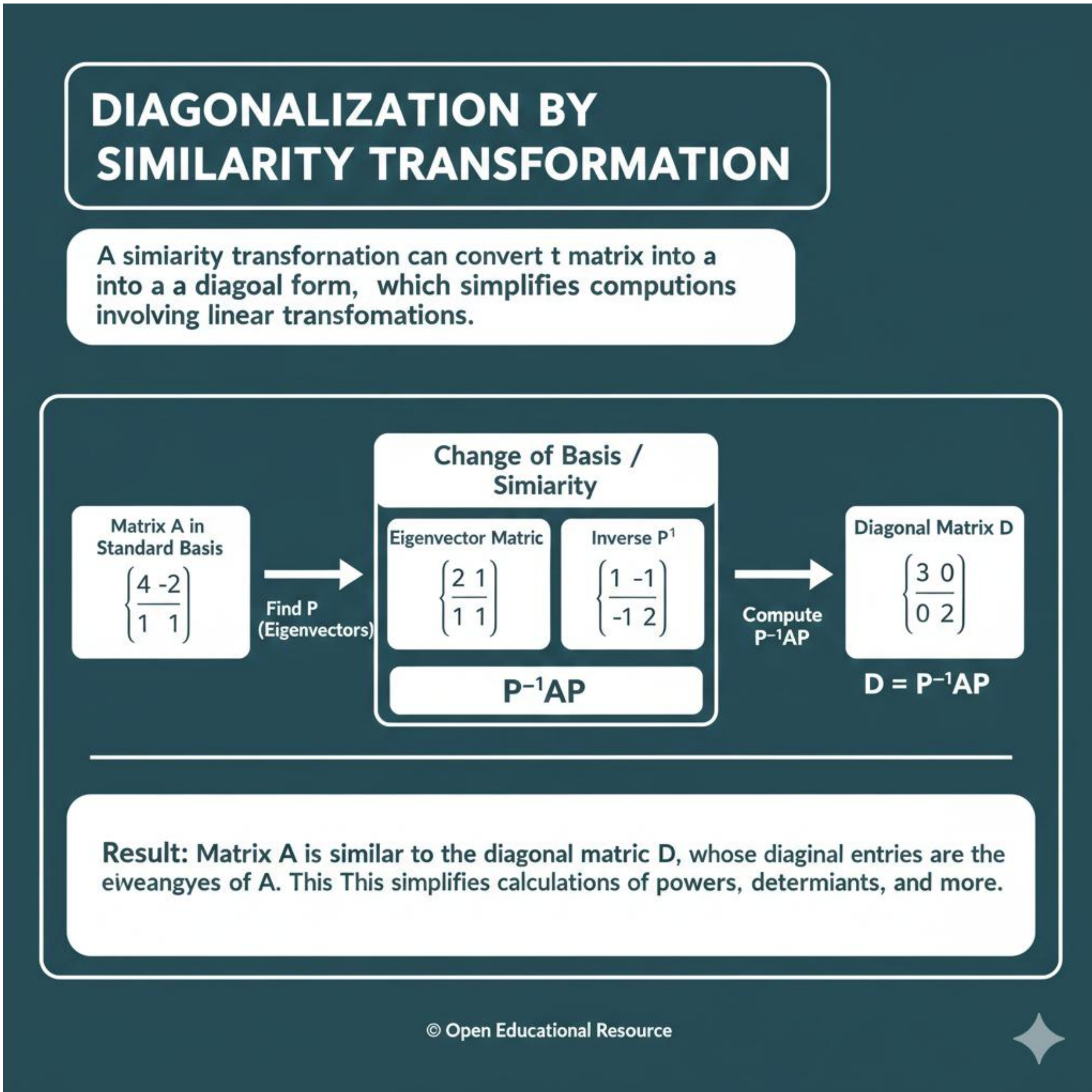
Here, Aand Bare similar, and Bis a diagonal matrix.

Applications:

Simplifying matrices to diagonal or Jordan form.

Analysing linear transformations in different bases.

Reducing computational complexity in solving problems.



Application of similarity in diagonalisation

Fill in the blank: Similarity is often used to transform a matrix into a more ______ form such as diagonal or Jordan form.

Pilot Questions 2.2

Two matrices Aand Bare similar if: A. They have the same determinant B. They can be transformed by an invertible matrix Psuch that B=P-1APC. They are both diagonal D. They have the same rank only

Which of the following properties is preserved under similarity? A. Determinant B. Trace C. Eigenvalues D. All of the above

If two matrices are similar, they represent: A. Different linear transformations B. The same linear transformation in different bases C. The same transformation in the same basis D. No transformation

In the example, the matrix A=2 1 0 2 was transformed into: A. A triangular matrix B. A diagonal matrix C. A zero matrix D. A rank-deficient matrix

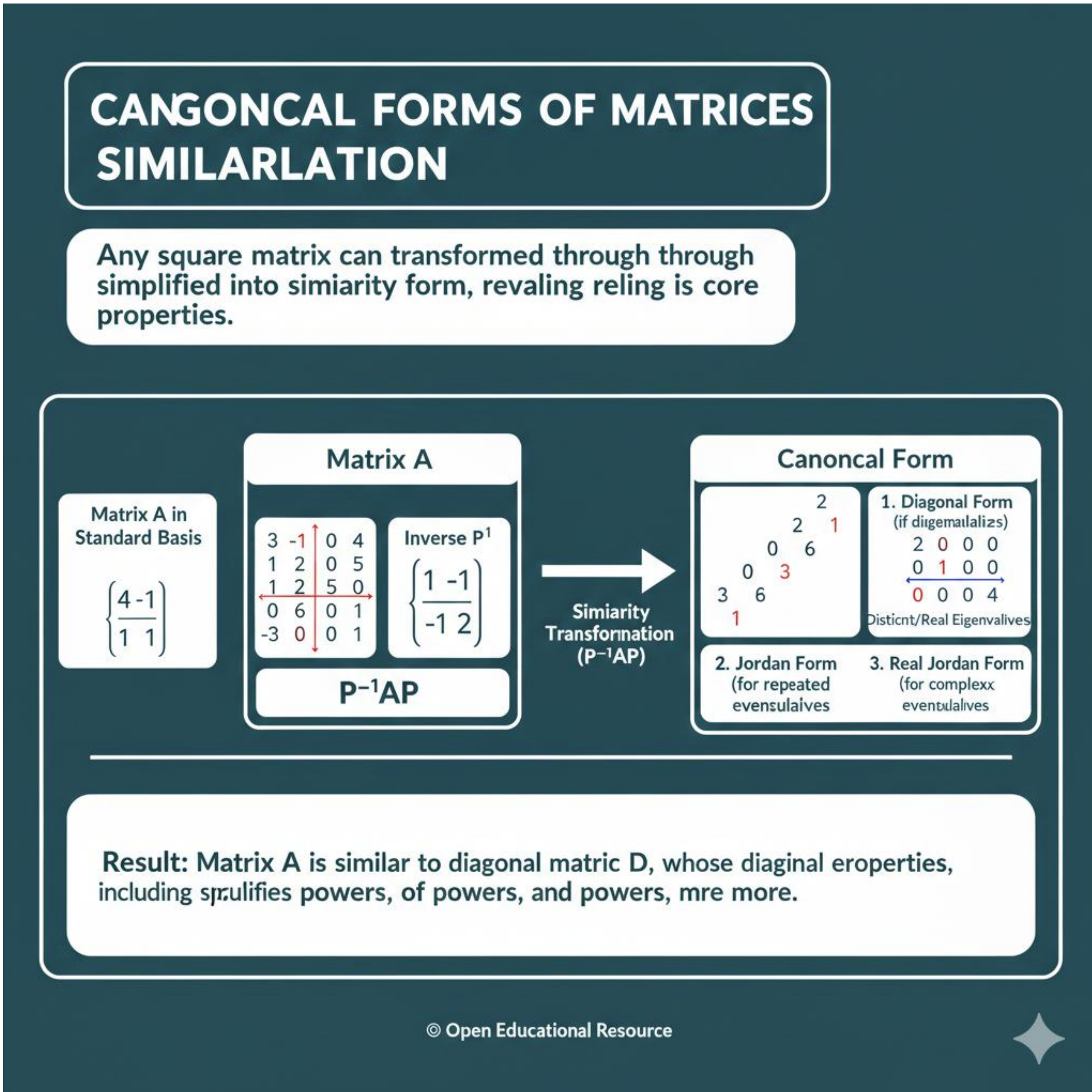
Similarity is useful in: A. Making matrices more complex B. Simplifying matrices to diagonal or Jordan form C. Eliminating eigenvalues D. Avoiding transformations

2.3 Canonical Forms

2.3.1 Concept of canonical forms

A **canonical form** is a simplified or standard representation of a matrix that retains its essential properties while making analysis easier. Canonical forms allow us to classify matrices into categories that reveal their structure more clearly. By transforming a matrix into a canonical form, we can study its behaviour without unnecessary complexity.

The most common canonical forms in linear algebra are the **diagonal form** and the **Jordan form**. These forms provide insight into eigenvalues, eigenvectors, and the structure of linear transformations.



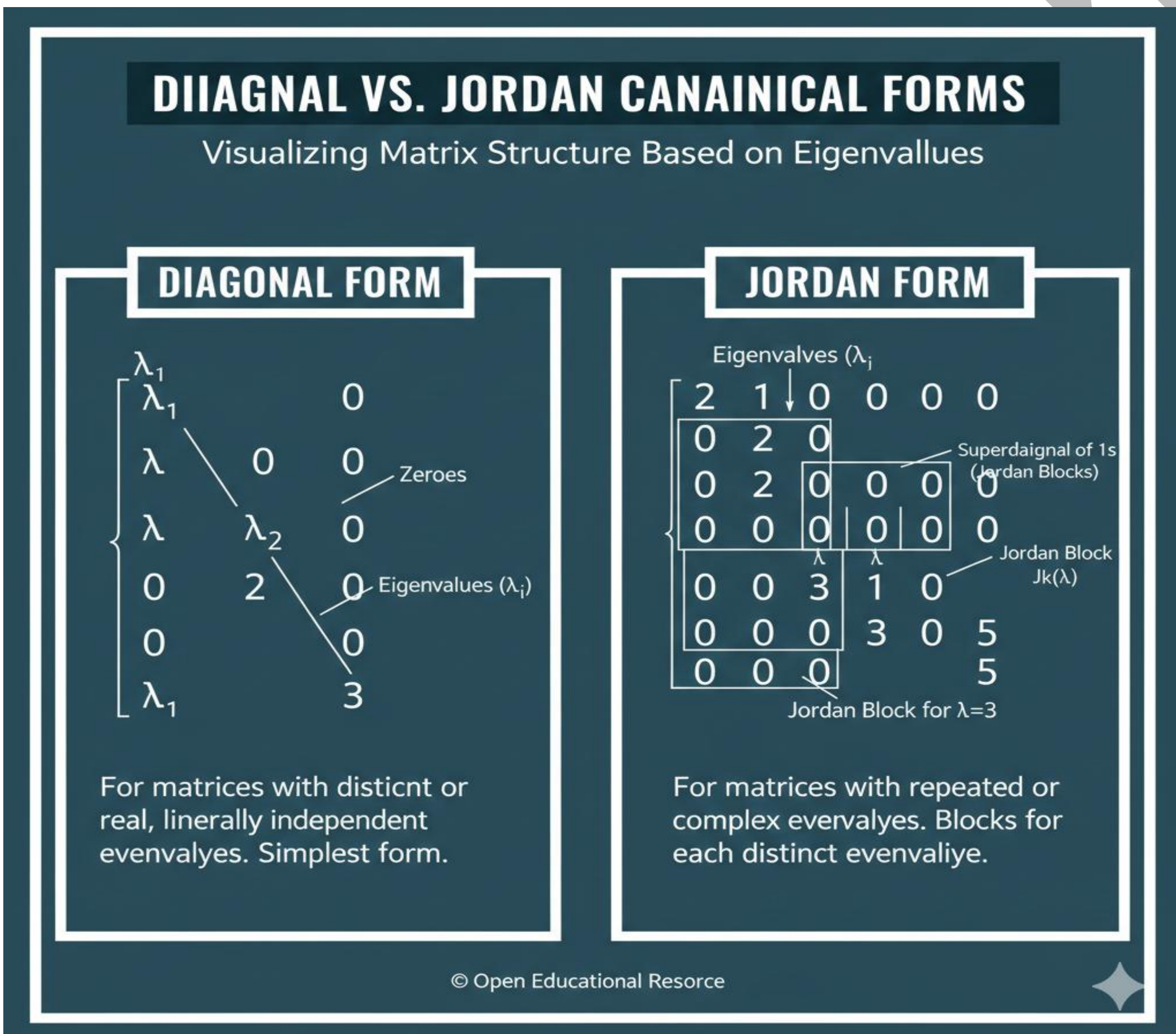
Transformation of a matrix into canonical

Fill in the blank: A canonical form is a simplified or ______ representation of a matrix.

2.3.2 Diagonal form and Jordan form

The **diagonal form** of a matrix is achieved when a matrix is similar to a diagonal matrix. This occurs when the matrix has enough linearly independent eigenvectors to form a basis. Diagonalisation simplifies computations, as powers of diagonal matrices are easy to calculate.

The **Jordan form** is a generalisation of diagonal form. When a matrix does not have enough independent eigenvectors, it can still be expressed in Jordan form, which consists of Jordan blocks along the diagonal. Each block corresponds to an eigenvalue and captures the structure of the matrix.



Diagonal and Jordan canonical forms AI prompt suggestion: “

Fill in the blank: A matrix that does not have enough independent eigenvectors can still be expressed in ______ form.

2.3.3 Applications of canonical forms in simplifying problems

Canonical forms are widely applied in mathematics and science:

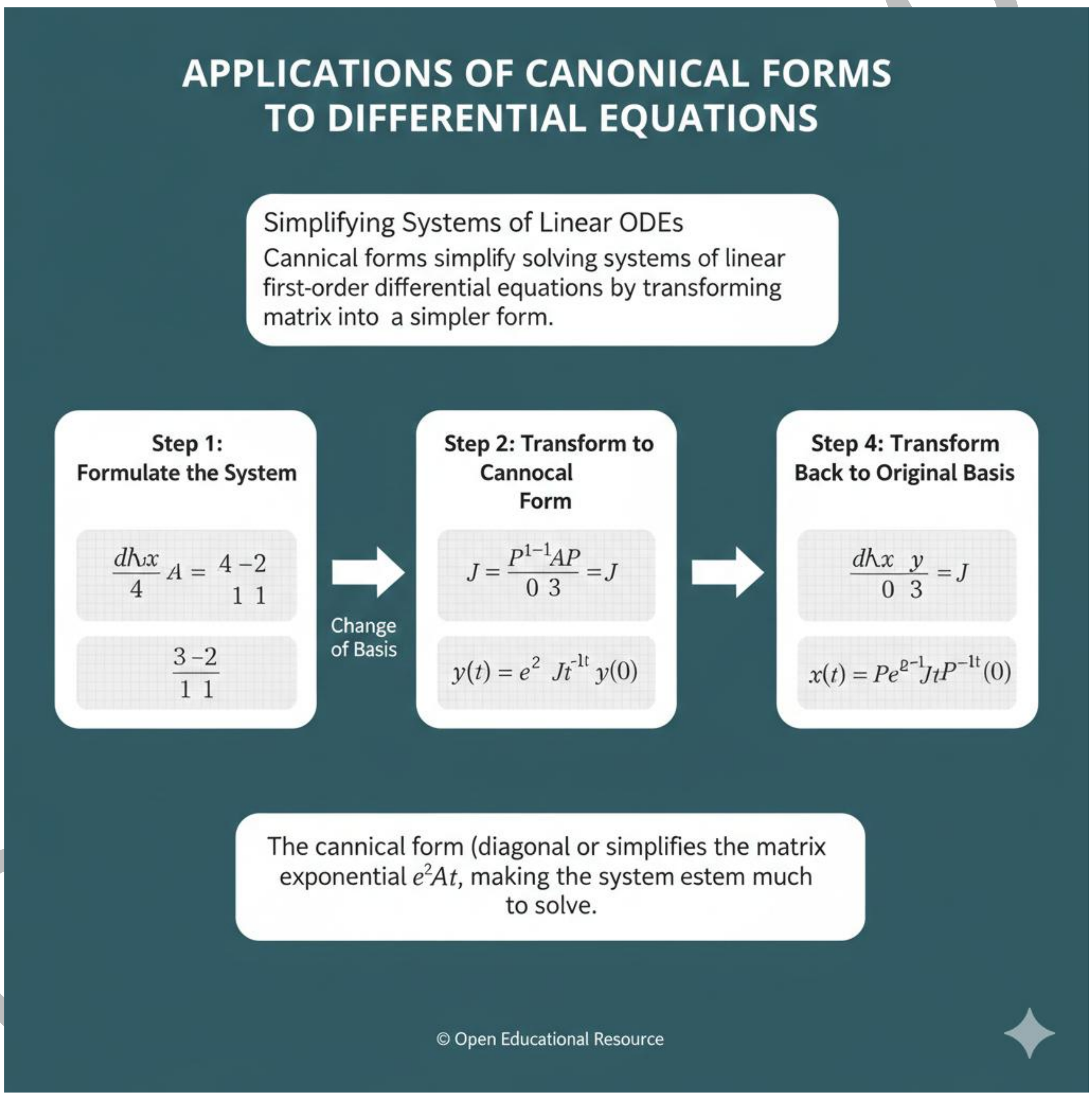
They simplify the computation of matrix powers and exponentials.

They make it easier to analyse linear transformations.

They are essential in solving differential equations using matrix methods.

They provide insight into the stability of systems in engineering and physics.

By reducing matrices to canonical forms, complex problems can be tackled more efficiently.



Application of canonical forms in solving differential equations

Fill in the blank: Canonical forms are especially useful in simplifying the computation of matrix ______.

Pilot Questions 2.3

A canonical form is: A. A random representation of a matrix B. A simplified or standard representation of a matrix C. A polynomial form of a matrix D. A determinant of a matrix

A matrix is diagonalised when: A. It has enough linearly independent eigenvectors B. It has no eigenvalues C. It is rank-deficient D. It is singular

The Jordan form consists of: A. Diagonal entries only B. Jordan blocks corresponding to eigenvalues C. Random entries D. Determinants

Canonical forms are useful in: A. Making problems more complex B. Simplifying computations and analysis C. Eliminating eigenvalues D. Avoiding transformations

In solving differential equations, canonical forms help by: A. Reducing matrices to simpler forms B. Increasing the number of variables C. Eliminating constants D. Avoiding eigenvectors

Series Summary

This Series explored how matrices can be classified and simplified through the concepts of **equivalence**, **similarity**, and **canonical forms**. We began with equivalence, focusing on elementary transformations and equivalence classes. We then examined similarity, highlighting its definition, properties, and applications in diagonalisation. Finally, we studied canonical forms, including diagonal and Jordan forms, and their role in simplifying complex problems such as solving differential equations.

Glossary of Terms

Canonical form: A simplified or standard representation of a matrix that retains its essential properties.

Diagonal form: A canonical form where a matrix is similar to a diagonal matrix, possible when enough independent eigenvectors exist.

Equivalence: A relation between matrices where one can be transformed into another using elementary row and column operations.

Equivalence class: A group of matrices that are equivalent to each other, sharing the same rank.

Elementary transformation: Basic operations such as swapping rows, multiplying by a non-zero constant, or adding multiples of rows or columns.

Jordan form: A canonical form consisting of Jordan blocks along the diagonal, used when a matrix lacks sufficient independent eigenvectors.

Rank: The number of linearly independent rows or columns in a matrix.

Similarity: A relation between matrices where one is obtained from another by a similarity transformation B=P-1AP.

Transition matrix: An invertible matrix Pused in similarity transformations to change bases.

Concept Check Questions 2

Objective questions

Two matrices are equivalent if: A. They have the same determinant B. They can be transformed into one another using elementary operations C. They are both diagonal D. They have the same eigenvalues (Linked to SLO 2.1)

Which property is preserved under similarity? A. Determinant B. Trace C. Eigenvalues D. All of the above (Linked to SLO 2.3)

Short-answer questions

Define an equivalence class of matrices and state its key invariant. (Linked to SLO 2.1)

Explain why similarity is considered a stronger relation than equivalence. (Linked to SLO 2.4)

Long-answer questions

Describe canonical forms and demonstrate how diagonal and Jordan forms simplify matrix analysis. (Linked to SLO 2.5, SLO 2.6, SLO 2.7)

Basic response: Define canonical forms and explain diagonal and Jordan forms.

Value-added response: Discuss applications in solving differential equations, stability analysis, and computational efficiency.

Answers to Concept Check Questions 2

B. They can be transformed into one another using elementary operations

D. All of the above

An equivalence class is a group of matrices that can be transformed into one another using elementary operations. The key invariant is rank.

Similarity preserves more properties than equivalence, including determinant, trace, and eigenvalues, making it a stronger relation.

Basic response: Canonical forms are simplified representations of matrices. Diagonal form applies when enough independent eigenvectors exist, while Jordan form generalises this when eigenvectors are insufficient.

Value-added response: Canonical forms are applied in solving differential equations, analysing system stability, and reducing computational complexity in engineering and physics.

Answers to Pilot Questions 2

From Part 2.1

They can be transformed into one another using elementary operations

Swapping two rows

Rank

Simplifying linear transformations

Reducing the augmented matrix to a simpler form

From Part 2.2

They can be transformed by an invertible matrix Psuch that B=P-1AP

All of the above

The same linear transformation in different bases

A diagonal matrix

Simplifying matrices to diagonal or Jordan form

From Part 2.3

A simplified or standard representation of a matrix

It has enough linearly independent eigenvectors

Jordan blocks corresponding to eigenvalues

Simplifying computations and analysis

Reducing matrices to simpler forms

3 Eigenvalues, Eigenvectors, And Characteristic Polynomials

Series Outline

For this 1-credit course, we will structure Series 3 into three Parts, ensuring clarity and progression:

Part 3.1 Eigenvalues and eigenvectors

3.1.1 Definition of eigenvalues and eigenvectors

3.1.2 Geometric interpretation

3.1.3 Applications in linear transformations

Part 3.2 Characteristic polynomials

3.2.1 Definition and computation

3.2.2 Relationship with eigenvalues

3.2.3 Examples and applications

Part 3.3 Applications of eigenvalues and eigenvectors

3.3.1 Diagonalisation of matrices

3.3.2 Stability analysis in systems

3.3.3 Applications in science and engineering

Introduction

Eigenvalues and eigenvectors are central concepts in linear algebra, providing deep insight into the behaviour of linear transformations. They reveal directions in which transformations act by simple stretching or compressing, and they play a crucial role in simplifying matrices through diagonalisation. Closely related to these is the characteristic polynomial, which provides a systematic way of finding eigenvalues.

The aim of this Series is to equip you with the ability to define, compute, and apply eigenvalues, eigenvectors, and characteristic polynomials, and to demonstrate their importance in both theoretical and practical contexts.

Series Learning Outcomes

Upon completing this Series, you should be able to:

SLO 3.1 define eigenvalues and eigenvectors clearly and accurately,

SLO 3.2 explain the geometric interpretation of eigenvectors in relation to linear transformations,

SLO 3.3 demonstrate how to compute eigenvalues and eigenvectors for given matrices,

SLO 3.4 define the characteristic polynomial of a matrix and show how it is derived,

SLO 3.5 calculate characteristic polynomials and use them to determine eigenvalues,

SLO 3.6 apply eigenvalues and eigenvectors in diagonalising matrices,

SLO 3.7 analyse the role of eigenvalues in stability analysis of systems, and

SLO 3.8 evaluate practical applications of eigenvalues and eigenvectors in science and engineering contexts.

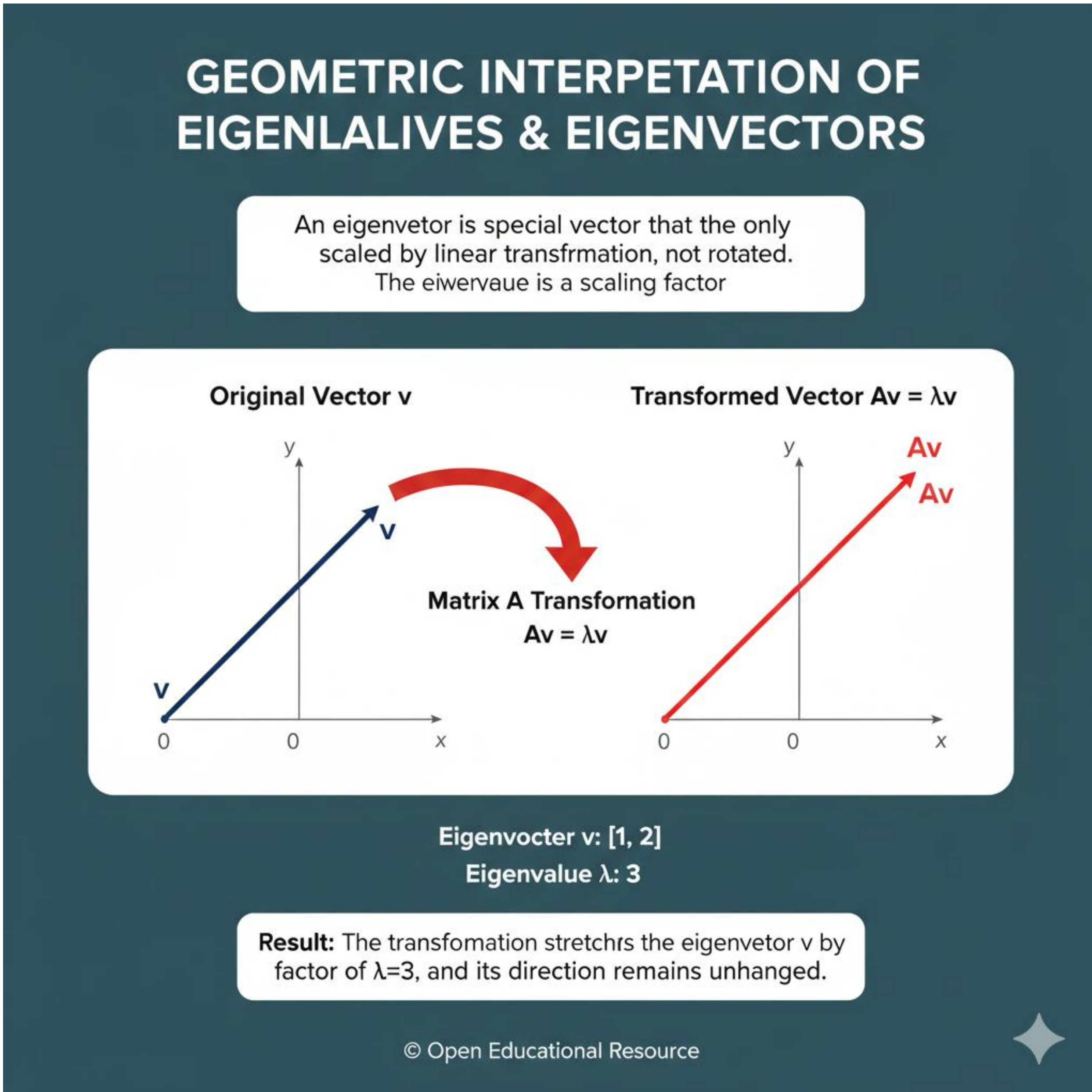
3.1 Eigenvalues And Eigenvectors

3.1.1 Definition of eigenvalues and eigenvectors

An eigenvalue of a square matrix Ais a scalar such that there exists a non-zero vector v(called an eigenvector) satisfying:

$$Av=\lambda v.$$

This means that when the matrix Aacts on the vector v, the result is simply a scaled version of v. The eigenvector points in a direction that is unchanged by the transformation, while the eigenvalue tells us how much it is stretched or compressed.



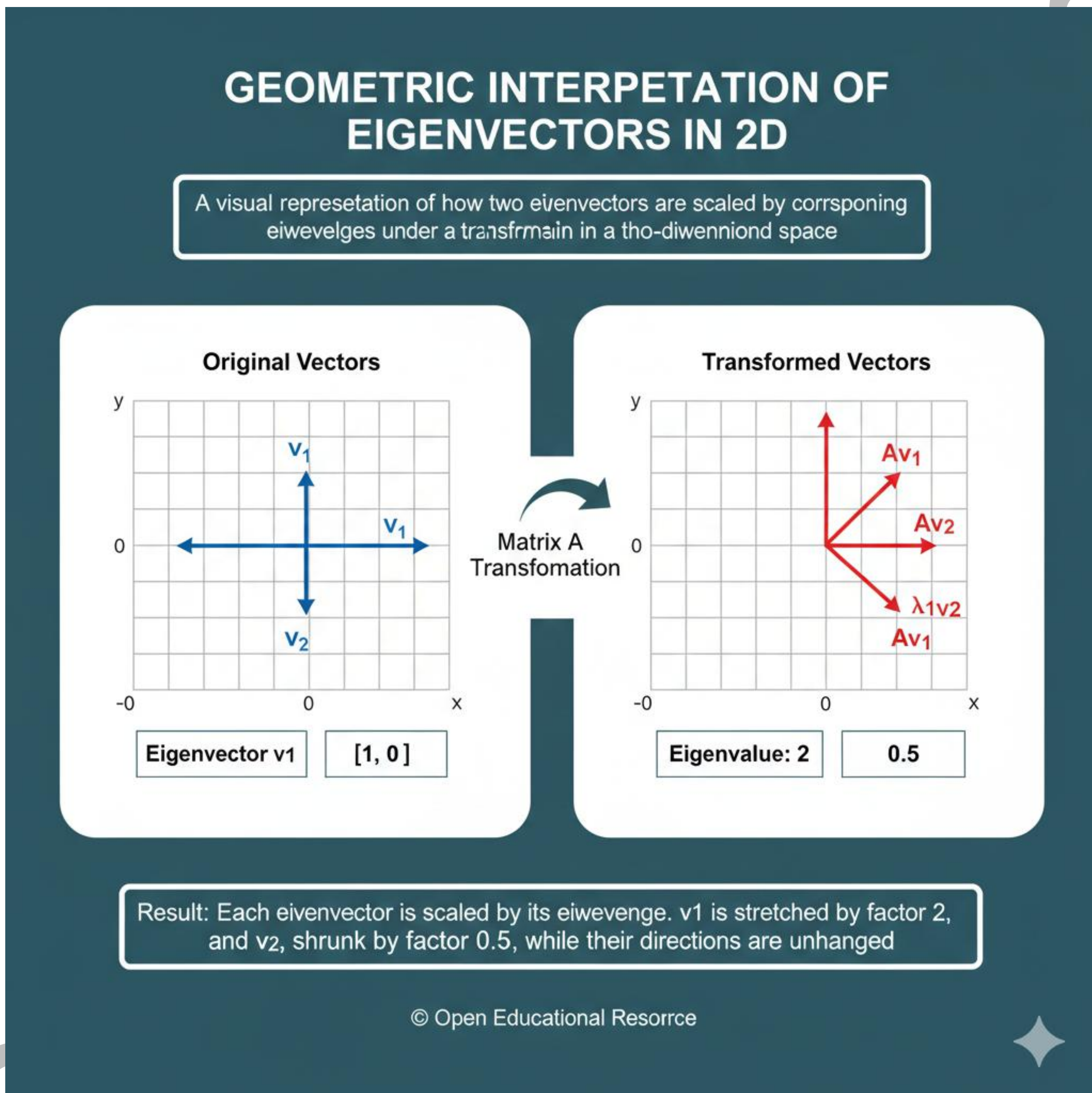
Geometric meaning of eigenvalues and eigenvectors

Fill in the blank: An eigenvector is a non-zero vector that remains in the same ______ after a matrix transformation.

3.1.2 Geometric interpretation

Geometrically, eigenvectors represent directions in space that are preserved under a linear transformation. The corresponding eigenvalues indicate how much vectors in those directions are scaled.

For example, in two dimensions, a transformation may rotate most vectors, but eigenvectors will remain aligned with their original direction, only stretched or compressed.



Geometric interpretation of eigenvectors in two dimensions

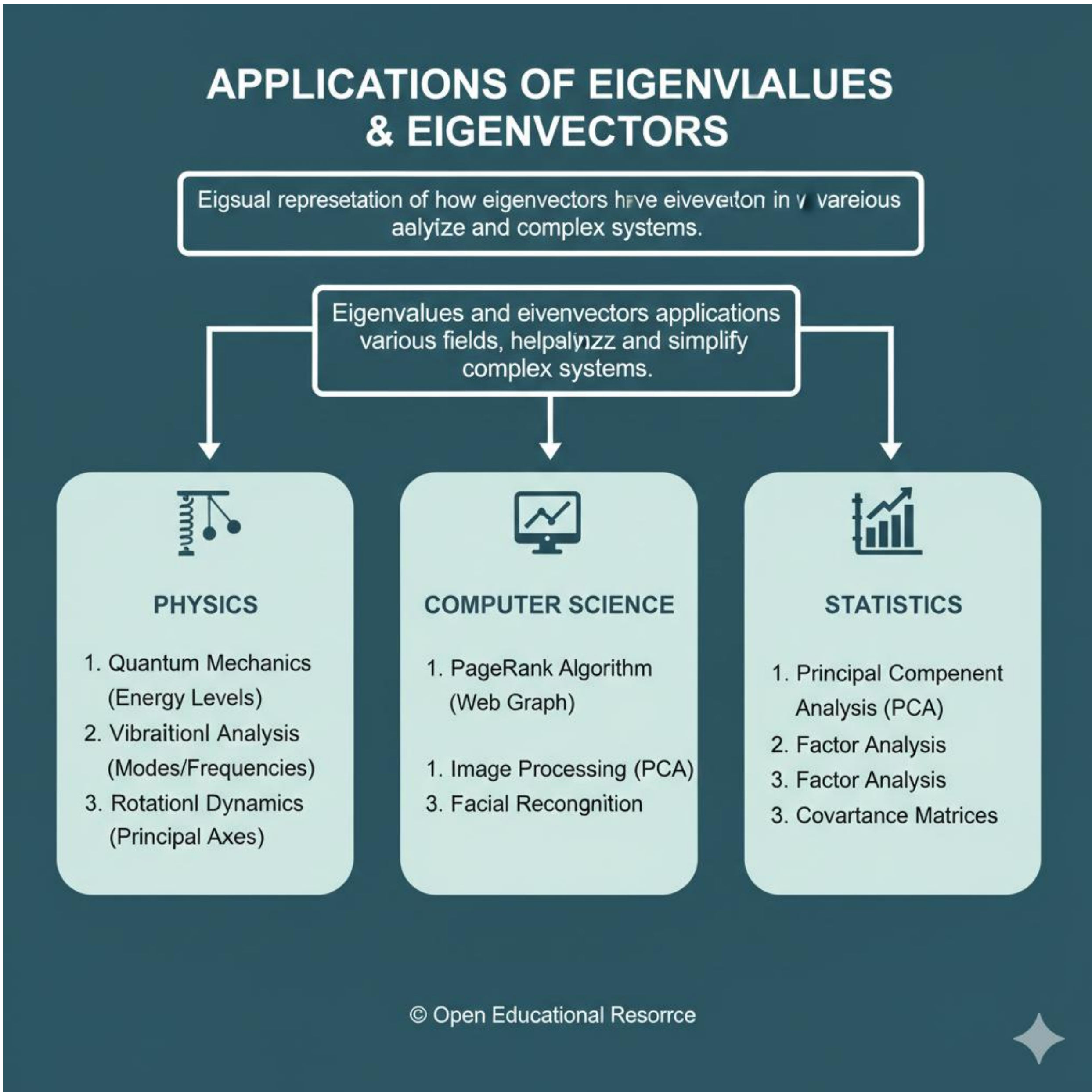
Fill in the blank: Eigenvalues indicate how much eigenvectors are ______ or compressed.

3.1.3 Applications in linear transformations

Eigenvalues and eigenvectors are widely applied in mathematics and science:

- In physics, they describe modes of vibration in mechanical systems.
- In computer science, they are used in algorithms such as Google's PageRank.
- In statistics, they underpin principal component analysis (PCA) for dimensionality reduction.
- In engineering, they help analyse stability and control systems.

By identifying eigenvalues and eigenvectors, we gain powerful tools for simplifying problems and understanding the behaviour of transformations.



Applications of eigenvalues and eigenvectors

Fill in the blank: In statistics, eigenvalues and eigenvectors are used in ______ component analysis.

Pilot Questions 3.1

An eigenvalue and eigenvector vsatisfy the equation: A. Av=vB. Av=λvC. Av=v+λD. Av=0

Eigenvectors represent: A. Directions that change under transformation B. Directions that remain unchanged except for scaling C. Random directions in space D. Only diagonal entries of a matrix

Eigenvalues indicate: A. How much eigenvectors are stretched or compressed B. The number of variables in a system C. The determinant of a matrix D. The trace of a matrix

Which of the following is an application of eigenvalues and eigenvectors? A. PageRank algorithm B. Principal component analysis C. Stability analysis in engineering D. All of the above

In physics, eigenvalues and eigenvectors are used to describe: A. Modes of vibration B. Random noise C. Polynomial equations D. Matrix rank

3.2 Characteristic Polynomials

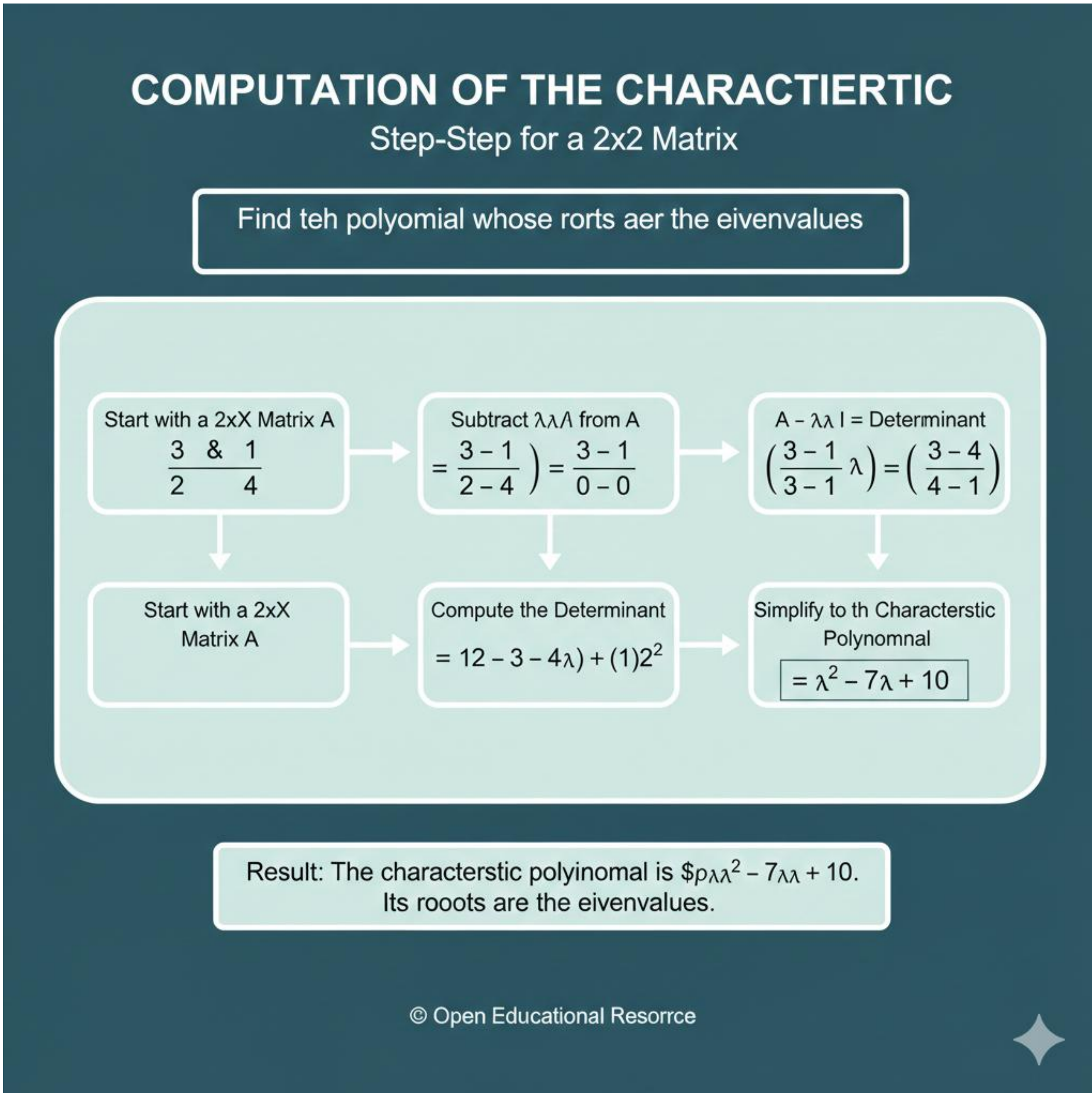
3.2.1 Definition and computation

The characteristic polynomial of a square matrix Ais defined as:

$$p(\lambda)=\det(A-\lambda I),$$

where is a scalar and Iis the identity matrix of the same size as A.

This polynomial encodes important information about the matrix. Its roots are the eigenvalues of A. Computing the characteristic polynomial involves subtracting from the diagonal entries of A, forming $A-\lambda I$, and then calculating the determinant.



Computation of the characteristic polynomial of a 2×2 matrix

Fill in the blank: The characteristic polynomial of a matrix is given by \det(A - \lambda ______).

3.2.2 Relationship with eigenvalues

The eigenvalues of a matrix are precisely the roots of its characteristic polynomial. If $p(\lambda)=0$, then is an eigenvalue of A.

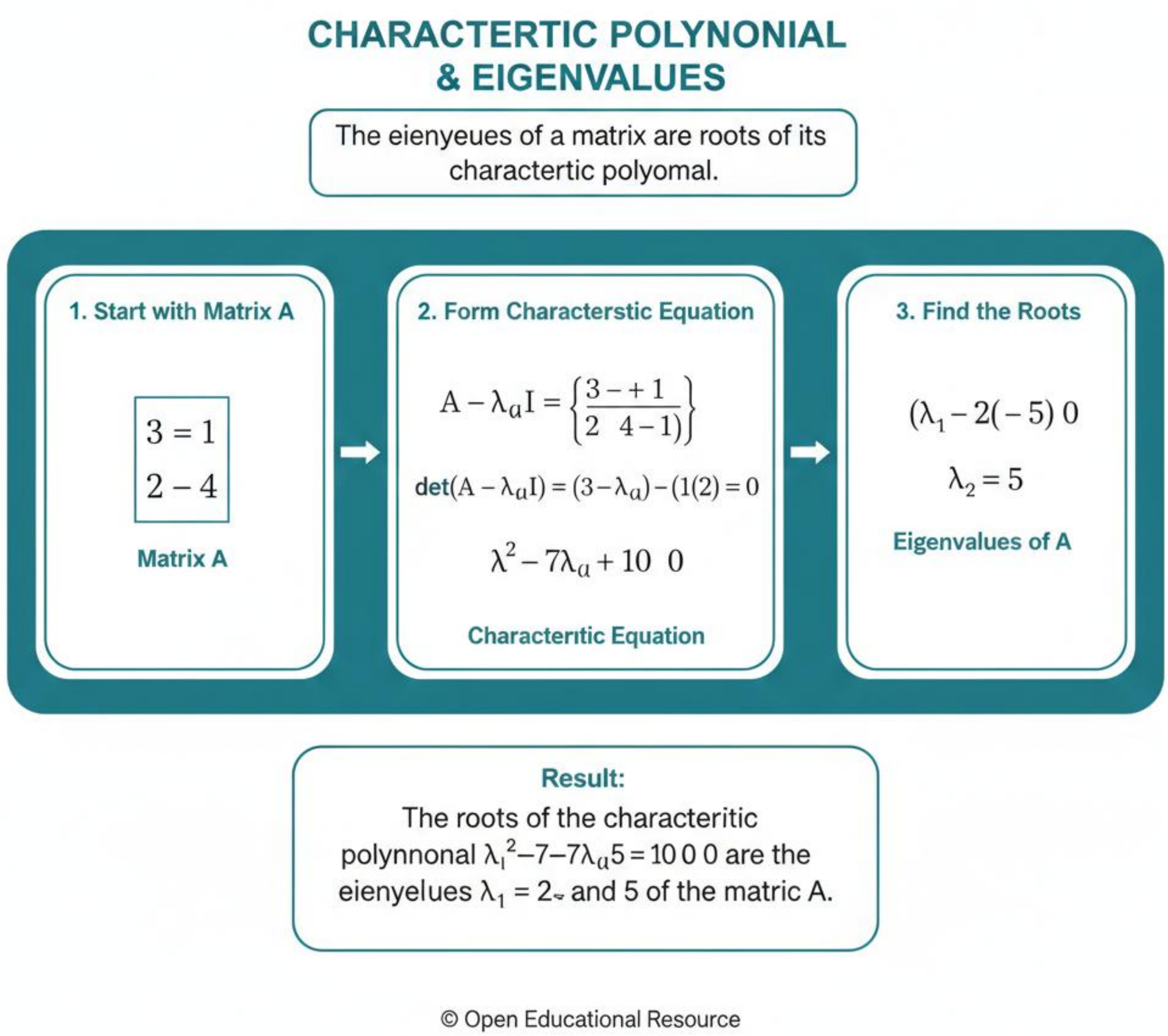
For example, for

$$A=2\ 1\ 0\ 3,$$

the characteristic polynomial is:

$$p(\lambda)=\det 2-\lambda\ 1\ 0\ 3-\lambda=(2-\lambda)(3-\lambda).$$

The roots are $\lambda=2$and $\lambda=3$, which are the eigenvalues of A.



Relationship between characteristic polynomial and eigenvalues

Fill in the blank: The eigenvalues of a matrix are the ______ of its characteristic polynomial.

3.2.3 Examples and applications

Characteristic polynomials are used to:

Find eigenvalues systematically.

Analyse stability of systems in engineering.

Classify matrices according to their eigenstructure.

Simplify computations in diagonalisation and Jordan form.

Example: For

$$A=4\ 2\ 1\ 3,$$

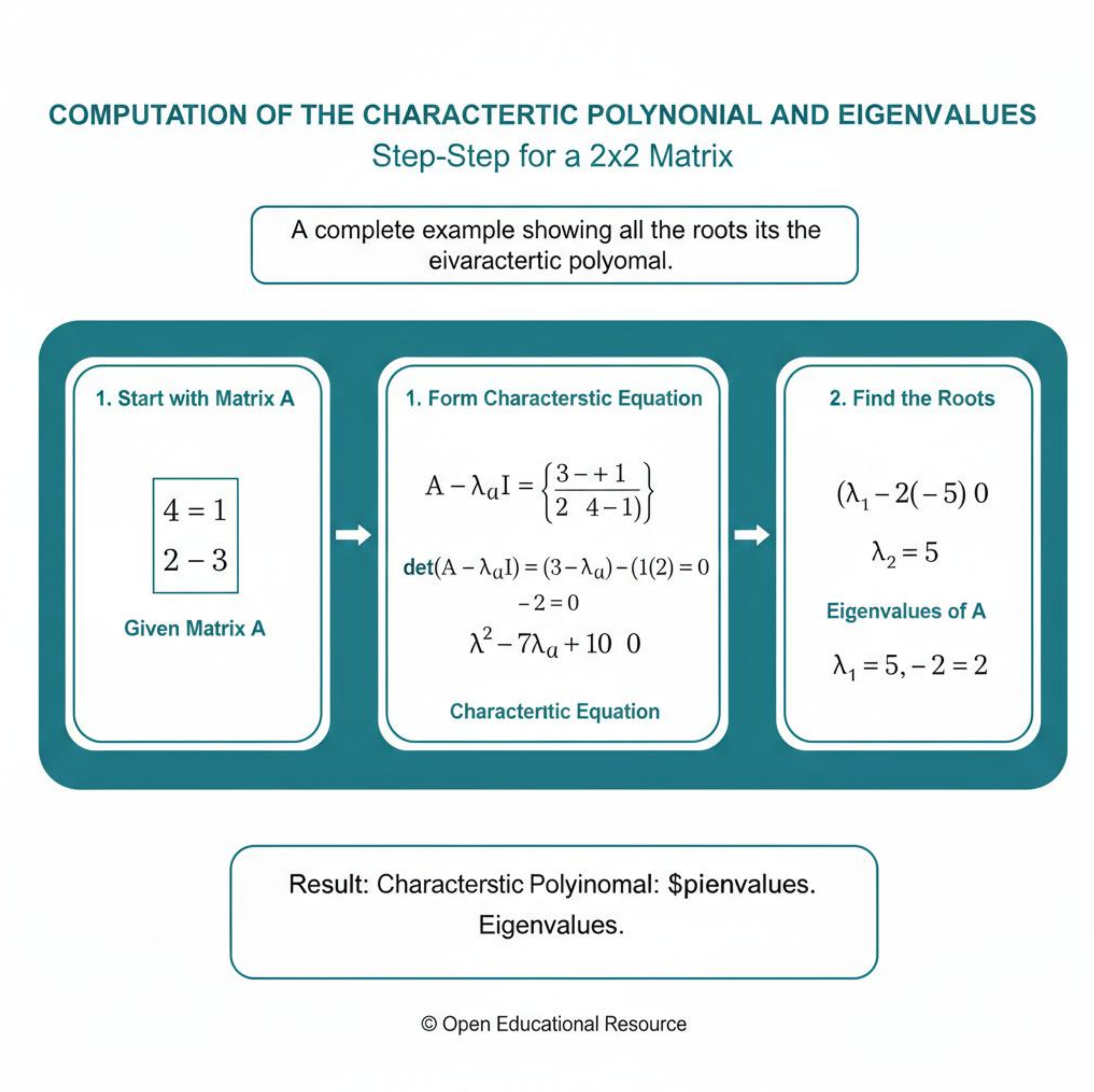
the characteristic polynomial is:

$$p(\lambda)=\det 4-\lambda\ 2\ 1\ 3-\lambda=(4-\lambda)(3-\lambda)-2.$$

Simplifying:

$$p(\lambda)=2-7\lambda+10.$$

The roots are $\lambda=5$and $\lambda=2$, which are the eigenvalues.



Example of characteristic polynomial and eigenvalues

Fill in the blank: Characteristic polynomials are especially useful in finding ______ systematically.

Pilot Questions 3.2

The characteristic polynomial of a matrix A is defined as: A. $\det(A+\lambda I)$ B. $\det(A-\lambda I)$ C. $\det(A)$ D. $\det(\lambda I)$

The eigenvalues of a matrix are: A. The coefficients of its characteristic polynomial B. The roots of its characteristic polynomial C. The trace of the matrix D. The determinant of the matrix

Which of the following is an application of characteristic polynomials? A. Finding eigenvalues B. Analysing stability of systems C. Diagonalisation D. All of the above

For the matrix

$$A=2\ 1\ 0\ 3\,,$$

the characteristic polynomial is: A. $2-\lambda)(3-\lambda$ B. $2+\lambda)(3+\lambda$ C. $2-\lambda)(3+\lambda$ D. None of the above

The roots of the polynomial $2-7\lambda+10$ are: A. 2 and 5 B. 3 and 4 C. 7 and 10 D. 1 and 10

3.3 Applications Of Eigenvalues And Eigenvectors

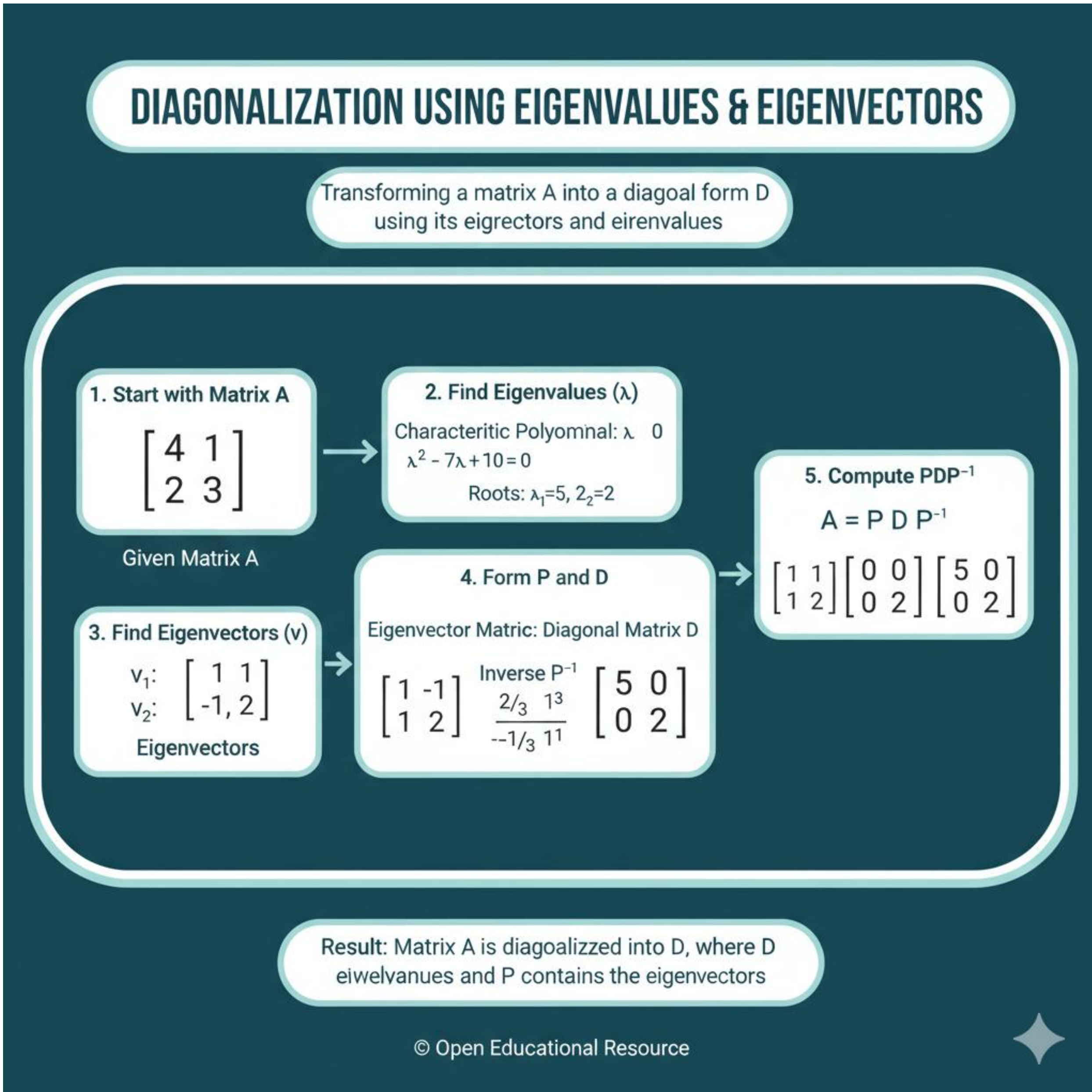
3.3.1 Diagonalisation of matrices

One of the most important applications of eigenvalues and eigenvectors is diagonalisation. A matrix A is diagonalised if it can be expressed as:

$$A=PDP-1,$$

where D is a diagonal matrix containing the eigenvalues of A, and P is a matrix whose columns are the corresponding eigenvectors.

Diagonalisation simplifies computations, especially when calculating powers of matrices or solving systems of differential equations.



Diagonalisation of a matrix using eigenvalues and eigenvectors

Fill in the blank: A matrix is diagonalised when it can be expressed as A=PDP-1, where Dcontains the ______.

3.3.2 Stability analysis in systems

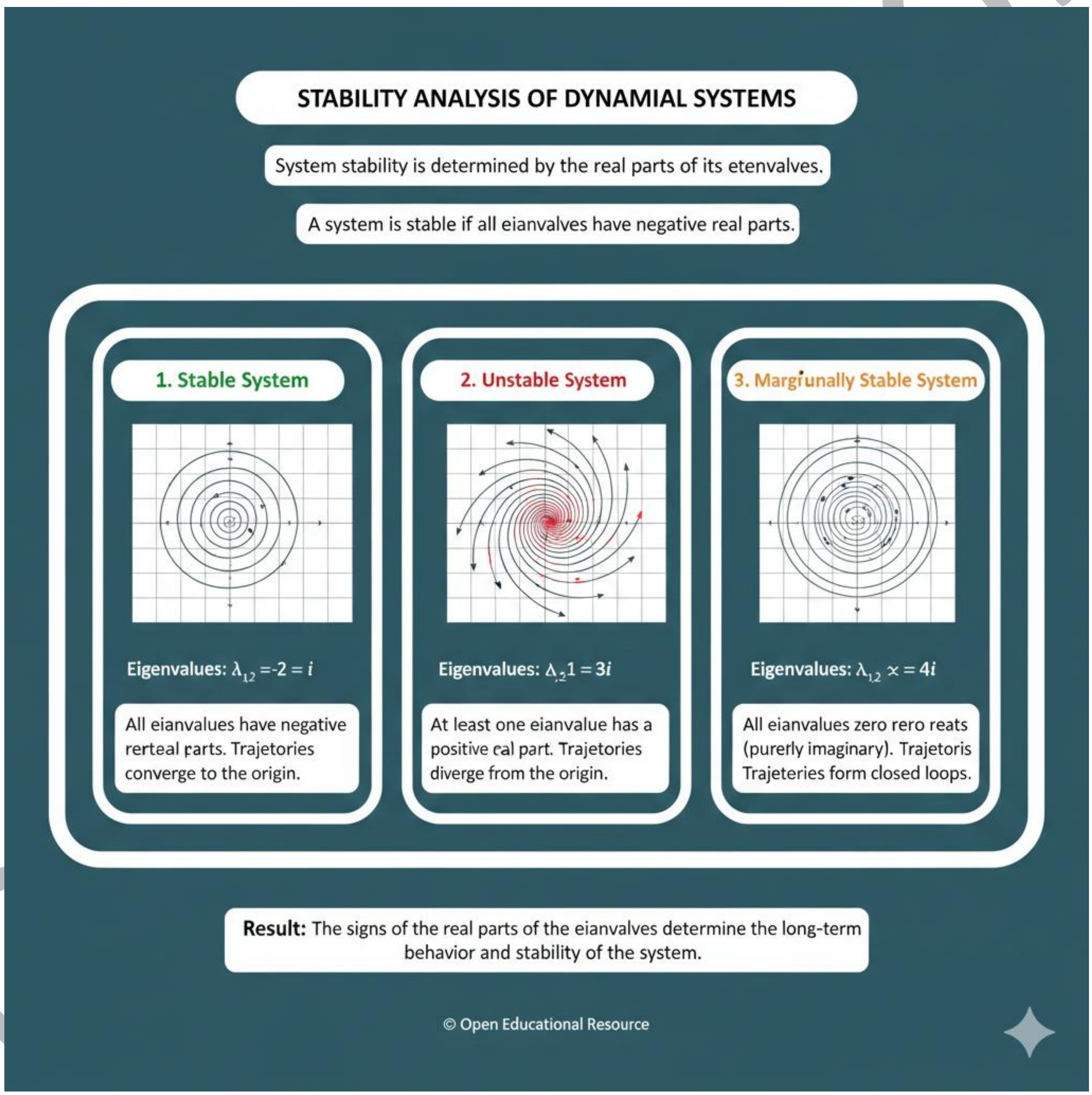
Eigenvalues play a crucial role in determining the stability of systems, especially in engineering and applied mathematics.

If all eigenvalues of a system's matrix have negative real parts, the system is stable.

If any eigenvalue has a positive real part, the system is unstable.

Eigenvalues with zero real parts indicate marginal stability.

This analysis is widely used in control theory, mechanical vibrations, and population dynamics.



Stability analysis using eigenvalues

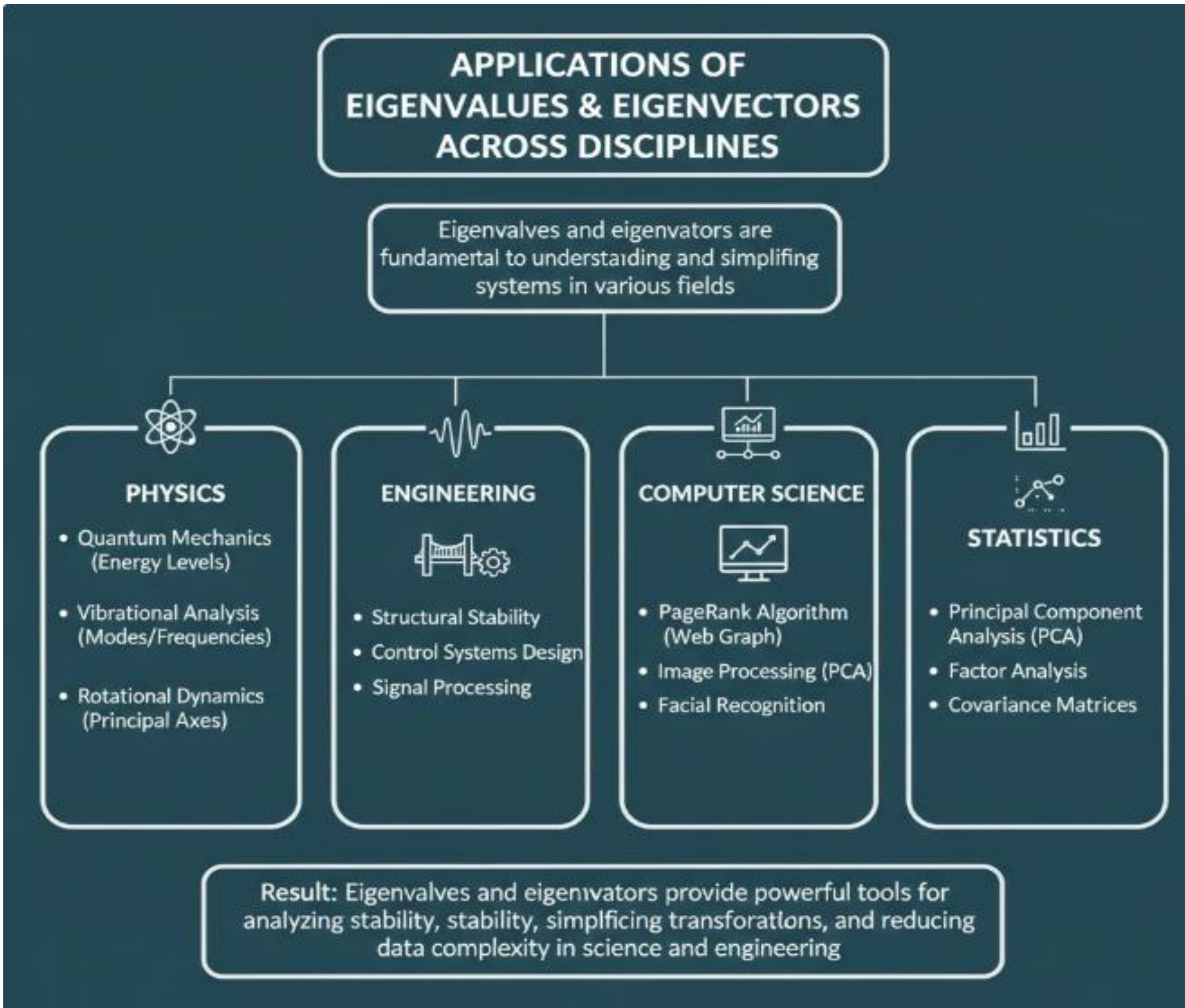
Fill in the blank: A system is stable if all eigenvalues have ______ real parts.

3.3.3 Applications in science and engineering

Eigenvalues and eigenvectors have broad applications:

- Physics: Analysing vibrations and quantum mechanics.
- Engineering: Studying structural stability and control systems.
- Computer science: Algorithms such as PageRank and machine learning.
- Statistics: Principal component analysis (PCA) for dimensionality reduction.

These applications highlight the versatility of eigenvalues and eigenvectors in both theoretical and practical contexts.



Applications of eigenvalues and eigenvectors in science and engineering

Fill in the blank: In computer science, eigenvalues and eigenvectors are used in algorithms such as ______.

Pilot Questions 3.3

- A matrix is diagonalised when: A. It is expressed as A=PDP-1B. It is expressed as A=P+DC. It is expressed as A=DPD. It is expressed as A=P-1DP
- Stability analysis using eigenvalues shows that a system is stable if: A. All eigenvalues have positive real parts B. All eigenvalues have negative real parts C. All eigenvalues are zero D. Eigenvalues are imaginary
- Which of the following fields use eigenvalues and eigenvectors? A. Physics B. Engineering C. Computer science D. All of the above

In control theory, eigenvalues with positive real parts indicate: A. Stability B. Instability C. Marginal stability D. No effect

In statistics, eigenvalues and eigenvectors are applied in: A. Regression analysis B. Principal component analysis C. Probability distributions D. Hypothesis testing

Series Summary

This Series introduced the fundamental concepts of **eigenvalues** and **eigenvectors**, which describe how matrices act on vectors by stretching or compressing them along preserved directions. We explored their geometric interpretation and wide-ranging applications in mathematics, physics, computer science, and statistics. We then examined the **characteristic polynomial**, a tool for systematically finding eigenvalues, and demonstrated its computation through examples. Finally, we studied applications of eigenvalues and eigenvectors, including diagonalisation of matrices, stability analysis in systems, and practical uses in science and engineering.

Glossary of Terms

Characteristic polynomial: A polynomial defined by $\det(A-\lambda I)$, whose roots are the eigenvalues of a matrix.

Diagonalisation: The process of expressing a matrix as $A=PDP^{-1}$, where D is diagonal and contains eigenvalues, and P contains eigenvectors.

Eigenvalue: A scalar such that $Av=\lambda v$ for some non-zero vector v.

Eigenvector: A non-zero vector v that remains in the same direction after a matrix transformation, scaled by its eigenvalue.

Principal component analysis (PCA): A statistical technique that uses eigenvalues and eigenvectors to reduce dimensionality of data.

Stability analysis: The study of system behaviour using eigenvalues to determine whether a system is stable, unstable, or marginally stable.

Concept Check Questions 3

Objective questions

An eigenvalue and eigenvector v satisfy the equation: A. $Av=v$ B. $Av=\lambda v$ C. $Av=v+\lambda$ D. $Av=0$ (Linked to SLO 3.1)

The eigenvalues of a matrix are: A. The coefficients of its characteristic polynomial B. The roots of its characteristic polynomial C. The trace of the matrix D. The determinant of the matrix (Linked to SLO 3.5)

Short-answer questions

Define an eigenvector and explain its geometric meaning. (Linked to SLO 3.2)

State the relationship between the characteristic polynomial and eigenvalues. (Linked to SLO 3.4, SLO 3.5)

Long-answer questions

Discuss applications of eigenvalues and eigenvectors in diagonalisation, stability analysis, and science/engineering. (Linked to SLO 3.6, SLO 3.7, SLO 3.8)

Basic response: Describe diagonalisation, stability analysis, and one scientific application.

Value-added response: Provide detailed examples such as PCA in statistics, PageRank in computer science, and vibration analysis in physics.

Answers to Concept Check Questions 3

B. $Av=\lambda v$

B. The roots of its characteristic polynomial

An eigenvector is a non-zero vector that remains in the same direction after a matrix transformation. Geometrically, it represents a preserved direction under the transformation.

The eigenvalues of a matrix are the roots of its characteristic polynomial.

Basic response: Eigenvalues and eigenvectors are used to diagonalise matrices, making computations easier. They also determine system stability and are applied in scientific contexts such as physics.

Value-added response: In statistics, PCA uses eigenvalues and eigenvectors to reduce dimensionality. In computer science, Google's PageRank algorithm relies on them. In physics, they describe modes of vibration in mechanical systems.

Answers to Pilot Questions 3

From Part 3.1

$Av=\lambda v$

Directions that remain unchanged except for scaling

How much eigenvectors are stretched or compressed

All of the above

Modes of vibration

From Part 3.2

$\det(A-\lambda I)$

The roots of its characteristic polynomial

All of the above

$2-\lambda)(3-\lambda$

2 and 5

From Part 3.3

It is expressed as A=PDP-1

All eigenvalues have negative real parts

All of the above

Instability

Principal component analysis

Series 4 Quadratic Forms And Orthogonality

Series 4 Outline

For this 1-credit course, Series 4 will be structured into three Parts:

Part 4.1 Quadratic forms

4.1.1 Definition of quadratic forms

4.1.2 Matrix representation of quadratic forms

4.1.3 Applications of quadratic forms

Part 4.2 Orthogonality

4.2.1 Definition of orthogonal vectors

4.2.2 Orthogonal matrices and their properties

4.2.3 Applications of orthogonality

Part 4.3 Applications of quadratic forms and orthogonality

4.3.1 Optimisation problems

4.3.2 Principal component analysis (PCA)

4.3.3 Applications in engineering and computer science

Introduction

Quadratic forms and orthogonality are powerful tools in linear algebra with wide-ranging applications. A quadratic form is a polynomial of degree two in several variables that can be expressed using matrices. It provides insight into optimisation problems and classification of matrices.

Orthogonality refers to the concept of perpendicularity in vector spaces. Orthogonal vectors and matrices play a crucial role in simplifying computations, preserving lengths and angles, and enabling efficient transformations.

The aim of this Series is to enable you to define and apply quadratic forms and orthogonality, understand their matrix representations, and explore their applications in optimisation, statistics, engineering, and computer science.

We shall commence this Series by examining quadratic forms, their definitions, matrix representations, and applications. Next, we will study orthogonality, including orthogonal vectors and matrices, and their properties. Finally, we will conclude with applications of quadratic forms and orthogonality in optimisation, PCA, and applied fields.

Series Learning Outcomes

By the end of this Series, you should be able to:

SLO 4.1 define quadratic forms and explain their matrix representation,

SLO 4.2 compute quadratic forms for given matrices and variables,

SLO 4.3 analyse applications of quadratic forms in optimisation and classification, SLO 4.4 define orthogonality of vectors and explain its geometric meaning,

SLO 4.5 describe properties of orthogonal matrices and demonstrate their use,

SLO 4.6 apply orthogonality in simplifying computations and transformations,

SLO 4.7 evaluate applications of quadratic forms and orthogonality in PCA, engineering, and computer science.

4.1 Quadratic Forms

4.1.1 Definition of quadratic forms

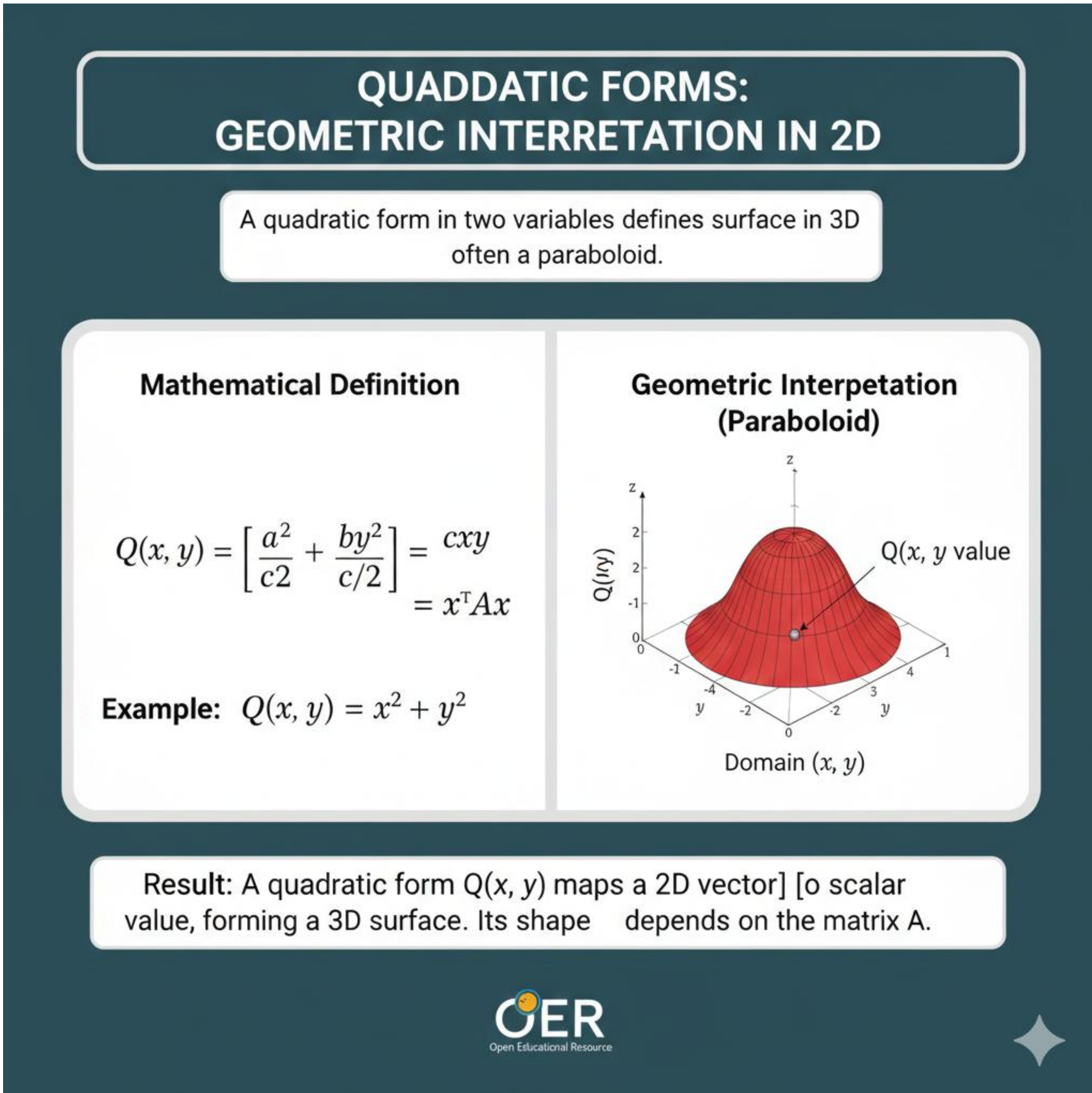
A quadratic form is a polynomial of degree two in several variables that can be expressed using matrices. For example, in two variables:

$$Q(x,y)=ax2+bxy+cy2$$

This can be written in matrix form as:

$$Q(x)=xTAx,$$

where $x=x\ y$ and Ais a symmetric matrix.



Quadratic form represented as a surface in 2D

Fill in the blank: A quadratic form in matrix notation is written as xTAx, where Ais a ______ matrix.

4.1.2 Matrix representation of quadratic forms

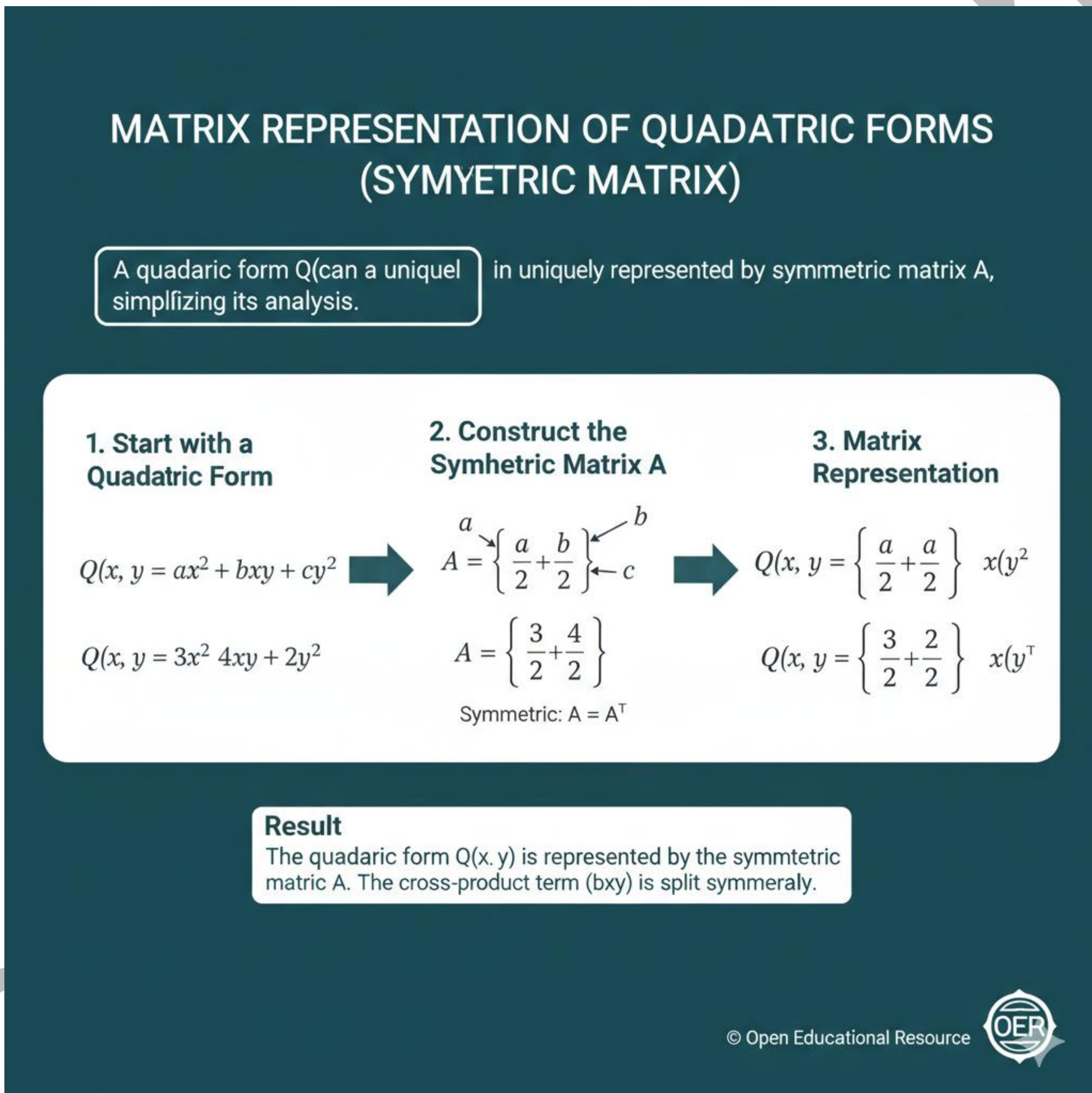
Quadratic forms can always be represented using symmetric matrices. For example:

$$Q(x,y,z)=2x2+3y2+4z2+2xy+yz$$

can be expressed as:

$$Q(x)=xTAx,x=x\ y\ z\ ,A=2\ 1\ 0\ 1\ 3\ 12\ 0\ 12\ 4\ .$$

This representation makes it easier to analyse properties such as definiteness (positive definite, negative definite, or indefinite).



Matrix representation of quadratic form

Fill in the blank: Quadratic forms are always represented using ______ matrices.

4.1.3 Applications of quadratic forms

Quadratic forms are widely applied in:

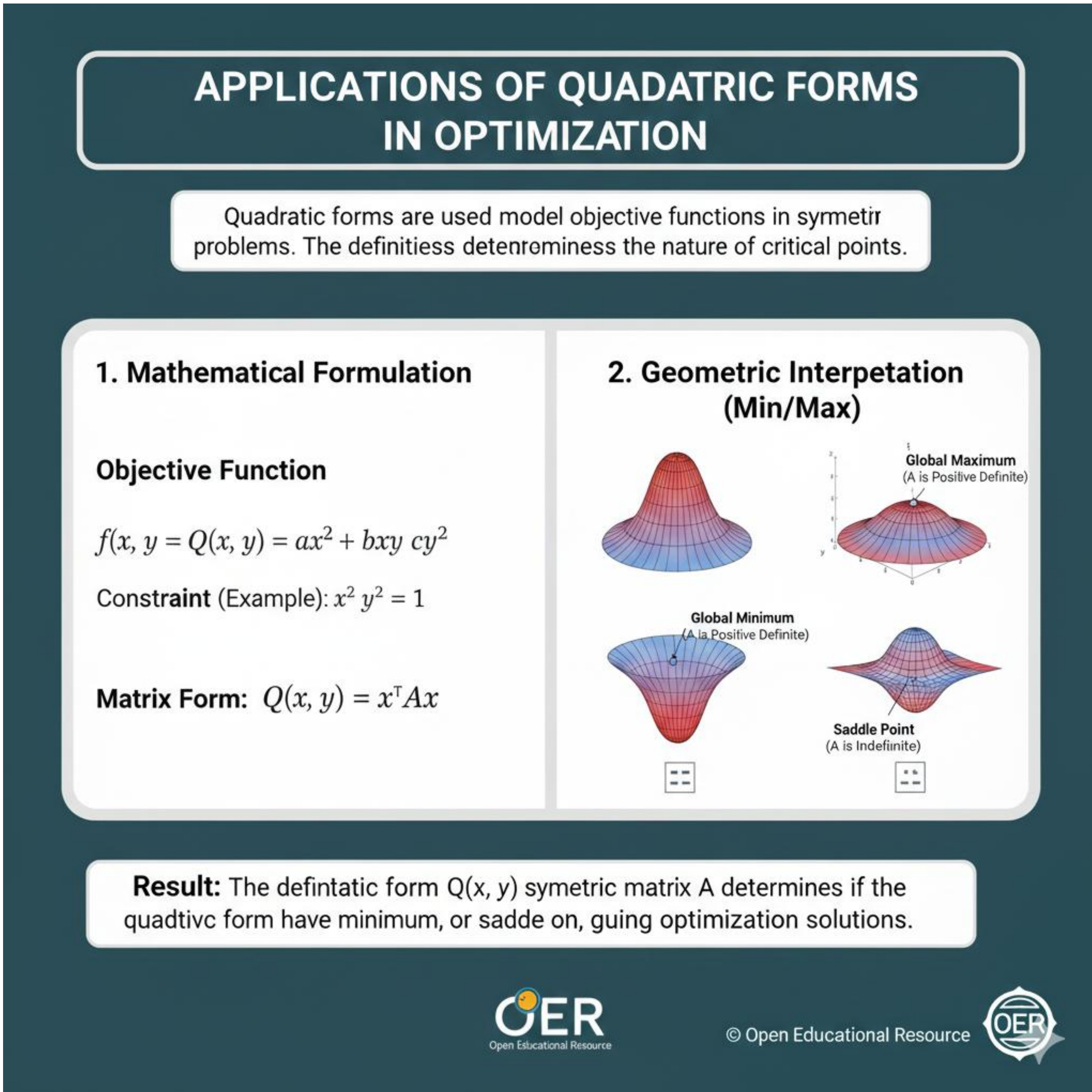
Optimisation: Determining maxima and minima of functions.

Statistics: Variance and covariance analysis.

Economics: Modelling cost functions.

Engineering: Analysing energy and stability in systems.

For example, in optimisation, quadratic forms help determine whether a critical point is a minimum, maximum, or saddle point by examining the definiteness of the matrix.



Application of quadratic forms in optimisation

Fill in the blank: In optimisation, quadratic forms are used to determine whether a critical point is a minimum, maximum, or ______ point.

Pilot Questions 4.1

A quadratic form in matrix notation is expressed as: A. xTAxB. AxC. xTxD. ATx

The matrix used to represent quadratic forms is always: A. Diagonal B. Symmetric C. Singular D. Orthogonal

Which of the following is an application of quadratic forms? A. Optimisation B. Statistics C. Engineering D. All of the above

In optimisation, quadratic forms help determine: A. Eigenvalues B. Critical points (minima, maxima, saddle points) C. Determinants D. Trace

A quadratic form in two variables can be represented as: A. A line B. A paraboloid surface C. A circle D. A cube

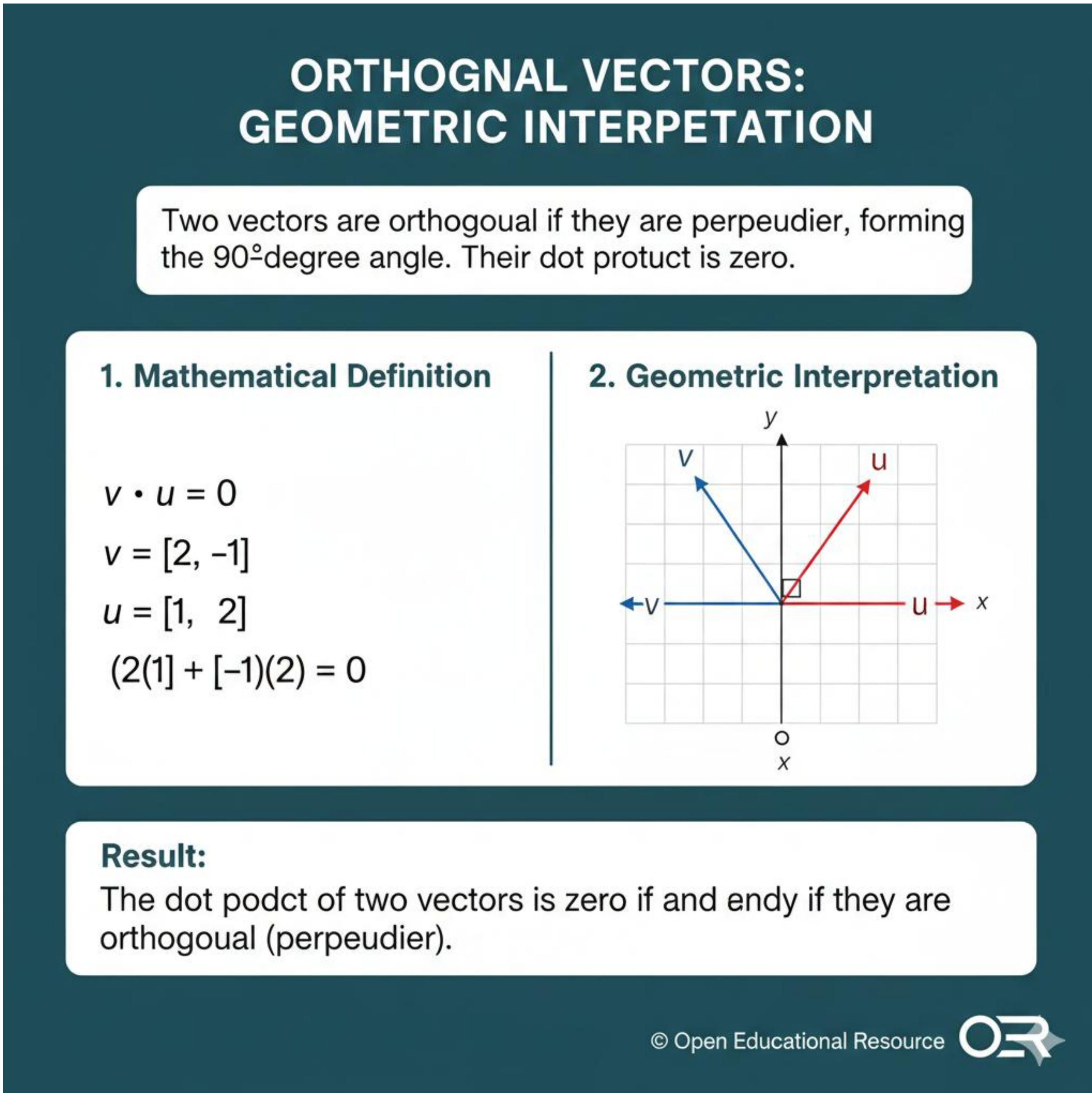
4.2 Orthogonality

4.2.1 Definition of orthogonal vectors

Two vectors are said to be **orthogonal** if their dot product is zero:

$$u \cdot v = 0.$$

This means the vectors are perpendicular to each other in the geometric sense. Orthogonality generalises the concept of perpendicularity to higher-dimensional vector spaces.



Orthogonal vectors in 2D space

Fill in the blank: Two vectors are orthogonal if their ______ product is zero.

4.2.2 Orthogonal matrices and their properties

A matrix Qis **orthogonal** if:

$$QTQ=I,$$

where QTis the transpose of Qand Iis the identity matrix.

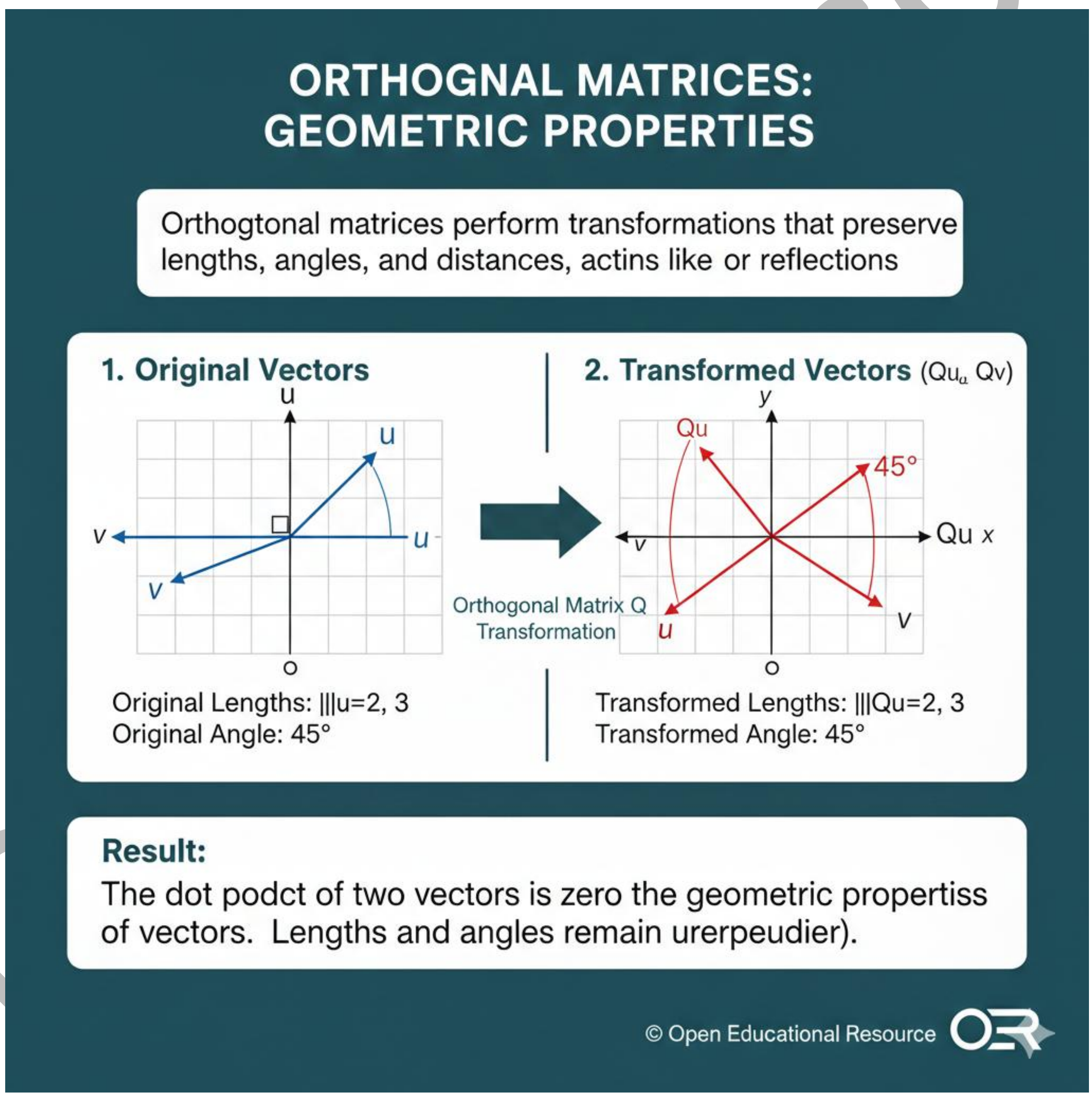
Properties of orthogonal matrices:

Columns (and rows) are orthonormal vectors.

They preserve lengths and angles.

The inverse of an orthogonal matrix is its transpose: Q-1=QT.

Determinant is either +1 or -1.



Properties of orthogonal matrices

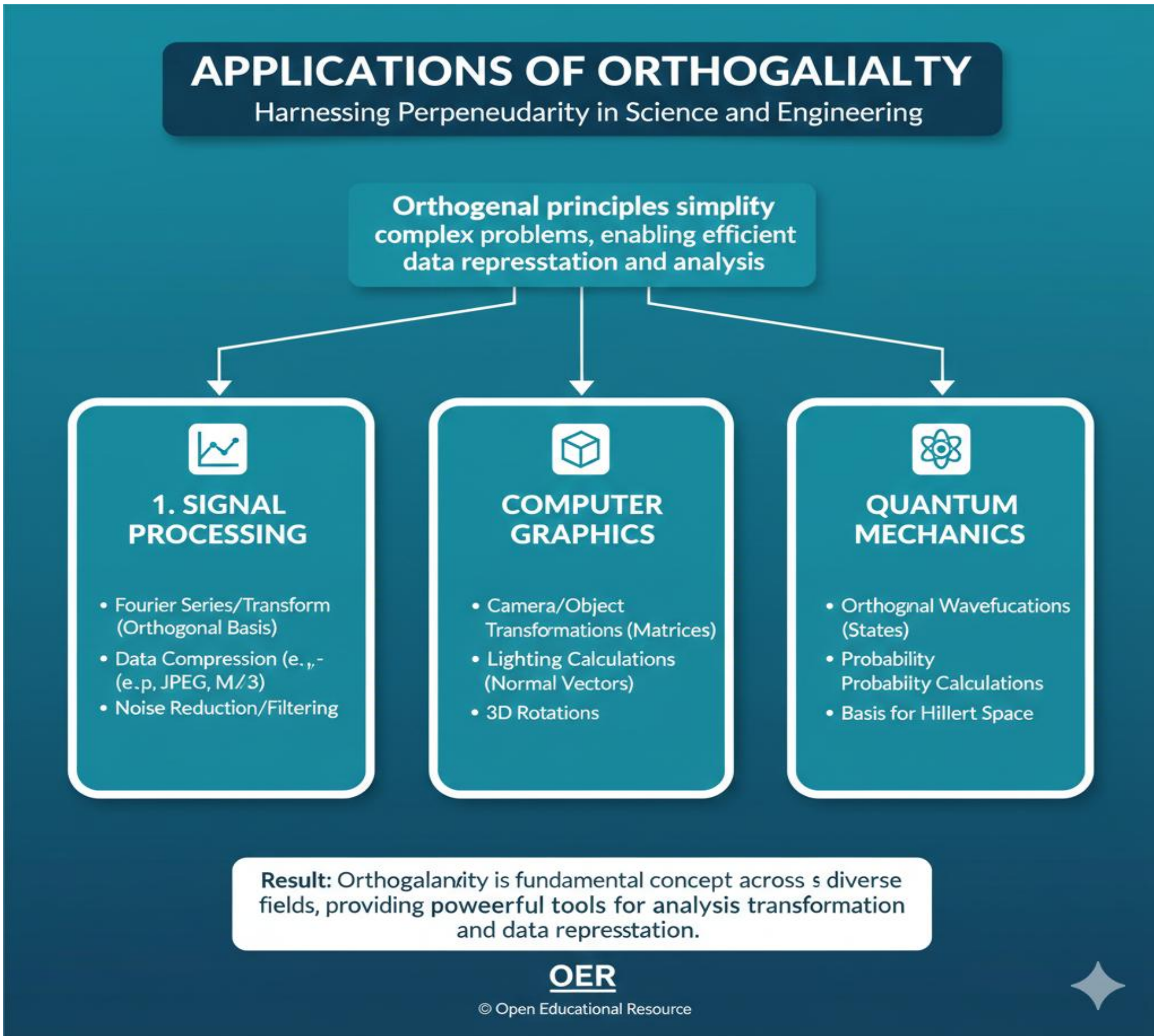
Fill in the blank: An orthogonal matrix satisfies the condition Q^T Q = ______.

4.2.3 Applications of orthogonality

Orthogonality is widely applied in mathematics and science:

- **Simplifying computations**: Orthogonal bases make calculations easier.
- **Signal processing**: Fourier transforms rely on orthogonal functions.
- **Computer graphics**: Orthogonal matrices are used in rotations and reflections.
- **Statistics**: Orthogonal transformations are used in regression and PCA.

By using orthogonal vectors and matrices, we can preserve geometric properties while simplifying analysis.



Applications of orthogonality

Fill in the blank: In computer graphics, orthogonal matrices are used in ______ and reflections.

Pilot Questions 4.2

Two vectors are orthogonal if: A. Their dot product is zero B. Their determinant is zero C. Their trace is zero D. Their sum is zero

An orthogonal matrix satisfies: A. QTQ=IB. QTQ=QC. QTQ=0D. QTQ=QT

Which property is true of orthogonal matrices? A. They preserve lengths and angles B. Their inverse is equal to their transpose C. Their determinant is ±1 D. All of the above

Orthogonality is applied in: A. Signal processing B. Computer graphics C. Statistics D. All of the above

In Fourier transforms, orthogonality is used to: A. Preserve eigenvalues B. Simplify computations with orthogonal functions C. Eliminate determinants D. Avoid diagonalisation

4.3 Applications Of Quadratic Forms And Orthogonality

4.3.1 Optimisation problems

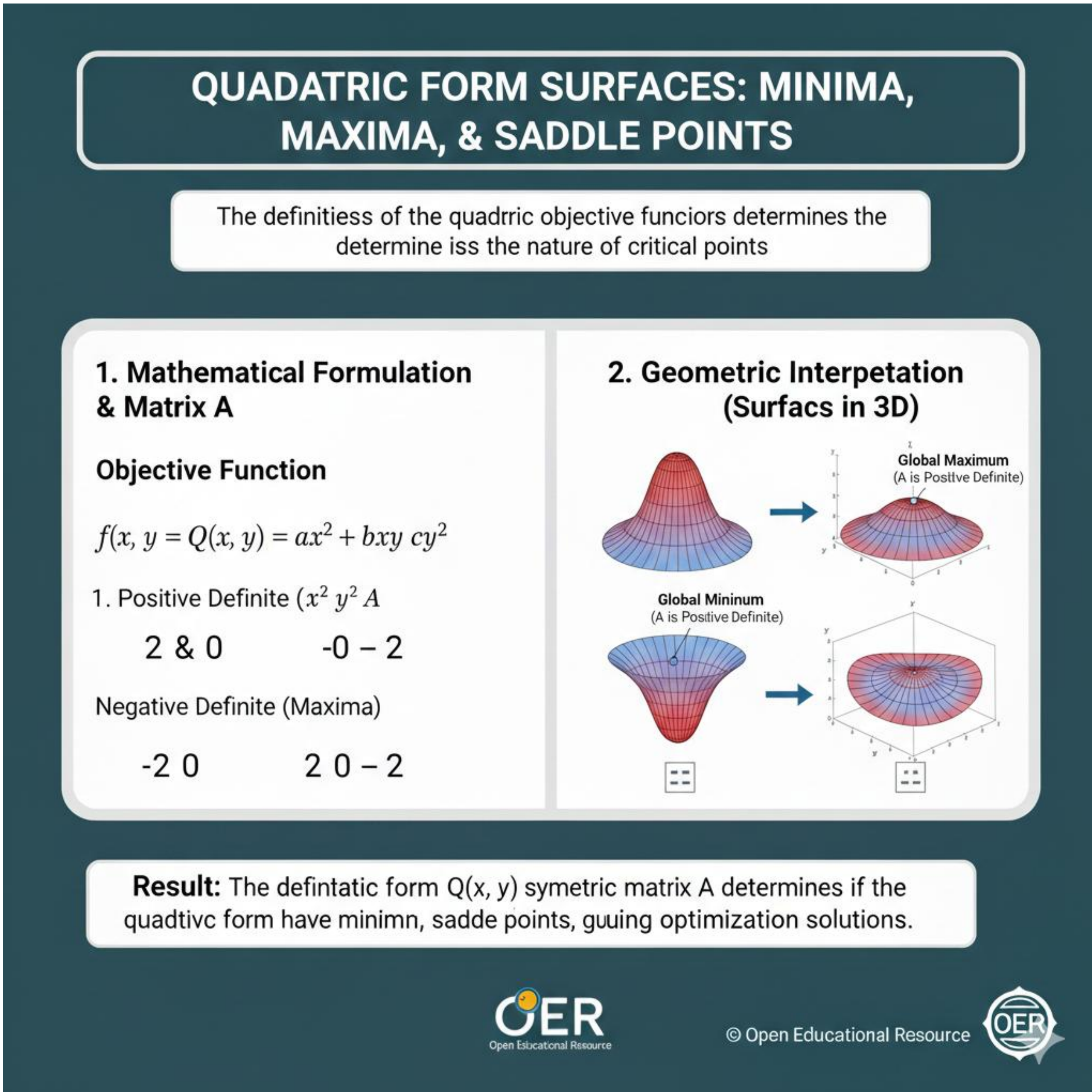
Quadratic forms are central to optimisation. By examining the definiteness of the matrix in a quadratic form, we can determine whether a critical point is a minimum, maximum, or saddle point.

Positive definite matrix → minimum point.

Negative definite matrix → maximum point.

Indefinite matrix → saddle point.

This is widely used in economics (cost minimisation), engineering (energy optimisation), and operations research.



Quadratic forms in optimisation problems

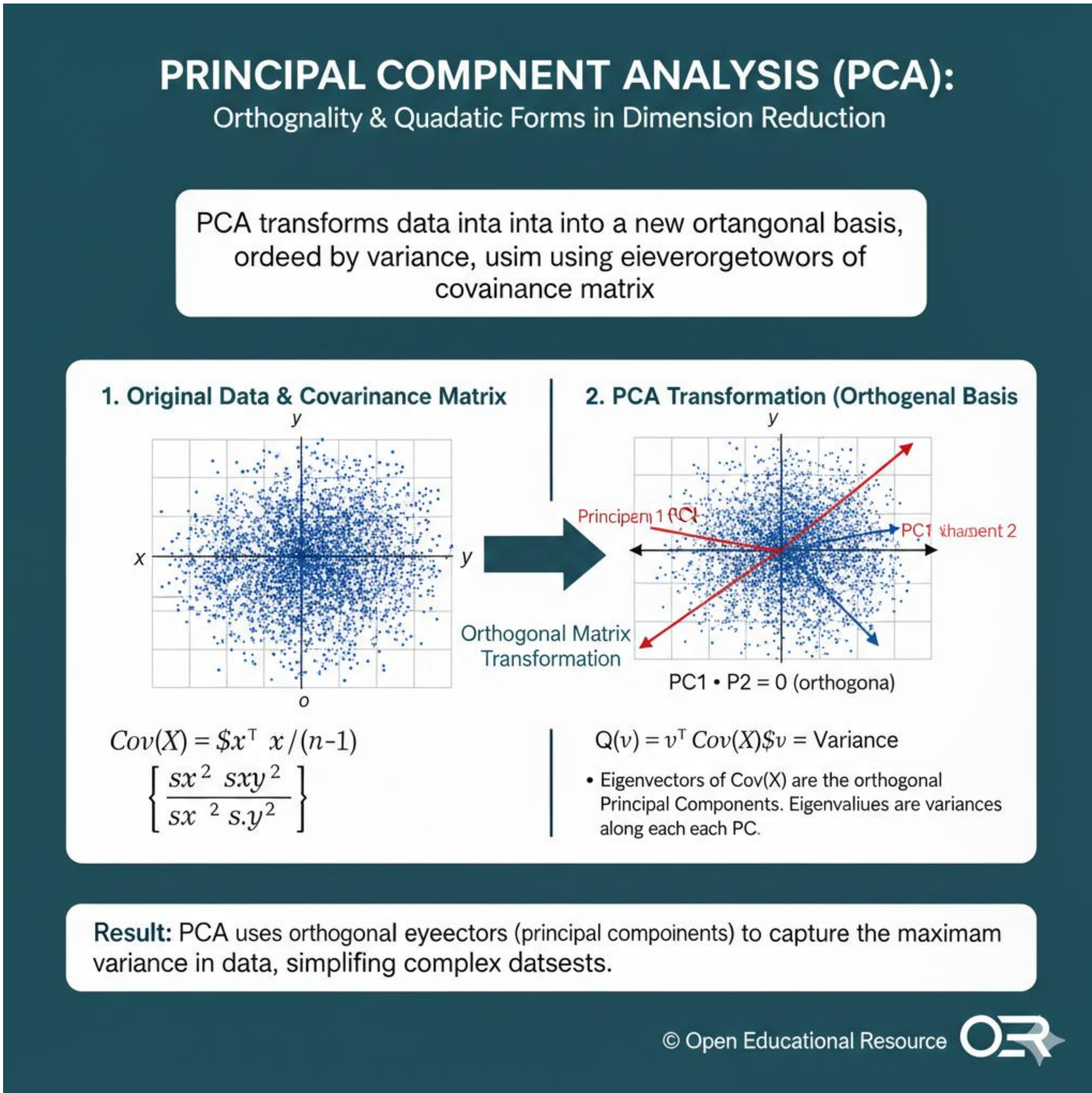
Fill in the blank: A positive definite matrix in a quadratic form corresponds to a ______ point in optimisation.

4.3.2 Principal component analysis (PCA)

Orthogonality and quadratic forms are key to **PCA**, a statistical technique for dimensionality reduction. PCA uses eigenvalues and eigenvectors of covariance matrices (quadratic forms) to identify principal components. These components are orthogonal directions that capture maximum variance in data.

Applications include:

- Data compression.
- Noise reduction.
- Feature extraction in machine learning.



PCA using quadratic forms and orthogonality

Fill in the blank: PCA identifies orthogonal directions that capture maximum ______ in data.

4.3.3 Applications in engineering and computer science

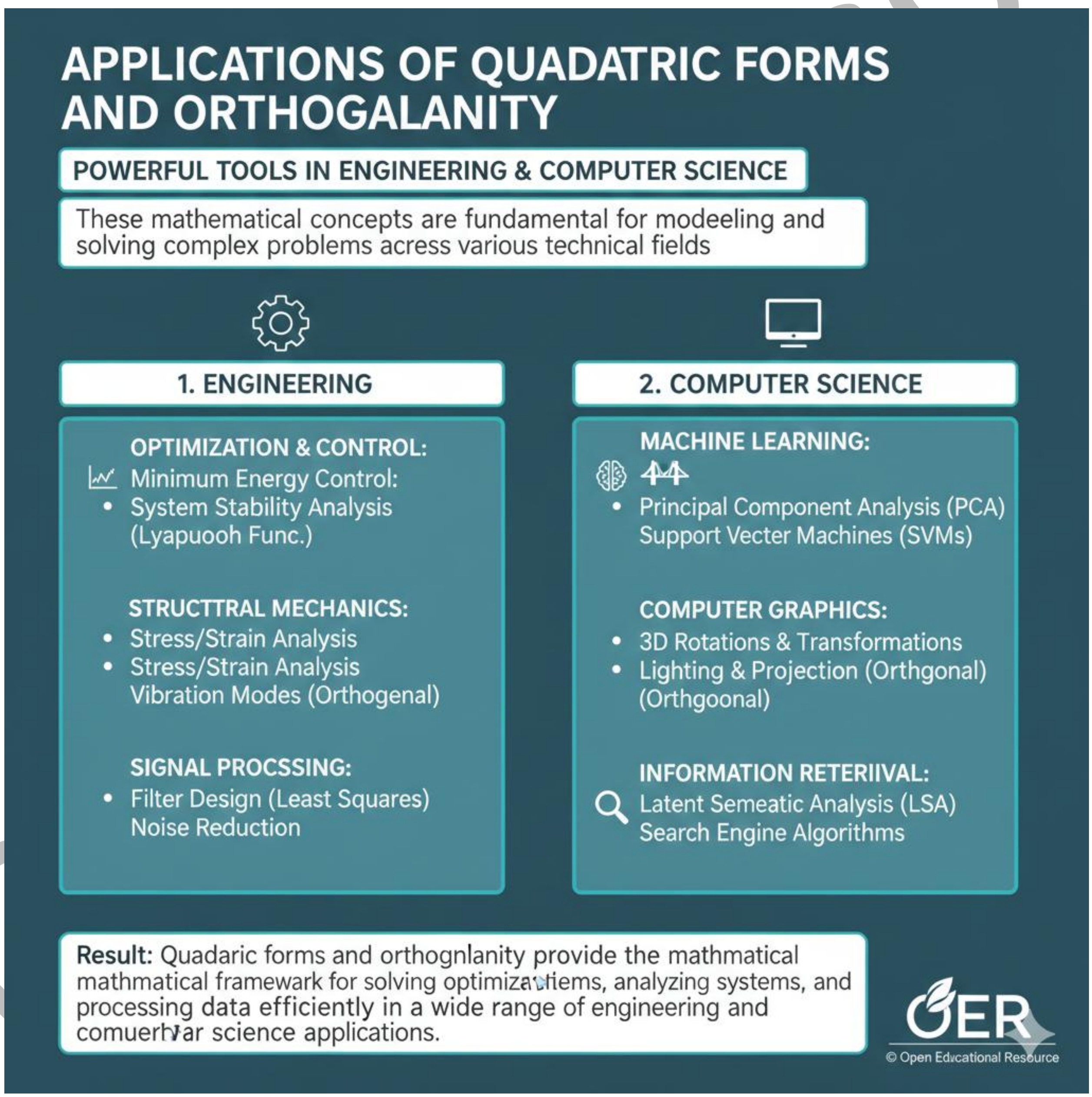
Quadratic forms and orthogonality are applied in many practical fields:

Engineering: Analysing vibrations, energy, and stability of structures.

Computer science: Graphics transformations using orthogonal matrices, machine learning algorithms using PCA.

Signal processing: Orthogonal functions used in Fourier transforms.

These applications highlight how abstract mathematical concepts directly impact real-world technologies.



Applications of quadratic forms and orthogonality in applied fields

Fill in the blank: In computer science, orthogonal matrices are widely used in graphics transformations such as ______.

Pilot Questions 4.3

- A positive definite quadratic form corresponds to: A. Maximum point B. Minimum point C. Saddle point D. No critical point
- PCA uses: A. Eigenvalues and eigenvectors of covariance matrices B. Determinants of matrices C. Trace of matrices D. Orthogonal matrices only
- Orthogonal principal components in PCA capture: A. Minimum variance B. Maximum variance C. Random noise D. Eigenvalues only
- Quadratic forms and orthogonality are applied in: A. Engineering B. Computer science C. Signal processing D. All of the above
- In computer graphics, orthogonal matrices are used in: A. Rotations and reflections B. Diagonalisation C. Regression analysis D. Probability distributions

Series Summary

This Series explored the concepts of **quadratic forms** and **orthogonality**, both of which are fundamental in linear algebra and widely applied across mathematics, science, and engineering. We began with quadratic forms, defining them, expressing them in matrix notation, and examining their applications in optimisation and statistics. We then studied orthogonality, focusing on orthogonal vectors, orthogonal matrices, and their properties. Finally, we examined applications of quadratic forms and orthogonality in optimisation problems, PCA, engineering, and computer science.

Glossary of Terms

- **Quadratic form**: A polynomial of degree two in several variables, expressed as xTAx.
- **Symmetric matrix**: A matrix equal to its transpose, used to represent quadratic forms.
- **Definiteness**: Property of a matrix (positive definite, negative definite, indefinite) that determines the nature of critical points in optimisation.

Orthogonality: Condition where two vectors are perpendicular, i.e., their dot product is zero.

Orthogonal matrix: A matrix Qsatisfying QTQ=I, with orthonormal columns and rows.

Principal component analysis (PCA): A statistical technique that uses eigenvalues, eigenvectors, and orthogonality to reduce dimensionality of data.

Fourier transform: A mathematical tool in signal processing that relies on orthogonal functions.

Concept Check Questions 4

Objective questions

A quadratic form in matrix notation is expressed as: A. xTAxB. AxC. xTxD. ATx(Linked to SLO 4.1)

An orthogonal matrix satisfies: A. QTQ=IB. QTQ=QC. QTQ=0D. QTQ=QT(Linked to SLO 4.5)

Short-answer questions

Define orthogonality of vectors and explain its geometric meaning. (Linked to SLO 4.4)

Explain how quadratic forms are used in optimisation problems. (Linked to SLO 4.3)

Long-answer questions

Discuss applications of quadratic forms and orthogonality in PCA, engineering, and computer science. (Linked to SLO 4.7)

Basic response: Describe PCA and one engineering application.

Value-added response: Provide detailed examples such as PCA in machine learning, orthogonal matrices in graphics transformations, and quadratic forms in stability analysis.

Answers to Concept Check Questions 4

A. xTAx

A. QTQ=I

Orthogonality means two vectors are perpendicular, with dot product equal to zero. Geometrically, they form a right angle in space.

Quadratic forms are used to classify critical points in optimisation: positive definite matrices yield minima, negative definite matrices yield maxima, and indefinite matrices yield saddle points.

Basic response: PCA uses orthogonal principal components to reduce dimensionality. In engineering, quadratic forms analyse vibrations and stability.

Value-added response: PCA is applied in machine learning for feature extraction. Orthogonal matrices are used in computer graphics for rotations and reflections. Quadratic forms are applied in energy optimisation and structural stability analysis.

Answers to Pilot Questions 4

From Part 4.1

xTAx

Symmetric

All of the above

Critical points (minima, maxima, saddle points)

A paraboloid surface

From Part 4.2

Their dot product is zero

QTQ=I

All of the above

All of the above

Simplify computations with orthogonal functions

From Part 4.3

Minimum point

Eigenvalues and eigenvectors of covariance matrices

Maximum variance

All of the above

Rotations and reflections

Series 5 Inner Product Spaces And Norms

Series Outline

For this 1-credit course, Series 5 will be structured into three Parts:

Part 5.1 Inner product spaces

5.1.1 Definition of inner product

5.1.2 Properties of inner product spaces

5.1.3 Examples of inner product spaces

Part 5.2 Norms

5.2.1 Definition of norms

5.2.2 Relationship between norms and inner products

5.2.3 Types of norms (Euclidean, p-norms, etc.)

Part 5.3 Applications of inner products and norms

5.3.1 Orthogonality and projections

5.3.2 Least squares problems

5.3.3 Applications in data science and engineering

Introduction

Inner product spaces and norms extend the concepts of orthogonality and length to more general vector spaces. An inner product generalises the dot product, allowing us to measure angles and lengths in abstract spaces. A norm is a function that measures the size or length of vectors, often derived from an inner product.

Together, these concepts provide the foundation for advanced topics such as orthogonal projections, least squares approximation, and applications in data science, engineering, and physics.

The aim of this Series is to enable you to define and apply inner products and norms, understand their properties, and explore their applications in solving practical problems.

Series Learning Outcomes

By the end of Series 5: Inner Product Spaces and Norms, you should be able to:

- SLO 5.1 define inner products and explain their role in generalising the dot product,
- SLO 5.2 describe and verify the properties of inner product spaces (linearity, symmetry, positivity),
- SLO 5.3 compute inner products in different vector spaces and provide examples,
- SLO 5.4 define norms and explain their relationship to inner products,
- SLO 5.5 calculate norms of vectors using Euclidean and other p-norms,
- SLO 5.6 apply inner products and norms to establish orthogonality and projections,
- SLO 5.7 solve least squares problems using inner product and norm concepts,
- SLO 5.8 evaluate applications of inner products and norms in data science, engineering, and physics.

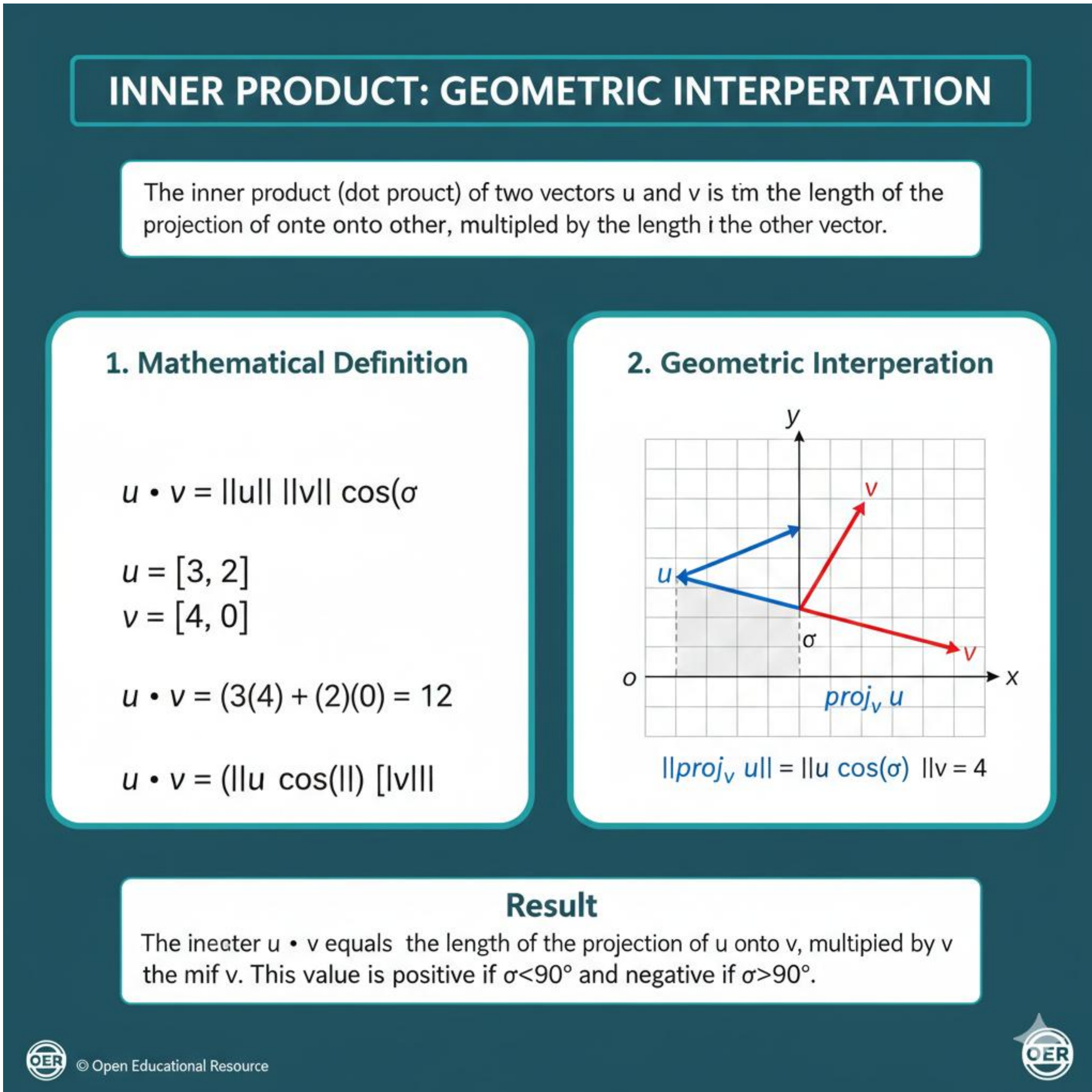
5.1 Inner Product Spaces

5.1.1 Definition of inner product

An inner product is a generalisation of the dot product that allows us to measure angles and lengths in abstract vector spaces. For vectors u,v∈Rn, the standard inner product is:

$$\langle u,v\rangle=uTv=i=1nuivi.$$

This operation returns a scalar and provides the foundation for defining orthogonality and norms.



Inner product as projection of one vector onto another

Fill in the blank: The inner product of two vectors generalises the concept of the ______ product.

5.1.2 Properties of inner product spaces

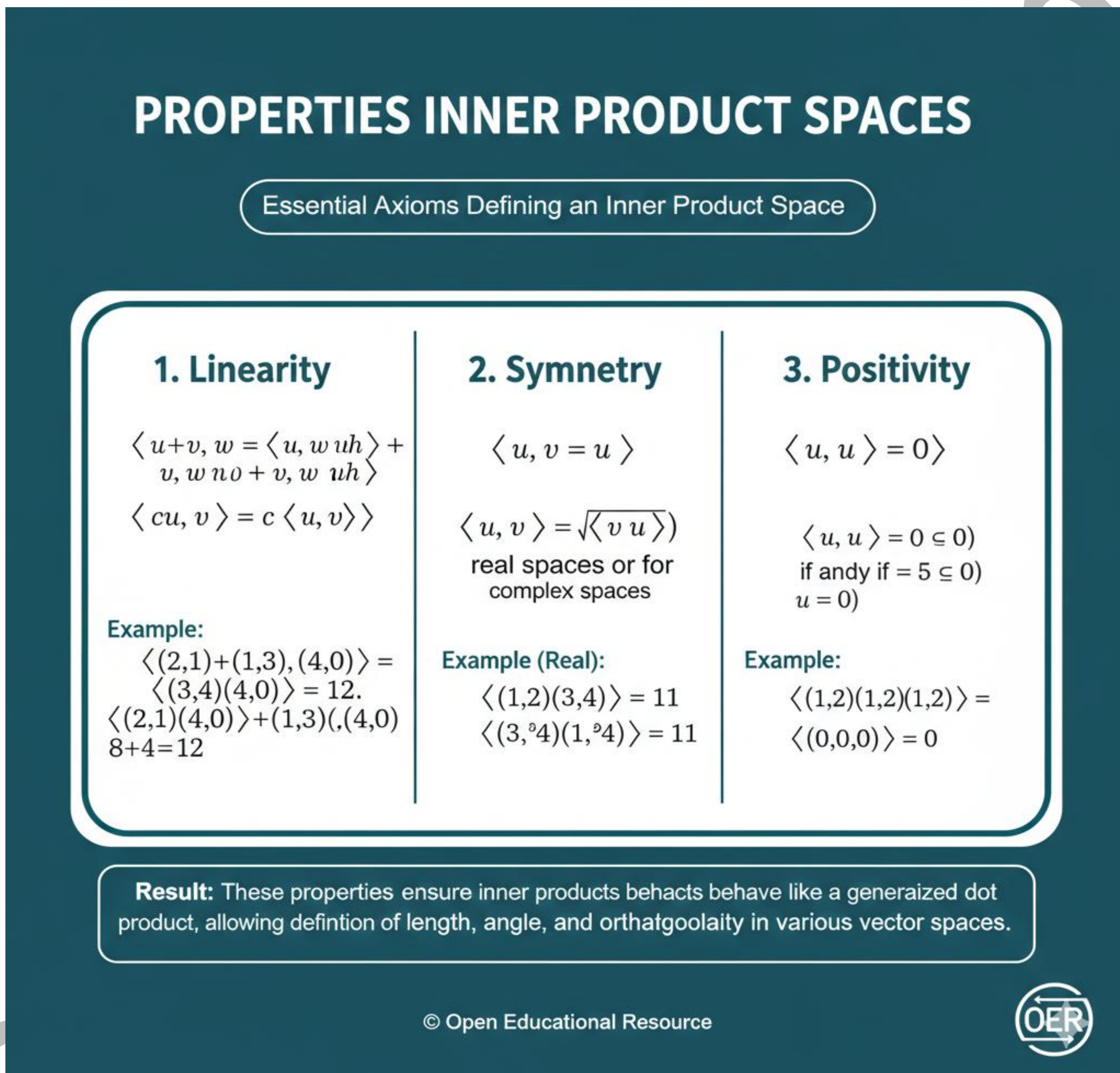
Inner products must satisfy the following properties:

Linearity: $\langle au+bv,w\rangle=a\langle u,w\rangle+b\langle v,w\rangle$.

Symmetry: $\langle u,v\rangle=\langle v,u\rangle$.

Positivity: ⟨v,v⟩≥0, with equality only if v=0.

These properties ensure that inner products behave consistently across different vector spaces.



Properties of inner product

Fill in the blank: The property ⟨v,v⟩≥0is known as ______.

5.1.3 Examples of inner product spaces

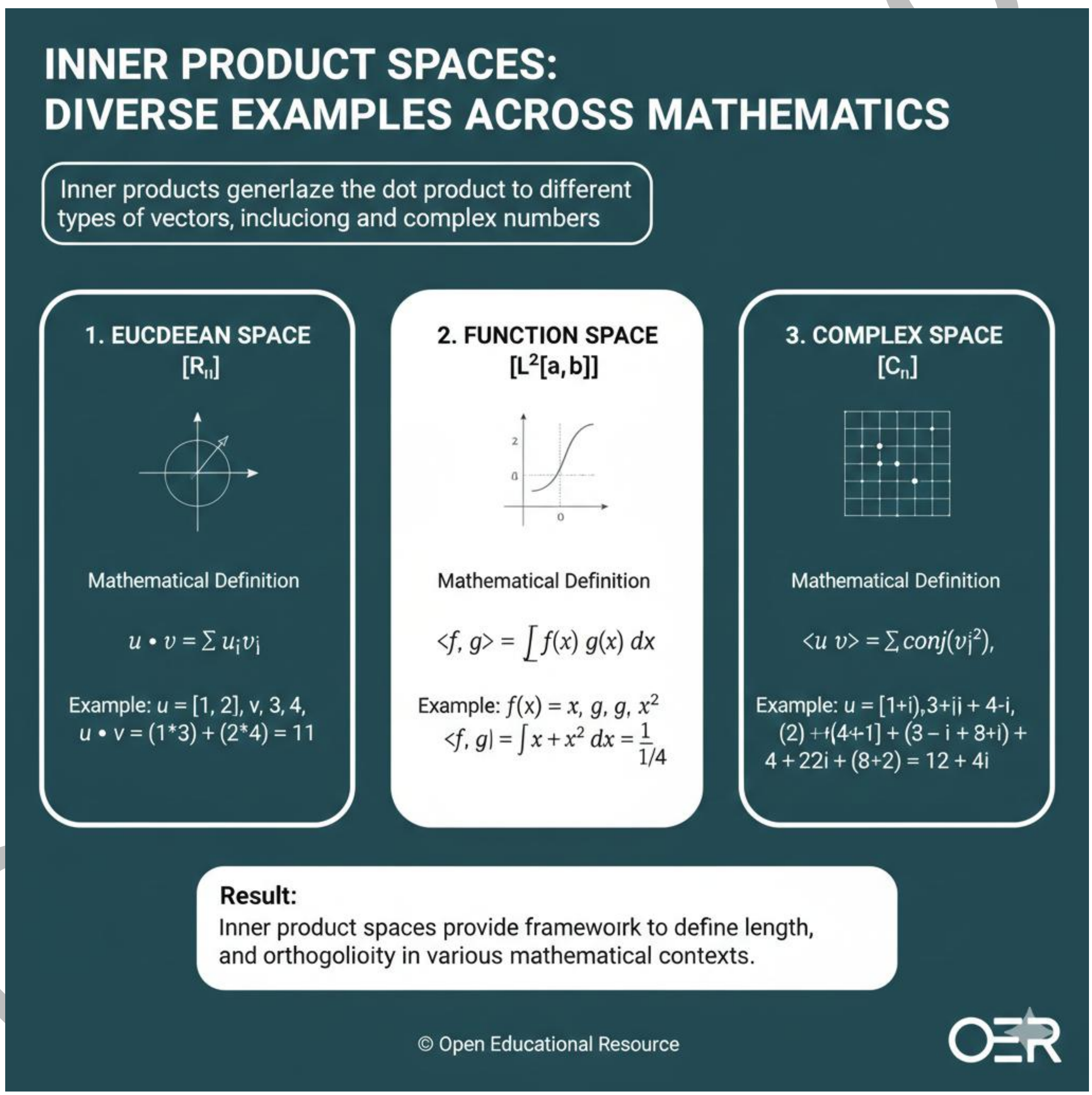
Examples include:

Euclidean space: Standard dot product in Rn.

Function spaces: Inner product defined as ⟨f,g⟩=abf(x)g(x) dx.

Complex vector spaces: Inner product defined with conjugation: ⟨u,v⟩=∑uivi‾.

These examples show how inner products extend beyond simple vectors to functions and complex numbers.



Examples of inner product spaces

Fill in the blank: In function spaces, the inner product is defined using an ______ over an interval.

Pilot Questions 5.1

- The inner product of two vectors in Rnis: A. A vector B. A scalar C. A matrix D. A polynomial
- Which property is NOT required for an inner product? A. Linearity B. Symmetry C. Positivity D. Determinant
- The property ⟨v,v⟩≥0ensures: A. Orthogonality B. Positivity C. Linearity D. Symmetry
- In complex vector spaces, the inner product uses: A. Conjugation B. Differentiation C. Integration D. Determinants
- In function spaces, the inner product is defined as: A. A determinant B. An integral of the product of functions C. A trace of a matrix D. A sum of eigenvalues

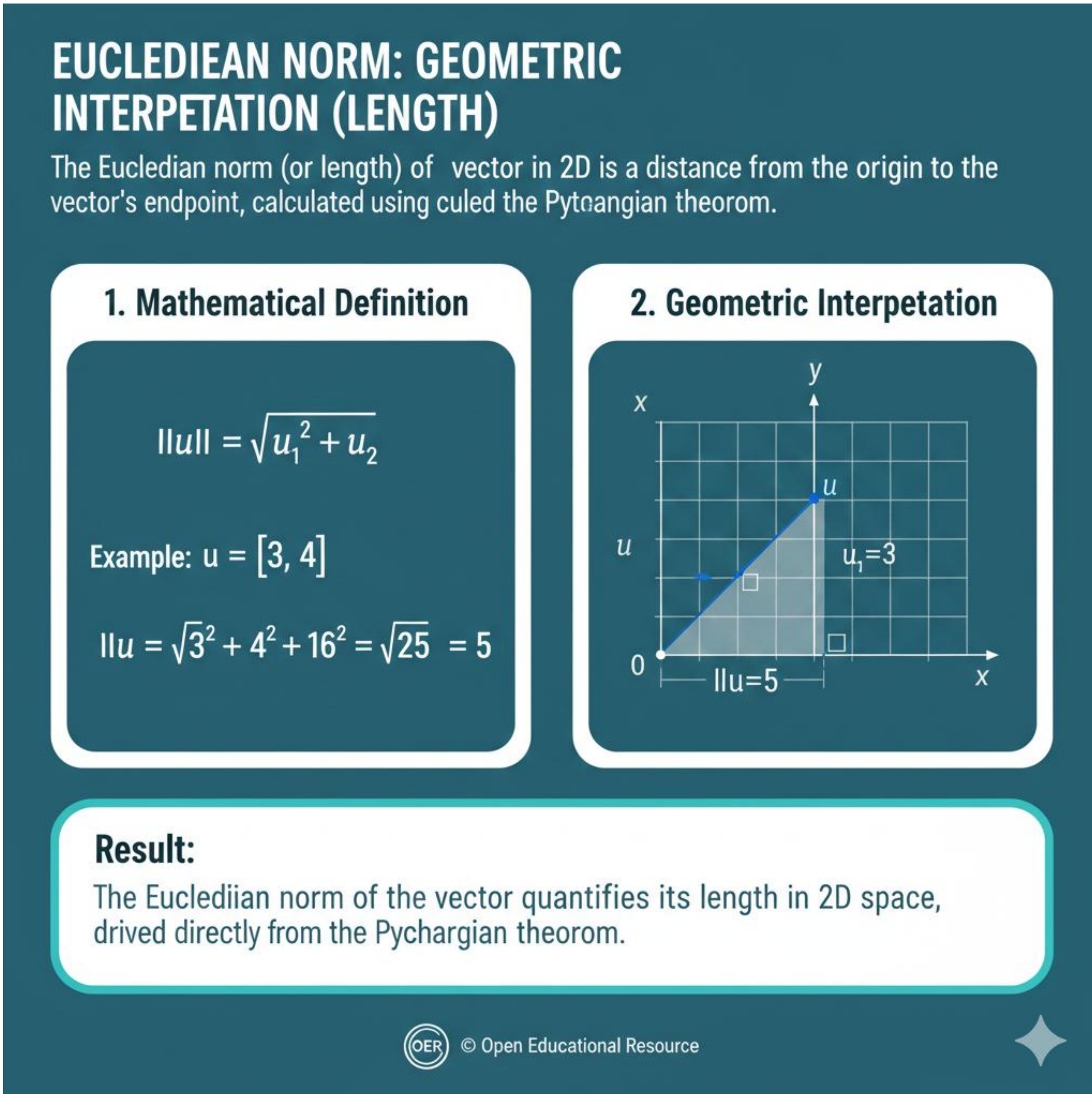
5.2 Norms

5.2.1 Definition of norms

A norm is a function that assigns a non-negative length or size to a vector. For a vector v, the norm is denoted as v.

The most common norm is the Euclidean norm (or 2-norm):

$$v2=v12+v22+\cdots+vn2.$$



Euclidean norm as vector length

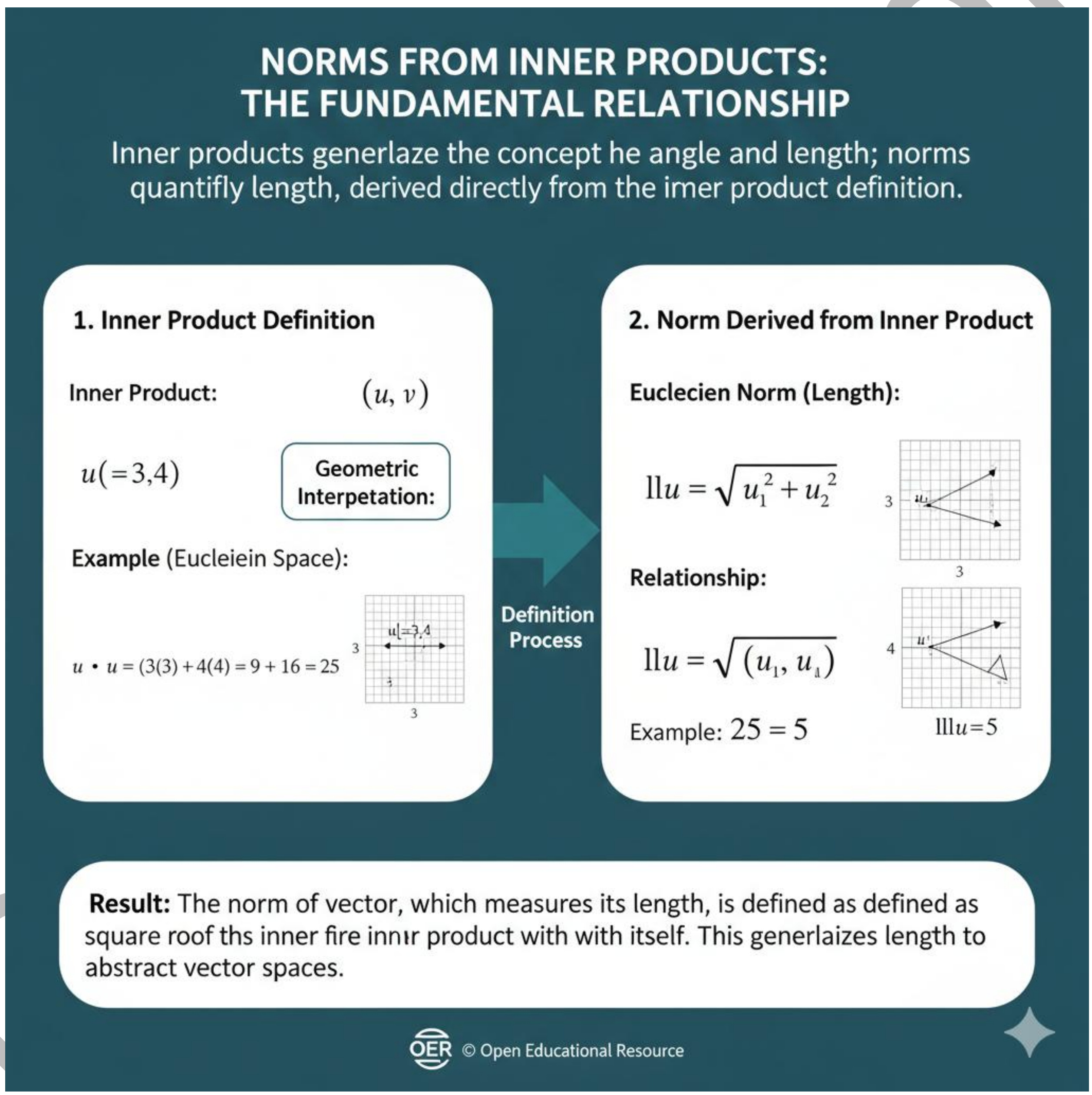
Fill in the blank: A norm is a function that assigns a non-negative ______ to a vector.

5.2.2 Relationship between norms and inner products

Norms can be derived from inner products. In an inner product space, the norm of a vector vis defined as:

$$v=\langle v,v\rangle.$$

This shows that norms measure the "size" of vectors, while inner products measure angles and projections.



Norm derived from inner product

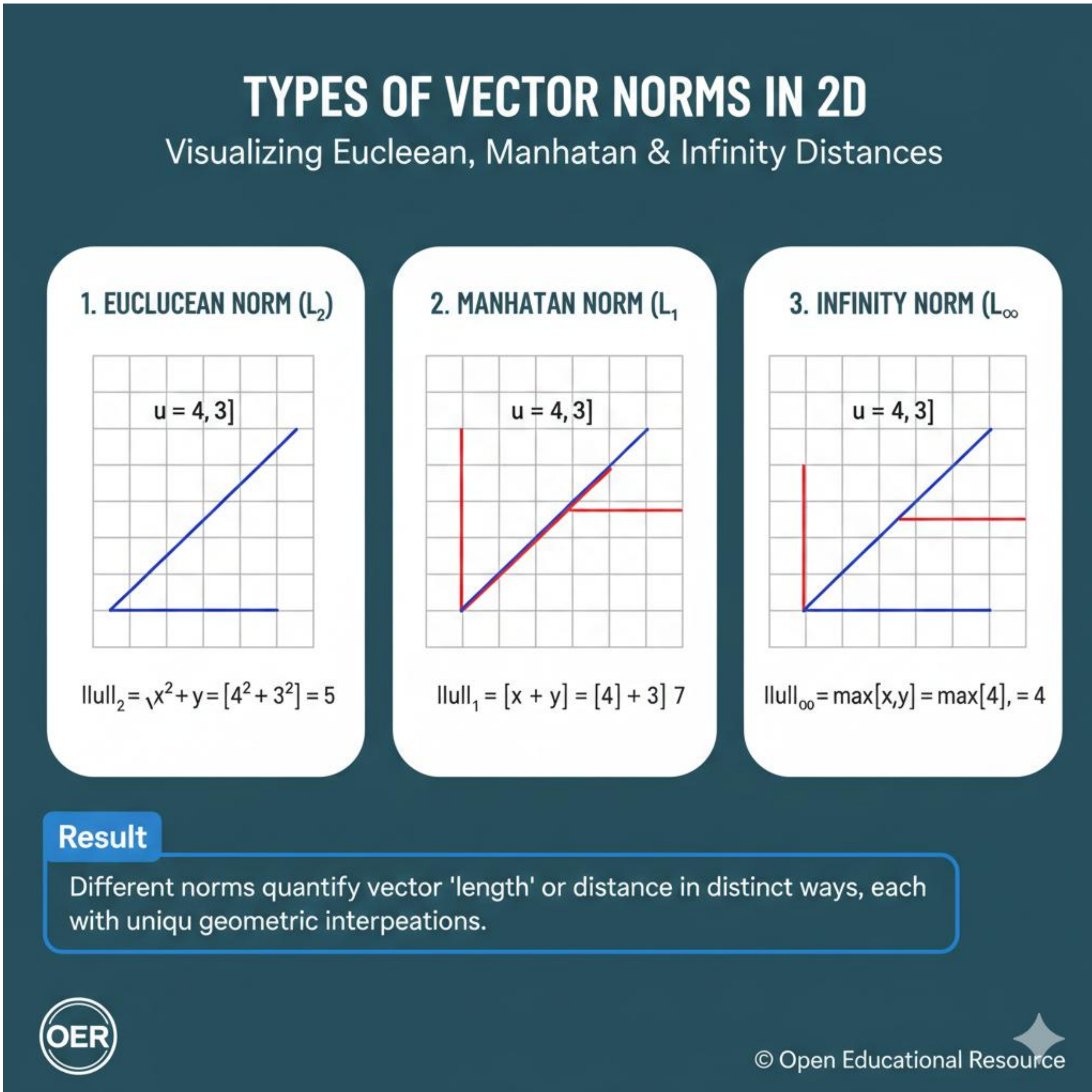
Fill in the blank: In an inner product space, the norm of a vector is defined as the square root of \langle v, ______ \rangle.

5.2.3 Types of norms

There are several types of norms:

- Euclidean norm (2-norm): $v2=\sum vi2$.
- 1-norm (Manhattan norm): $v1=\sum|vi|$.
- Infinity norm (max norm): $v=\max|vi|$.
- p-norm: $vp=|vi|p1/p$.

Each norm has different applications depending on the context, such as optimisation, machine learning, and numerical analysis.



Comparison of different norms in 2D space

Fill in the blank: The infinity norm of a vector is defined as the ______ of the absolute values of its components.

Pilot Questions 5.2

A norm assigns to a vector: A. A matrix B. A scalar length C. A polynomial D. A determinant

The Euclidean norm of vector v=(3,4)is: A. 5 B. 7 C. 12 D. 25

In an inner product space, the norm of a vector is defined as: A. ⟨v,v⟩B. ⟨v,v⟩C. ⟨v,u⟩D. ∑vi

The 1-norm of vector v=(2,-3)is: A. 1 B. 5 C. 6 D. -1

The infinity norm of vector v=(2,-7,4)is: A. 2 B. 4 C. 7 D. -7

5.3 Applications of Inner Products and Norms

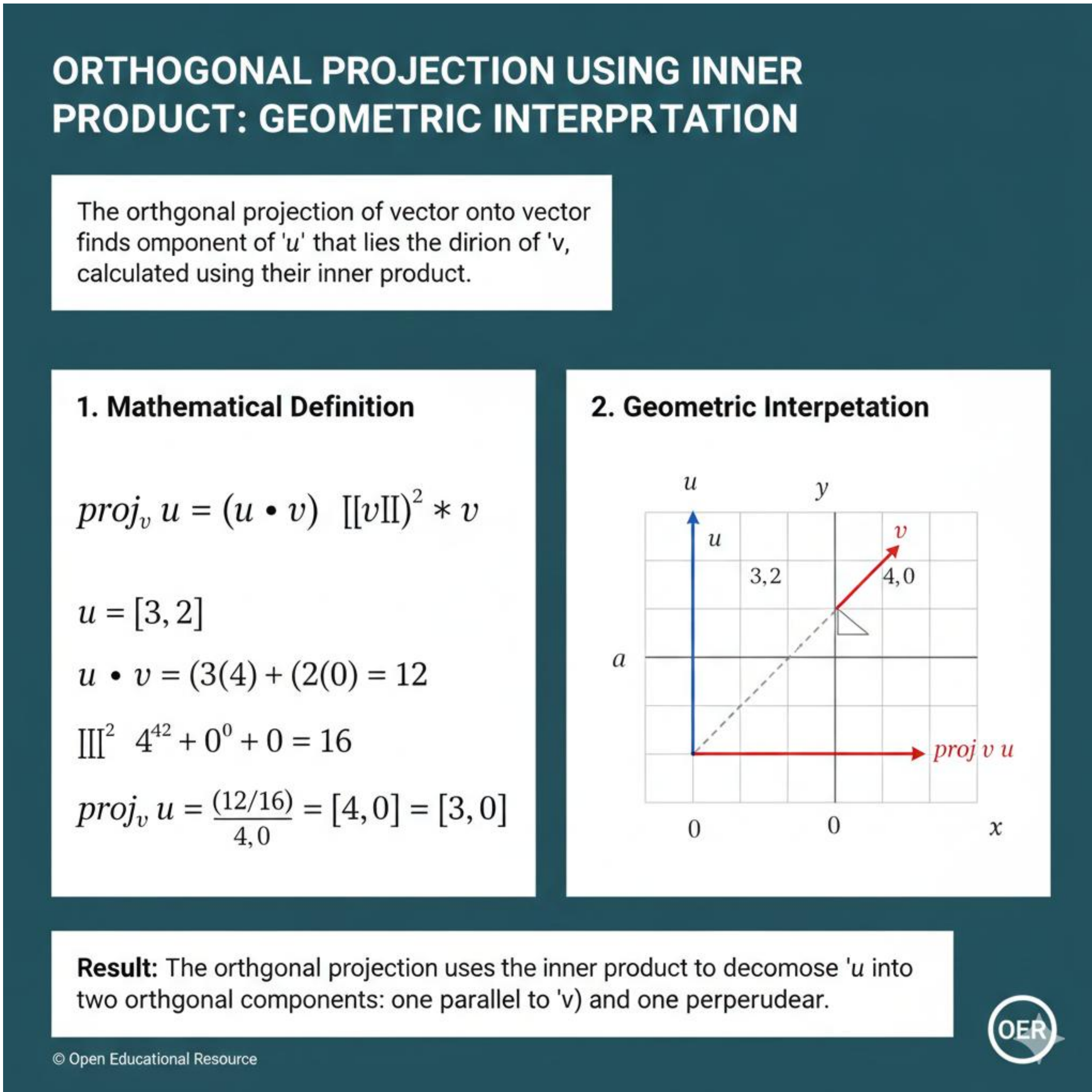
5.3.1 Orthogonality and projections

Inner products allow us to define orthogonality in abstract spaces. Two vectors are orthogonal if their inner product is zero.

Projections are also defined using inner products:

$$\text{proju}(v)=\langle v,u\rangle\langle u,u\rangle u.$$

This formula projects vector vonto vector u, which is fundamental in geometry, regression, and approximation problems.



Orthogonal projection using inner product

Fill in the blank: Two vectors are orthogonal if their ______ is zero.

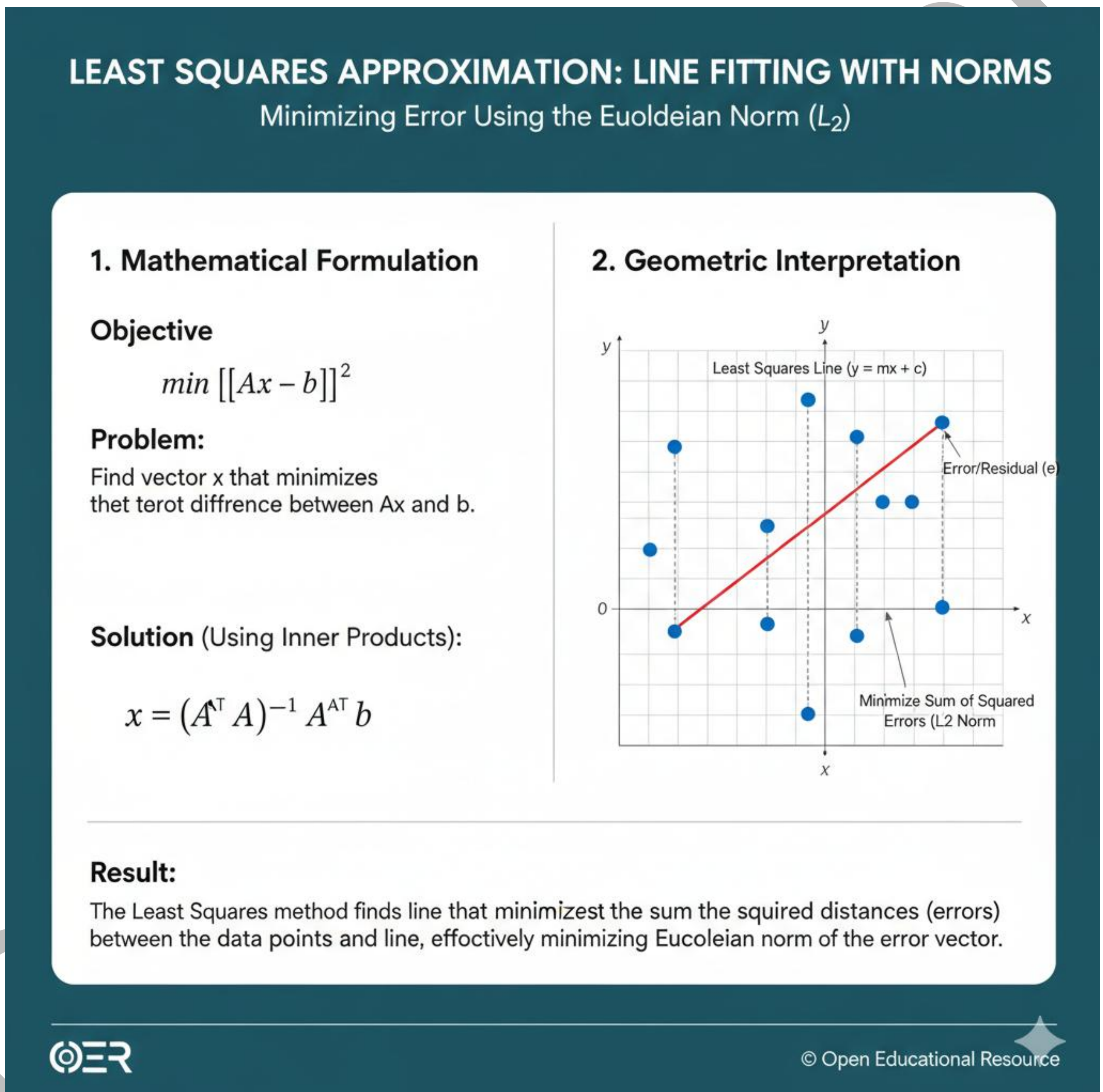
5.3.2 Least squares problems

Norms and inner products are essential in solving least squares problems, which minimise the error between observed data and a fitted model.

For a system $Ax \approx b$, the least squares solution minimises:

$$Ax-b2,$$

which is derived using inner products and projections. This method is widely used in statistics, data fitting, and machine learning.



Least squares approximation using norms

Fill in the blank: Least squares problems minimise the ______ between observed data and a fitted model.

5.3.3 Applications in data science and engineering

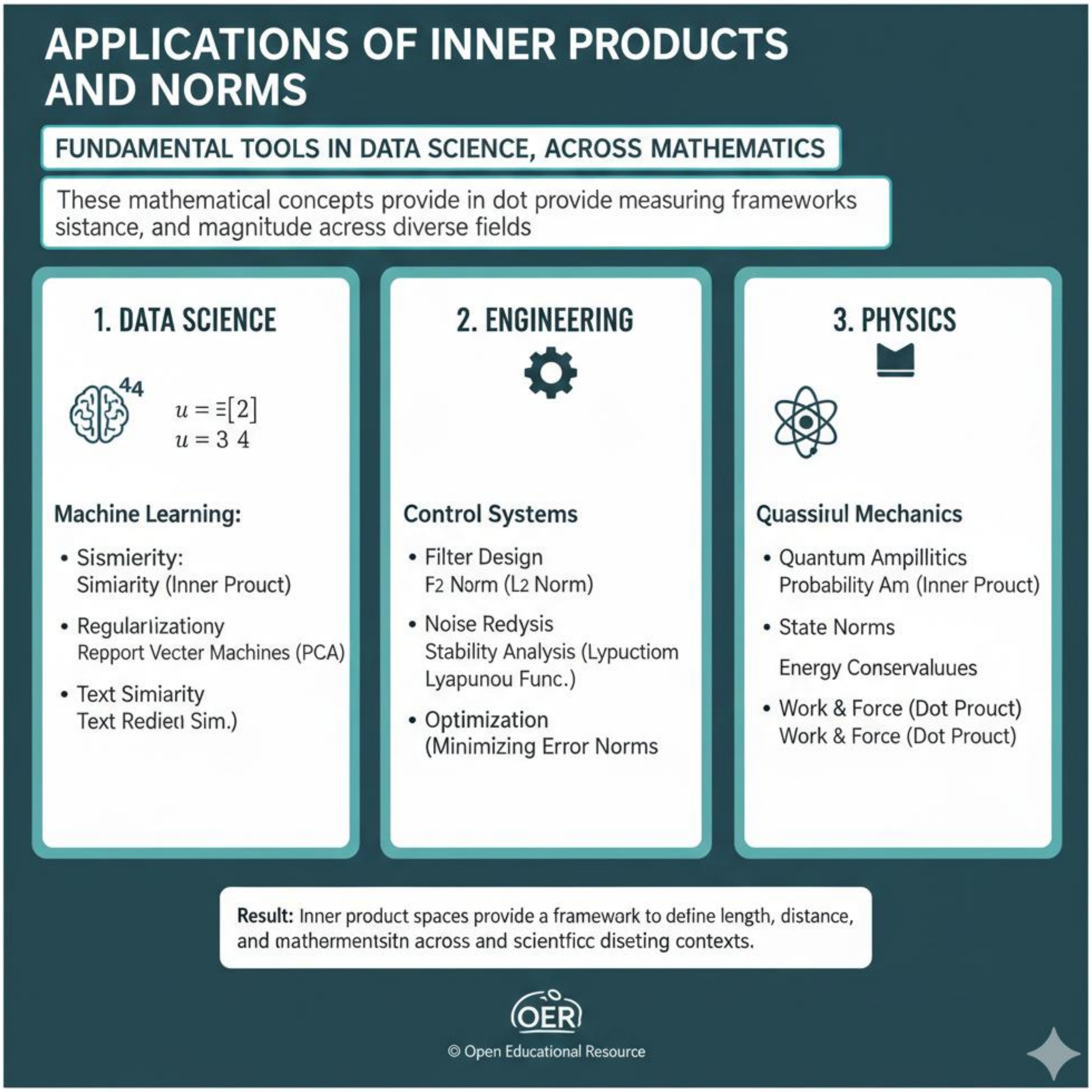
Inner products and norms have broad applications:

Data science: Measuring similarity between data vectors, PCA, regression.

Engineering: Signal processing, error minimisation, stability analysis.

Physics: Quantum mechanics uses inner products to calculate probabilities.

These applications highlight how abstract mathematical tools directly support practical problem-solving.



Applications of inner products and norms

Fill in the blank: In quantum mechanics, inner products are used to calculate ______.

Pilot Questions 5.3

Two vectors are orthogonal if: A. Their inner product is zero B. Their determinant is zero C. Their trace is zero D. Their sum is zero

The projection of vector vonto uis given by: A. ⟨v,u⟩⟨u,u⟩uB. ⟨v,u⟩uC. ⟨u,v⟩vD. uv

Least squares problems minimise: A. The determinant of a matrix B. The error between observed data and a fitted model C. The trace of a matrix D. The eigenvalues of a system

Inner products and norms are applied in: A. Data science B. Engineering C. Physics D. All of the above

In quantum mechanics, inner products are used to calculate: A. Probabilities B. Determinants C. Eigenvalues D. Trace

Series Summary

This Series introduced the concepts of inner product spaces and norms, which generalise the ideas of dot products and vector lengths to abstract spaces. We began by defining inner products, exploring their properties (linearity, symmetry, positivity), and examining examples in Euclidean, function, and complex spaces. We then studied norms, their relationship to inner products, and different types such as Euclidean, 1-norm, infinity norm, and p-norms. Finally, we explored applications of inner products and norms in orthogonality, projections, least squares problems, and practical fields like data science, engineering, and physics.

Glossary of Terms

Inner product: A generalisation of the dot product that measures angles and lengths in vector spaces.

Inner product space: A vector space equipped with an inner product.

Norm: A function that assigns a non-negative length or size to a vector.

Euclidean norm (2-norm): The standard length of a vector in Rn.

1-norm (Manhattan norm): The sum of absolute values of vector components.

Infinity norm (max norm): The maximum absolute value among vector components.

Projection: The component of one vector along the direction of another, defined using inner products.

Least squares problem: A method of minimising error between observed data and a fitted model using norms.

Concept Check Questions 5

Objective questions

The inner product of two vectors in Rnis: A. A vector B. A scalar C. A matrix D. A polynomial (Linked to SLO 5.1)

The Euclidean norm of vector v=(3,4)is: A. 5 B. 7 C. 12 D. 25 (Linked to SLO 5.5)

Short-answer questions

Define orthogonality of vectors using inner products. (Linked to SLO 5.6)

Explain how least squares problems use norms to minimise error. (Linked to SLO 5.7)

Long-answer questions

Discuss applications of inner products and norms in data science, engineering, and physics. (Linked to SLO 5.8)

Basic response: Describe one application in data science and one in engineering.

Value-added response: Provide detailed examples such as PCA in data science, error minimisation in engineering, and probability calculations in quantum mechanics.

Answers to Concept Check Questions 5

B. A scalar

A. 5

Two vectors are orthogonal if their inner product is zero. This means they are perpendicular in the geometric sense.

Least squares problems minimise the squared norm of the error vector Ax-b2, ensuring the best approximation of data by a model.

Basic response: In data science, inner products measure similarity between data vectors. In engineering, norms are used to minimise error in signal processing.

Value-added response: PCA uses inner products and norms to reduce dimensionality in data science. In engineering, least squares methods minimise error in system modelling. In physics, quantum mechanics uses inner products to calculate probabilities of states.

Answers to Pilot Questions 5

From Part 5.1

A scalar

Determinant

Positivity

Conjugation

An integral of the product of functions

From Part 5.2

A scalar length

5

⟨v,v⟩

5

7

From Part 5.3

Their inner product is zero

⟨v,u⟩⟨u,u⟩u

The error between observed data and a fitted model

All of the above

Probabilities