MATRICES

Introduction

Matrices - Introduction

Matrix algebra has at least two advantages:

•Reduces complicated systems of equations to simple expressions

•Adaptable to systematic method of mathematical treatment and well suited to computers

Definition:

A matrix is a set or group of numbers arranged in a square or rectangular array enclosed by two brackets

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \qquad \begin{bmatrix} 4 & 2 \\ -3 & 0 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Matrices - Introduction

Properties:

•A specified number of rows and a specified number of columns

•Two numbers (rows x columns) describe the dimensions or size of the matrix.

Examples:

3x3 matrix

2x4 matrix

1x2 matrix

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

Matrices - Introduction

A matrix is denoted by a bold capital letter and the elements within the matrix are denoted by lower case letters

e.g. matrix [**A**] with elements a_{ij}

$$\mathbf{A}_{mxn}= {}_{m}\mathbf{A}^{n} \quad \begin{bmatrix} a_{11} & a_{12}\ldots & a_{ij} & a_{in} \\ a_{21} & a_{22}\ldots & a_{ij} & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{ij} & a_{mn} \end{bmatrix}$$

i goes from 1 to m

j goes from 1 to n

Matrices - Introduction

TYPES OF MATRICES

1. Column matrix or vector:

The number of rows may be any integer but the number of columns is always 1

$$\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ -3 \end{bmatrix} \qquad \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

Matrices - Introduction

TYPES OF MATRICES

2. Row matrix or vector

Any number of columns but only one row

$$\begin{bmatrix} 1 & 1 & 6 \end{bmatrix} \qquad \begin{bmatrix} 0 & 3 & 5 & 2 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \cdots & a_{1n} \end{bmatrix}$$

Matrices - Introduction

TYPES OF MATRICES

3. Rectangular matrix

Contains more than one element and number of rows is not equal to the number of columns

$$\begin{bmatrix} 1 & 1 \\ 3 & 7 \\ 7 & -7 \\ 7 & 6 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 3 & 3 & 0 \end{bmatrix}$$

$$m \neq n$$

Matrices - Introduction

TYPES OF MATRICES

4. Square matrix

The number of rows is equal to the number of columns

(a square matrix $\mathbf{A}_{m \times m}$ has an order of m)

$$\begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 \\ 9 & 9 & 0 \\ 6 & 6 & 1 \end{bmatrix}$$

The principal or main diagonal of a square matrix is composed of all elements a_{ij} for which $i=j$

Matrices - Introduction

TYPES OF MATRICES

5. Diagonal matrix

A square matrix where all the elements are zero except those on the main diagonal

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

i.e. $a_{ij} = 0$ for all $i \neq j$

$a_{ij} \neq 0$ for some or all $i = j$

Matrices - Introduction

TYPES OF MATRICES

6. Unit or Identity matrix - I

A diagonal matrix with ones on the main diagonal

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} a_{ij} & 0 \\ 0 & a_{ij} \end{bmatrix}$$

i.e. $a_{ij} = 0$ for all $i \neq j$

$a_{ij} = 1$ for some or all $i = j$

Matrices - Introduction

TYPES OF MATRICES

7. Null (zero) matrix - 0

All elements in the matrix are zero

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

$$a_{ij} = 0 \qquad \text{For all } i,j$$

Matrices - Introduction

TYPES OF MATRICES

8. Triangular matrix

A square matrix whose elements above or below the main diagonal are all zero

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 8 & 9 \\ 0 & 1 & 6 \\ 0 & 0 & 3 \end{bmatrix}$$

Matrices - Introduction

TYPES OF MATRICES

8a. Upper triangular matrix

A square matrix whose elements below the main diagonal are all zero

$$\begin{bmatrix} a_{ij} & a_{ij} & a_{ij} \\ 0 & a_{ij} & a_{ij} \\ 0 & 0 & a_{ij} \end{bmatrix} \quad \begin{bmatrix} 1 & 8 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 7 & 4 & 4 \\ 0 & 1 & 7 & 4 \\ 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

i.e. $a_{ij} = 0$ for all $i > j$

Matrices - Introduction

TYPES OF MATRICES

8b. Lower triangular matrix

A square matrix whose elements above the main diagonal are all zero

$$\begin{bmatrix} a_{ij} & 0 & 0 \\ a_{ij} & a_{ij} & 0 \\ a_{ij} & a_{ij} & a_{ij} \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 2 & 3 \end{bmatrix}$$

i.e. $a_{ij} = 0$ for all $i < j$

Matrices – Introduction

TYPES OF MATRICES

9. Scalar matrix

A diagonal matrix whose main diagonal elements are equal to the same scalar

A scalar is defined as a single number or constant

$$\begin{bmatrix} a_{ij} & 0 & 0 \\ 0 & a_{ij} & 0 \\ 0 & 0 & a_{ij} \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

i.e. $a_{ij} = 0$ for all $i \neq j$
$a_{ij} = a$ for all $i = j$

MATRICES

Matrix Operations

Matrices - Operations

EQUALITY OF MATRICES

Two matrices are said to be equal only when all corresponding elements are equal

Therefore their size or dimensions are equal as well

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

Matrices - Operations

Some properties of equality:

•IIf **A** = **B**, then **B** = **A** for all **A** and **B**

•IIf **A** = **B**, and **B** = **C**, then **A** = **C** for all **A**, **B** and **C**

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 2 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

If **A** = **B** then $a_{ij} = b_{ij}$

Matrices - Operations

ADDITION AND SUBTRACTION OF MATRICES

The sum or difference of two matrices, **A** and **B** of the same size yields a matrix **C** of the same size

$$c_{ij} = a_{ij} + b_{ij}$$

Matrices of different sizes cannot be added or subtracted

Matrices - Operations

Commutative Law:

A + **B** = **B** + **A**

Associative Law:

A + (**B** + **C**) = (**A** + **B**) + **C** = **A** + **B** + **C**

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

A	**B**	**C**
2x3	2x3	2x3

Matrices - Operations

$\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$

$\mathbf{A} + (\mathbf{-A}) = \mathbf{0}$ (where $-\mathbf{A}$ is the matrix composed of $-a_{ij}$ as elements)

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

Matrices - Operations

SCALAR MULTIPLICATION OF MATRICES

Matrices can be multiplied by a scalar (constant or single element)

Let k be a scalar quantity; then

kA = Ak

Ex. If k=4 and

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

Matrices - Operations

$$4\times\begin{bmatrix}3 & -1\\2 & 1\\2 & -3\\4 & 1\end{bmatrix}=\begin{bmatrix}3 & -1\\2 & 1\\2 & -3\\4 & 1\end{bmatrix}\times 4=\begin{bmatrix}12 & -4\\8 & 4\\8 & -12\\16 & 4\end{bmatrix}$$

Properties:

- k (**A** + **B**) = k**A** + k**B**
- (k + g)**A** = k**A** + g**A**
- k(**AB**) = (k**A**)**B** = **A**(k)**B**
- k(g**A**) = (kg)**A**

Matrices - Operations

MULTIPLICATION OF MATRICES

The product of two matrices is another matrix

Two matrices **A** and **B** must be **conformable** for multiplication to be possible

i.e. the number of columns of **A** must equal the number of rows of **B**

Example.

$$\mathbf{A} \quad \text{x} \quad \mathbf{B} \quad = \quad \mathbf{C}$$

$$(1\text{x}3) \quad (3\text{x}1) \quad (1\text{x}1)$$

Matrices - Operations

$$\mathbf{B} \times \mathbf{A} = \text{Not possible!}$$

(2x1) (4x2)

$$\mathbf{A} \times \mathbf{B} = \text{Not possible!}$$

(6x2) (6x3)

Example

$$\mathbf{A} \times \mathbf{B} = \mathbf{C}$$

(2x3) (3x2) (2x2)

Matrices - Operations

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$(a_{11} \times b_{11}) + (a_{12} \times b_{21}) + (a_{13} \times b_{31}) = c_{11}$$

$$(a_{11} \times b_{12}) + (a_{12} \times b_{22}) + (a_{13} \times b_{32}) = c_{12}$$

$$(a_{21} \times b_{11}) + (a_{22} \times b_{21}) + (a_{23} \times b_{31}) = c_{21}$$

$$(a_{21} \times b_{12}) + (a_{22} \times b_{22}) + (a_{23} \times b_{32}) = c_{22}$$

Successive multiplication of row *i* of **A** with column *j* of **B** – row by column multiplication

Matrices - Operations

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

$$= \begin{bmatrix} 31 & 21 \\ 63 & 57 \end{bmatrix}$$

Remember also:

IA = **A**

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 31 & 21 \\ 63 & 57 \end{bmatrix} = \begin{bmatrix} 31 & 21 \\ 63 & 57 \end{bmatrix}$$

Matrices - Operations

Assuming that matrices **A**, **B** and **C** are conformable for the operations indicated, the following are true:

1. **AI** = **IA** = **A**
2. **A**(**BC**) = (**AB**)**C** = **ABC** - (associative law)
3. **A**(**B**+**C**) = **AB** + **AC** - (first distributive law)
4. (**A**+**B**)**C** = **AC** + **BC** - (second distributive law)

Caution!

1. **AB** not generally equal to **BA**, **BA** may not be conformable
2. If **AB** = **0**, neither **A** nor **B** necessarily = **0**
3. If **AB** = **AC**, **B** not necessarily = **C**

Matrices - Operations

AB not generally equal to **BA**, **BA** may not be conformable

$$T = \begin{bmatrix} 1 & 2 \\ 5 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}$$

$$TS = \begin{bmatrix} 1 & 2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 15 & 20 \end{bmatrix}$$

$$ST = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 23 & 6 \\ 10 & 0 \end{bmatrix}$$

Matrices - Operations

If **AB** = **0**, neither **A** nor **B** necessarily = **0**

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

Matrices - Operations

TRANSPOSE OF A MATRIX

If :

$$\underset{2x3}{A} = {}_2A^3 = \begin{bmatrix} 2 & 4 & 7 \\ 5 & 3 & 1 \end{bmatrix}$$

Then transpose of A, denoted A^T is:

$$A^T = {}_2A^{3^T} = \begin{bmatrix} 2 & 5 \\ 4 & 3 \\ 7 & 1 \end{bmatrix}$$

$$a_{ij} = a^T_{ji} \quad \text{For all } i \text{ and } j$$

Matrices - Operations

To transpose:

Interchange rows and columns

The dimensions of $\mathbf{A}^T$ are the reverse of the dimensions of $\mathbf{A}$

$$A = {}_2A^3 = \begin{bmatrix} 2 & 4 & 7 \\ 5 & 3 & 1 \end{bmatrix} \qquad 2 \text{ x } 3$$

$$A^T = {}_3A^{T\,2} = \begin{bmatrix} 2 & 5 \\ 4 & 3 \\ 7 & 1 \end{bmatrix} \qquad 3 \text{ x } 2$$

Matrices - Operations

Properties of transposed matrices:

1. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
2. $(\mathbf{AB})^T = \mathbf{B}^T\,\mathbf{A}^T$
3. $(k\mathbf{A})^T = k\mathbf{A}^T$
4. $(\mathbf{A}^T)^T = \mathbf{A}$

Matrices - Operations

1. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

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

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

Matrices - Operations

$(\mathbf{AB})^T = \mathbf{B}^T\,\mathbf{A}^T$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 8 \end{bmatrix}$$

Matrices - Operations

SYMMETRIC MATRICES

A Square matrix is symmetric if it is equal to its transpose:

$$\mathbf{A} = \mathbf{A}^{\mathrm{T}}$$

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

$$A^T = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

Matrices - Operations

When the original matrix is square, transposition does not affect the elements of the main diagonal

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

The identity matrix, **I**, a diagonal matrix **D**, and a scalar matrix, **K**, are equal to their transpose since the diagonal is unaffected.

Matrices - Operations

INVERSE OF A MATRIX

Consider a scalar k. The inverse is the reciprocal or division of 1 by the scalar.

Example:

k=7 the inverse of k or $k^{-1} = 1/k = 1/7$

Division of matrices is not defined since there may be **AB** = **AC** while $\mathbf{B} \neq \mathbf{C}$

Instead matrix inversion is used.

The inverse of a square matrix, **A**, if it exists, is the unique matrix $\mathbf{A}^{-1}$ where:

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

Matrices - Operations

Example:

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

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

Because:

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

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

Matrices - Operations

Properties of the inverse:

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

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

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

$$(kA)^{-1} = \frac{1}{k}A^{-1}$$

A square matrix that has an inverse is called a nonsingular matrix

A matrix that does not have an inverse is called a singular matrix

Square matrices have inverses except when the determinant is zero

When the determinant of a matrix is zero the matrix is singular

Matrices - Operations

DETERMINANT OF A MATRIX

To compute the inverse of a matrix, the determinant is required

Each square matrix **A** has a unit scalar value called the determinant of **A**, denoted by det **A** or |**A**|

If $$A = \begin{bmatrix} 1 & 2 \\ 6 & 5 \end{bmatrix}$$

then $$|A| = \begin{vmatrix} 1 & 2 \\ 6 & 5 \end{vmatrix}$$

Matrices - Operations

If **A** = [**A**] is a single element (1x1), then the determinant is defined as the value of the element

Then |**A**| =det **A** = a_{11}

If **A** is (n x n), its determinant may be defined in terms of order (n-1) or less.

Matrices - Operations

MINORS

If **A** is an n x n matrix and one row and one column are deleted, the resulting matrix is an (n-1) x (n-1) submatrix of **A**.

The determinant of such a submatrix is called a minor of **A** and is designated by m_{ij} , where *i* and *j* correspond to the deleted

row and column, respectively.

m_{ij} is the minor of the element a_{ij} in **A**.

Matrices - Operations

eg.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Each element in **A** has a minor

Delete first row and column from **A** .

The determinant of the remaining 2 x 2 submatrix is the minor of a_{11}

$$m_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

Matrices - Operations

Therefore the minor of a_{12} is:

$$m_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

And the minor for a_{13} is:

$$m_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Matrices - Operations

COFACTORS

The cofactor C_{ij} of an element a_{ij} is defined as:

$$C_{ij} = (-1)^{i+j} m_{ij}$$

When the sum of a row number *i* and column *j* is even, $c_{ij} = m_{ij}$ and when *i*+*j* is odd, $c_{ij} = -m_{ij}$

$$c_{11}(i = 1, j = 1) = (-1)^{1+1} m_{11} = +m_{11}$$

$$c_{12}(i = 1, j = 2) = (-1)^{1+2} m_{12} = -m_{12}$$

$$c_{13}(i = 1, j = 3) = (-1)^{1+3} m_{13} = +m_{13}$$

Matrices - Operations

DETERMINANTS CONTINUED

The determinant of an n x n matrix **A** can now be defined as

$$|A| = \det A = a_{11}c_{11} + a_{12}c_{12} + \ldots + a_{1n}c_{1n}$$

The determinant of **A** is therefore the sum of the products of the elements of the first row of **A** and their corresponding cofactors.

(It is possible to define |**A**| in terms of any other row or column but for simplicity, the first row only is used)

Matrices - Operations

Therefore the 2 x 2 matrix :

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

Has cofactors :

$$c_{11} = m_{11} = |a_{22}| = a_{22}$$

And:

$$c_{12} = -m_{12} = -|a_{21}| = -a_{21}$$

And the determinant of **A** is:

$$|A| = a_{11}c_{11} + a_{12}c_{12} = a_{11}a_{22} - a_{12}a_{21}$$

Matrices - Operations

Example 1:

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

$$|A| = (3)(2) - (1)(1) = 5$$

Matrices - Operations

For a 3 x 3 matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The cofactors of the first row are:

$$c_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}$$

$$c_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -(a_{21}a_{33} - a_{23}a_{31})$$

$$c_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}$$

Matrices - Operations

The determinant of a matrix A is:

$$|A| = a_{11}c_{11} + a_{12}c_{12} = a_{11}a_{22} - a_{12}a_{21}$$

Which by substituting for the cofactors in this case is:

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

Matrices - Operations

Example 2:

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

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$|A| = (1)(2-0) - (0)(0+3) + (1)(0+2) = 4$$

Matrices - Operations

ADJOINT MATRICES

A cofactor matrix **C** of a matrix **A** is the square matrix of the same order as **A** in which each element a_{ij} is replaced by its cofactor c_{ij} .

Example:

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

The cofactor C of A is $$C = \begin{bmatrix} 4 & 3 \\ -2 & 1 \end{bmatrix}$$

Matrices - Operations

The adjoint matrix of **A**, denoted by adj **A**, is the transpose of its cofactor matrix

$$adj A = C^T$$

It can be shown that:

$$\mathbf{A}(\text{adj } \mathbf{A}) = (\text{adj}\mathbf{A})\,\mathbf{A} = |\mathbf{A}|\,\mathbf{I}$$

Example:

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

$$|A| = (1)(4) - (2)(-3) = 10$$

$$adj A = C^T = \begin{bmatrix} 4 & -2 \\ 3 & 1 \end{bmatrix}$$

Matrices - Operations

$$A(adjA)=\begin{bmatrix}1 & 2\\ -3 & 4\end{bmatrix}\begin{bmatrix}4 & -2\\ 3 & 1\end{bmatrix}=\begin{bmatrix}10 & 0\\ 0 & 10\end{bmatrix}=10I$$

$$(adjA)A=\begin{bmatrix}4 & -2\\ 3 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\ -3 & 4\end{bmatrix}=\begin{bmatrix}10 & 0\\ 0 & 10\end{bmatrix}=10I$$

Matrices - Operations

USING THE ADJOINT MATRIX IN MATRIX INVERSION

Since

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$$

and

$$\mathbf{A}(\mathrm{adj}\ \mathbf{A}) = (\mathrm{adj}\mathbf{A})\ \mathbf{A} = |\mathbf{A}|\ \mathbf{I}$$

then

$$A^{-1} = \frac{adjA}{|A|}$$

Matrices - Operations

Example

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

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 4 & -2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.2 \\ 0.3 & 0.1 \end{bmatrix}$$

To check $\quad \mathbf{AA}^{-1} = \mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} 0.4 & -0.2 \\ 0.3 & 0.1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$A^{-1}A = \begin{bmatrix} 0.4 & -0.2 \\ 0.3 & 0.1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Matrices - Operations

Example 2

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

The determinant of **A** is

|**A**| = (3)(-1-0)-(-1)(-2-0)+(1)(4-1) = -2

The elements of the cofactor matrix are

$$c_{11} = +(-1), \qquad c_{12} = -(-2), \qquad c_{13} = +(3),$$
$$c_{21} = -(-1), \qquad c_{22} = +(-4), \qquad c_{23} = -(7),$$
$$c_{31} = +(-1), \qquad c_{32} = -(-2), \qquad c_{33} = +(5),$$

Matrices - Operations

The cofactor matrix is therefore

$$C = \begin{bmatrix} -1 & 2 & 3 \\ 1 & -4 & -7 \\ -1 & 2 & 5 \end{bmatrix}$$

so

$$adjA = C^T = \begin{bmatrix} -1 & 1 & -1 \\ 2 & -4 & 2 \\ 3 & -7 & 5 \end{bmatrix}$$

and

$$A^{-1} = \frac{adjA}{|A|} = \frac{1}{-2}\begin{bmatrix} -1 & 1 & -1 \\ 2 & -4 & 2 \\ 3 & -7 & 5 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ -1.0 & 2.0 & -1.0 \\ -1.5 & 3.5 & -2.5 \end{bmatrix}$$

Matrices - Operations

The result can be checked using

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

The determinant of a matrix must not be zero for the inverse to exist as there will not be a solution

Nonsingular matrices have non-zero determinants

Singular matrices have zero determinants

MATRIX INVERSION

Simple 2 x 2 case

Simple 2 x 2 case

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and

$$A^{-1} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

Since it is known that

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

then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Simple 2 x 2 case

Multiplying gives

$$aw + by = 1$$

$$ax + bz = 0$$

$$cw + dy = 0$$

$$cx + dz = 1$$

It can simply be shown that

$$|A| = ad - bc$$

Simple 2 x 2 case

thus

$$y = \frac{1 - aw}{b}$$

$$y = \frac{-cw}{d}$$

$$\frac{1 - aw}{b} = \frac{-cw}{d}$$

$$w = \frac{d}{da - bc} = \frac{d}{|A|}$$

Simple 2 x 2 case

$$z = \frac{-ax}{b}$$

$$z = \frac{1 - cx}{d}$$

$$\frac{-ax}{b} = \frac{1 - cx}{d}$$

$$x = \frac{b}{-da + bc} = -\frac{b}{|A|}$$

Simple 2 x 2 case

$$w = \frac{1 - by}{a}$$

$$w = \frac{-dy}{c}$$

$$\frac{1 - by}{a} = \frac{-dy}{c}$$

$$y = \frac{c}{-ad + cb} = -\frac{c}{|A|}$$

Simple 2 x 2 case

$$x = \frac{-bz}{a}$$

$$x = \frac{1-dz}{c}$$

$$\frac{-bz}{a} = \frac{1-dz}{c}$$

$$z = \frac{a}{ad-bc} = \frac{a}{|A|}$$

Simple 2 x 2 case

So that for a 2 x 2 matrix the inverse can be constructed in a simple fashion as

$$A^{-1} = \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} \frac{d}{|A|} & \frac{b}{|A|} \\ \frac{-c}{|A|} & \frac{a}{|A|} \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

- Exchange elements of main diagonal
- Change sign in elements off main diagonal
- Divide resulting matrix by the determinant

Simple 2 x 2 case

Example

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

$$A^{-1} = -\frac{1}{10}\begin{bmatrix} 1 & -3 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} -0.1 & 0.3 \\ 0.4 & -0.2 \end{bmatrix}$$

Check inverse

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

$$-\frac{1}{10}\begin{bmatrix} 1 & -3 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

MATRICES AND LINEAR EQUATIONS

Linear Equations

Linear Equations

Linear equations are common and important for survey problems

Matrices can be used to express these linear equations and aid in the computation of unknown values

Example

n equations in *n* unknowns, the a_{ij} are numerical coefficients, the b_i are constants and the x_j are unknowns

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots & \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$

Linear Equations

The equations may be expressed in the form

$$\mathbf{AX} = \mathbf{B}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n1} \cdots & a_{nn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ and } \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

n x n　　　n x 1　　　n x 1

Number of unknowns = number of equations = n

Linear Equations

If the determinant is nonzero, the equation can be solved to produce n numerical values for x that satisfy all the simultaneous equations

To solve, premultiply both sides of the equation by $\mathbf{A}^{-1}$ which exists because $|\mathbf{A}| \neq \mathbf{0}$

$$\mathbf{A}^{-1}\,\mathbf{AX} = \mathbf{A}^{-1}\,\mathbf{B}$$

Now since

$$\mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$$

We get

$$\mathbf{X} = \mathbf{A}^{-1}\,\mathbf{B}$$

So if the inverse of the coefficient matrix is found, the unknowns, **X** would be determined

Linear Equations

Example

$$3x_1 - x_2 + x_3 = 2$$
$$2x_1 + x_2 = 1$$
$$x_1 + 2x_2 - x_3 = 3$$

The equations can be expressed as

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

Linear Equations

When $\mathbf{A}^{-1}$ is computed the equation becomes

$$X = A^{-1}B = \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ -1.0 & 2.0 & -1.0 \\ -1.5 & 3.5 & -2.5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 7 \end{bmatrix}$$

Therefore

$$x_1 = 2,$$

$$x_2 = -3,$$

$$x_3 = -7$$

Linear Equations

The values for the unknowns should be checked by substitution back into the initial equations

$$x_1 = 2, \quad x_2 = -3, \quad x_3 = -7$$

$$3x_1 - x_2 + x_3 = 2$$
$$2x_1 + x_2 = 1$$
$$x_1 + 2x_2 - x_3 = 3$$

$$3\times(2)-(-3)+(-7)=2$$
$$2\times(2)+(-3)=1$$
$$(2)+2\times(-3)-(-7)=3$$

EIGENVALUES AND EIGENVECTORS

Eigen values and Eigen vectors

Definition

Let A be an $n \times n$ matrix. A <u>*scalar*</u> λ is called an **eigenvalue** of A if there exists a nonzero vector $\mathbf{x}$ in $\mathbf{R}^n$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The vector $\mathbf{x}$ is called an **eigenvector** corresponding to λ.

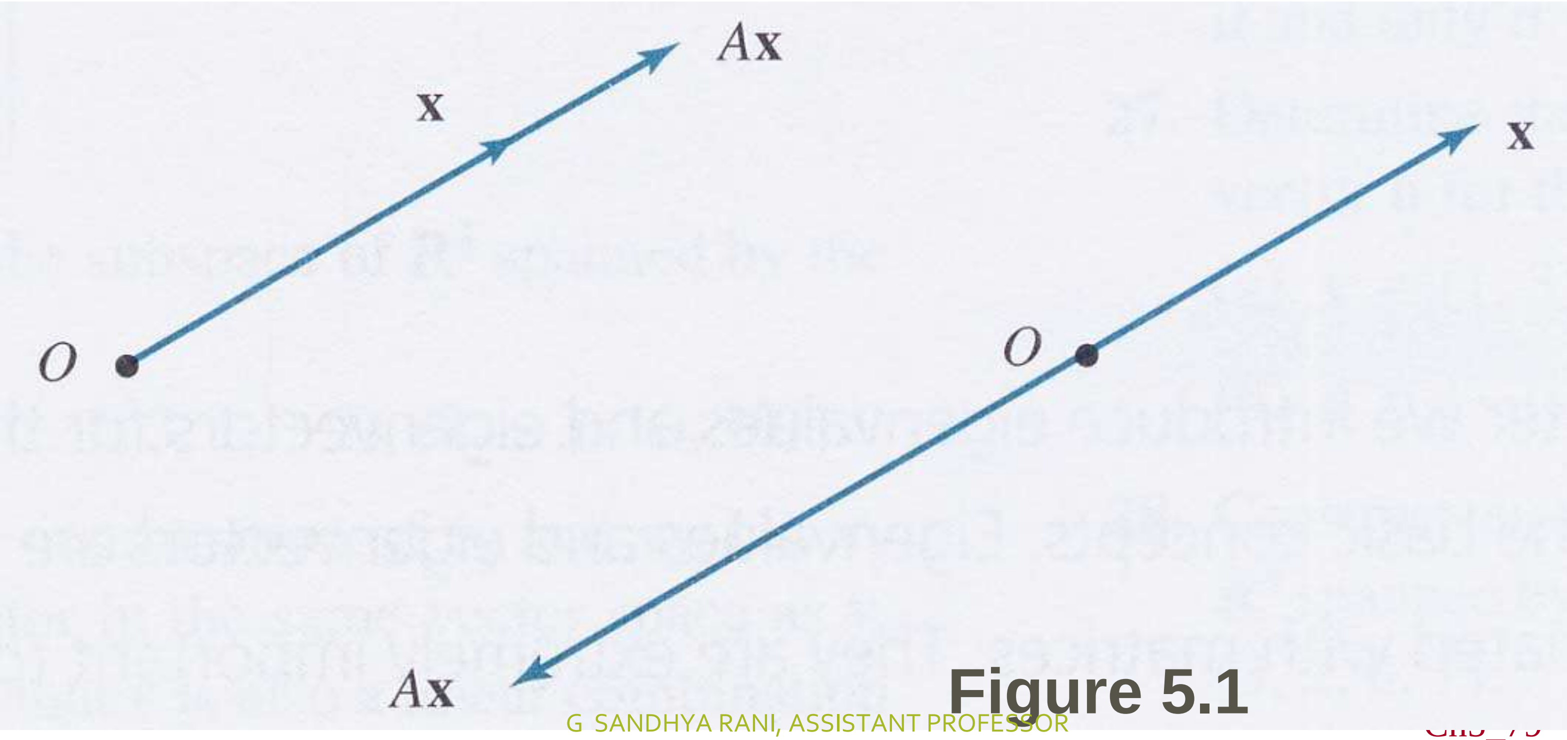


Figure 5.1

Computation of Eigen values and Eigen vectors

Let A be an $n \times n$ matrix with eigenvalue λ and corresponding eigenvector $\mathbf{x}$. Thus $A\mathbf{x} = \lambda\mathbf{x}$. This equation may be written

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$$

given

$$(A - \lambda I_n)\mathbf{x} = \mathbf{0}$$

Solving the equation $|A - \lambda I_n| = 0$ for λ leads to all the eigenvalues of A.

On expending the determinant $|A - \lambda I_n|$, we get a polynomial in λ.

This polynomial is called the **characteristic polynomial** of A.

The equation $|A - \lambda I_n| = 0$ is called the **characteristic equation** of A.

Find the eigenvalues and eigenvectors of the matrix

Example 1

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$$

Solution Let us first derive the characteristic polynomial of A.

We get

$$A - \lambda I_2 = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4-\lambda & -6 \\ 3 & 5-\lambda \end{bmatrix}$$

$$|A - \lambda I_2| = (-4-\lambda)(5-\lambda) + 18 = \lambda^2 - \lambda - 2$$

We now solve the characteristic equation of A.

$$\lambda^2 - \lambda - 2 = 0 \Rightarrow (\lambda - 2)(\lambda + 1) = 0 \Rightarrow \lambda = 2 \text{ or } -1$$

The eigenvalues of A are 2 and –1.

The corresponding eigenvectors are found by using these values of λ in the equation$(A - \lambda I_2)\mathbf{x} = \mathbf{0}$. There are many eigenvectors corresponding to each eigenvalue.

- For $\underline{\lambda = 2}$

  We solve the equation $(A - 2I_2)\mathbf{x} = \mathbf{0}$ for $\mathbf{x}$.

The matrix $(A - 2I_2)$ is obtained by subtracting 2 from the diagonal elements of A. We get

$$\begin{bmatrix} -6 & -6 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

This leads to the system of equations

$$-6x_1 - 6x_2 = 0$$
$$3x_1 + 3x_2 = 0$$

giving $x_1 = -x_2$. The solutions to this system of equations are $x_1 = -r$, $x_2 = r$, where r is a scalar. Thus the eigenvectors of A corresponding to $\lambda = 2$ are nonzero vectors of the form

$$\mathrm{v}_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = r\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- For $\lambda = -1$

  We solve the equation $(A + 1I_2)x = 0$ for x.

The matrix $(A + 1I_2)$ is obtained by adding 1 to the diagonal elements of A. We get

$$\begin{bmatrix} -3 & -6 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

This leads to the system of equations

$$-3x_1 - 6x_2 = 0$$
$$3x_1 + 6x_2 = 0$$

Thus $x_1 = -2x_2$. The solutions to this system of equations are $x_1 = -2s$ and $x_2 = s$, where s is a scalar. Thus the **eigenvectors** of A corresponding to $\lambda = -1$ are nonzero vectors of the form

$$\mathrm{v}_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Let A be an $n \times n$ matrix and λ an eigenvalue of A. The set of all eigenvectors corresponding to λ, together with the zero vector, is a subspace of $\boldsymbol{R}^n$. This subspace is called the **eigenspace** of λ.

Proof

Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two vectors in the eigenspace of λ and let c be a scalar. Then $A\mathbf{x}_1 = \lambda\mathbf{x}_1$ and $A\mathbf{x}_2 = \lambda\mathbf{x}_2$. Hence,

$$A\mathbf{x}_1 + A\mathbf{x}_2 = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2$$

$$A(\mathbf{x}_1 + \mathbf{x}_2) = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$$

Thus $\mathbf{x}_1 + \mathbf{x}_2$ is a vector in the eigenspace of λ. The set is closed under addition.

Further, since $A\mathbf{x}_1 = \lambda\mathbf{x}_1$,

$$cA\mathbf{x}_1 = c\lambda\mathbf{x}_1$$

$$A(c\mathbf{x}_1) = \lambda(c\mathbf{x}_1)$$

Therefore $c\mathbf{x}_1$ is a vector in the eigenspace of λ. The set is closed scalar multiplication.

Thus the set is a subspace of $\boldsymbol{R}^n$.

Example2: Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

Sol The matrix $A - \lambda I_3$ is obtained by subtracting λ from the diagonal elements of A. Thus

$$A - \lambda I_3 = \begin{bmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{bmatrix}$$

The characteristic polynomial of A is $|A - \lambda I_3|$. Using row and column operations to simplify determinants, we get

$$\left|A - \lambda I_3\right| = \begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & -1+\lambda & 0 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 1-\lambda & 0 & 0 \\ 4 & 9-\lambda & 2 \\ 2 & 4 & 2-\lambda \end{vmatrix}$$

$$= (1-\lambda)[(9-\lambda)(2-\lambda)-8] = (1-\lambda)[\lambda^2 - 11\lambda + 10]$$

$$= (1-\lambda)(\lambda-10)(\lambda-1) = -(\lambda-10)(\lambda-1)^2$$

We now solving the characteristic equation of *A*:

$$-(\lambda-10)(\lambda-1)^2 = 0$$

$$\lambda = 10 \text{ or } 1$$

The eigenvalues of *A* are 10 and 1.

The corresponding eigenvectors are found by using three values of λ in the equation $(A - \lambda I_3)\mathbf{x} = \mathbf{0}$.

- $\lambda_1 = 10$

  We get

$$(A - 10I_3)\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

The solution to this system of equations are $x_1 = 2r$, $x_2 = 2r$, and $x_3 = r$, where r is a scalar.
Thus the eigenspace of $\lambda_1 = 10$ is the one-dimensional space of vectors of the form.

$$r\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

- $\lambda_2 = 1$

  Let $\lambda = 1$ in $(A - \lambda I_3)\mathbf{x} = \mathbf{0}$. We get

$$(A - 1I_3)\mathbf{x} = \mathbf{0}$$

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

The solution to this system of equations can be shown to be $x_1 = -s - t$, $x_2 = s$, and $x_3 = 2t$, where s and t are scalars. Thus the eigenspace of $\lambda_2 = 1$ is the space of vectors of the form.

$$\begin{bmatrix} -s-t \\ s \\ 2t \end{bmatrix}$$

Separating the parameters s and t, we can write

$$\begin{bmatrix} -s-t \\ s \\ 2t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

Thus the eigenspace of $\lambda = 1$ is a two-dimensional subspace of $\mathbf{R}^3$ with basis

$$\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}\right\}$$

If an eigenvalue occurs as a k times repeated root of the characteristic equation, we say that it is of **multiplicity** k. Thus λ=10 has multiplicity 1, while λ=1 has multiplicity 2 in this example.

Practice problems

Ex: Prove that if A is a diagonal matrix, then its eigenvalues are the diagonal elements.

Ex: Prove that if A and A^t have the same eigenvalues.

Ex: Prove that the constant term of the characteristic polynomial of a matrix A is $|A|$.

Diagonalization of Matrices

Definition

Let A and B be square matrices of the same size. B is said to be **similar** to A if there exists an invertible matrix C such that $B = C^{-1}AC$. The transformation of the matrix A into the matrix B in this manner is called a **similarity transformation**.

Example : Consider the following matrices A and C with C is invertible. Use the similarity transformation $C^{-1}AC$ to transform A into a matrix B.

$$A = \begin{bmatrix} 7 & -10 \\ 3 & -4 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

Solution

$$B = C^{-1}AC = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 7 & -10 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 7 & -10 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

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

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

Theorem : Similar matrices have the same eigenvalues.

Proof

Let A and B be similar matrices. Hence there exists a matrix C such that $B = C^{-1}AC$.

The characteristic polynomial of B is $|B - \lambda I_n|$. Substituting for B and using the multiplicative properties of determinants, we get

$$
\begin{aligned}
|B-\lambda I| &= |C^{-1}AC-\lambda I| = |C^{-1}(A-\lambda I)C| \\
&= |C^{-1}||A-\lambda I||C| = |A-\lambda I||C^{-1}||C| \\
&= |A-\lambda I||C^{-1}C| = |A-\lambda I||I| \\
&= |A-\lambda I|
\end{aligned}
$$

The characteristic polynomials of A and B are identical. This means that their eigenvalues are the same.

Definition

A *square matrix* A is said to be **diagonalizable** if there exists a matrix C such that $D = C^{-1}AC$ is a diagonal matrix.

Theorem

Let A be an $n \times n$ matrix.

(a) If A has n linearly independent eigenvectors, it is diagonalizable. *The matrix C whose columns consist of n linearly independent eigenvectors can be used in a similarity transformation $C^{-1}AC$ to give a diagonal matrix D. The diagonal elements of D will be the eigenvalues of A.*

(b) If A is diagonalizable, then it has n linearly independent eigenvectors

Example :

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$$

(a) Show that the matrix is diagonalizable.

(b) Find a diagonal matrix D that is similar to A.

(c) Determine the similarity transformation that diagonalizes A.

Solution

(a) The eigenvalues and corresponding eigenvector of this matrix were found in Example 1 of Section 5.1. They are

$$\lambda_1 = 2, \mathbf{v}_1 = r\begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \lambda_2 = -1, \text{and } \mathbf{v}_2 = s\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Since A, a 2×2 matrix, has two linearly independent eigenvectors, it is diagonalizable.

(b) A is similar to the diagonal matrix D, which has diagonal elements $\lambda_1 = 2$ and $\lambda_2 = -1$. Thus

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} \text{ is similar to } D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

(c) Select two convenient linearly independent eigenvectors, say

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Let these vectors be the column vectors of the diagonalizing matrix C.

$$C = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \quad C^{-1}AC = \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$$

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

Note

If A is similar to a diagonal matrix D under the transformation $C^{-1}AC$, then it can be shown that $A^k = CD^kC^{-1}$.
This result can be used to compute A^k. Let us derive this result and then apply it.

$$D^k = (C^{-1}AC)^k = \underbrace{(C^{-1}AC) \quad \cdots \quad (C^{-1}AC)}_{k \text{ times}} = C^{-1}A^kC$$

This leads to

$$\boxed{A^k = CD^kC^{-1}}$$

Example : Compute A^9 for the following matrix A.

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$$

Solution

A is the matrix of the previous example. Use the values of C and D from that example. We get

$$D^9 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}^9 = \begin{bmatrix} 2^9 & 0 \\ 0 & (-1)^9 \end{bmatrix} = \begin{bmatrix} 512 & 0 \\ 0 & -1 \end{bmatrix}$$

$$A^9 = CD^9C^{-1}$$

$$= \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 512 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -514 & -1026 \\ 513 & 1025 \end{bmatrix}$$

Example : Show that the following matrix A is not diagonalizable.

$$A = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}$$

Solution

$$A - \lambda I_2 = \begin{bmatrix} 5-\lambda & -3 \\ 3 & -1-\lambda \end{bmatrix}$$

The characteristic equation is

$$|A - \lambda I_2| = 0 \Rightarrow (5-\lambda)(-1-\lambda) + 9 = 0$$

$$\Rightarrow \lambda^2 - 4\lambda + 4 = 0 \Rightarrow (\lambda - 2)(\lambda - 2) = 0$$

There is a single eigenvalue, $\lambda = 2$. We find he corresponding eigenvectors. $(A - 2I)\,\mathbf{x} = \mathbf{0}$ gives

$$\begin{bmatrix} 3 & -3 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \quad \Rightarrow \quad 3x_1 - 3x_2 = 0.$$

Thus $x_1 = r$, $x_2 = r$. The eigenvectors are nonzero vectors of the form

$$r\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenspace is a one-dimensional space. A is a 2×2 matrix, but it does not have two linearly independent eigenvectors. Thus A is not diagonalizable.

Theorem : Let A be an $n \times n$ symmetric matrix.

(a) All the eigenvalues of A are real numbers.

(b) The dimension of an eigenspace of A is the multiplicity of the eigenvalues as a root of the characteristic equation.

(c) A has n linearly independent eigenvectors.

Symmetric Matrices and Quadratic Forms

QUADRATIC FORMS

- **Example 1:** Let $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^T A\mathbf{x}$ for the following matrices.

a. $$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$$

b. $$A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

QUADRATIC FORMS

$$\mathrm{x}^T A\mathrm{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2$$

$$\mathrm{x}^T A\mathrm{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{bmatrix}$$

$$= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2)$$

$$= 3x_1^2 - 2x_1x_2 - 2x_2x_1 + 7x_2^2$$

$$= 3x_1^2 - 4x_1x_2 + 7x_2^2$$

QUADRATIC FORMS

- The presence of in the quadratic form in Example 1(b) is due to the entries off the diagonal in the matrix *A*.

- In contrast, the quadratic form associated with the diagonal matrix *A* in Example 1(a) has no x_1x_2 *cross-product* term.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- If **x** represents a variable vector in , then a **change of variable** is an equation of the form

$$\mathrm{x} = P\mathrm{y} \quad \text{, or equivalently,} \quad \mathrm{y} = P^{-1}\mathrm{x} \quad \text{----(1)}$$

  where *P* is an invertible matrix and **y** is a new variable vector in .

- Here **y** is the coordinate vector of **x** relative to the basis of determined by the columns of *P*.

- If the change of variable (1) is made in a quadratic form $\mathbf{x}^T A\mathbf{x}$, then

$$\mathrm{x}^T A\mathrm{x} = (P\mathrm{y})^T A(P\mathrm{y}) = \mathrm{y}^T P^T AP\mathrm{y} = \mathrm{y}^T (P^T AP)\mathrm{y} \quad \text{----(2)}$$

  and the new matrix of the quadratic form is P^TAP.

- Since A is symmetric, Theorem 2 guarantees that there is an orthogonal matrix P such that P^TAP is a diagonal matrix D, and the quadratic form in (2) becomes $\mathbf{y}^T D\mathbf{y}$.
- **Example 2:** Make a change of variable that transforms the quadratic form $Q(\mathrm{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.
- **Solution:** The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

- The first step is to orthogonally diagonalize *A*.
- Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$
- Associated unit eigenvectors are

$$\lambda = 3 : \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \lambda = -7 : \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

- These vectors are automatically orthogonal (because they correspond to distinct eigenvalues) and so provide an orthonormal basis for $\mathbb{R}^2$.

CHANGE OF VARIBALE IN A QUADRATIC FORM

- Let

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

- Then $A = PDP^{-1}$ and $D = P^{-1}AP = P^{T}AP$

- A suitable change of variable is

$$\mathrm{x} = P\mathrm{y}$$

where $\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$.

- Then

$$x_1^2 - 8x_1x_2 - 5x_2^2 = \mathrm{x}^T A\mathrm{x} = (P\mathrm{y})^T A(P\mathrm{y})$$
$$= \mathrm{y}^T P^T AP\mathrm{y} = \mathrm{y}^T D\mathrm{y}$$
$$= 3\mathrm{y}_1^2 - 7\mathrm{y}_2^2$$

- To illustrate the meaning of the equality of quadratic forms in Example 2, we can compute *Q* (**x**) for $\mathrm{x} = (2, -2)$ using the new quadratic form.

- First, since $\mathrm{x} = P\mathrm{y}$

$$\mathrm{y} = P^{-1}\mathrm{x} = P^{T}\mathrm{x}$$

so

$$y = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

- Hence

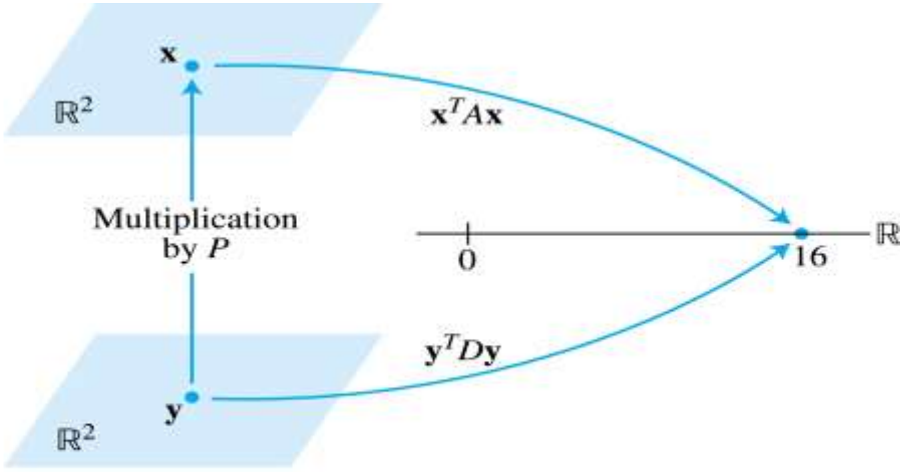
$$3\mathrm{y}_1^2 - 7\mathrm{y}_2^2 = 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5)$$

$$= 80/5 = 16$$

- This is the value of *Q* (**x**) when $\mathrm{x} = (2, -2)$

THE PRINCIPAL AXIS THEOREM

- See the figure below.



Change of variable in $\mathbf{x}^T A\mathbf{x}$.

- **Theorem 4:** Let *A* be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathrm{x} = P\mathrm{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

THE PRINCIPAL AXIS THEOREM

- The columns of *P* in theorem 4 are called the **principal axes** of the quadratic form $\mathbf{x}^T A\mathbf{x}$.

- The vector **y** is the coordinate vector of **x** relative to the orthonormal basis of given by these principal axes.

- **A Geometric View of Principal Axes**

- Suppose $Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x}$, where *A* is an invertible 2×2
- symmetric matrix, and let *c* be a constant.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- It can be shown that the set of all **x** in that satisfy

$$\mathrm{x}^T A \mathrm{x} = c$$

----(3)

either corresponds to an ellipse (or circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.

- If *A* is a diagonal matrix, the graph is in *standard position*, such as in the figure below.

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \quad a > b > 0$$

ellipse

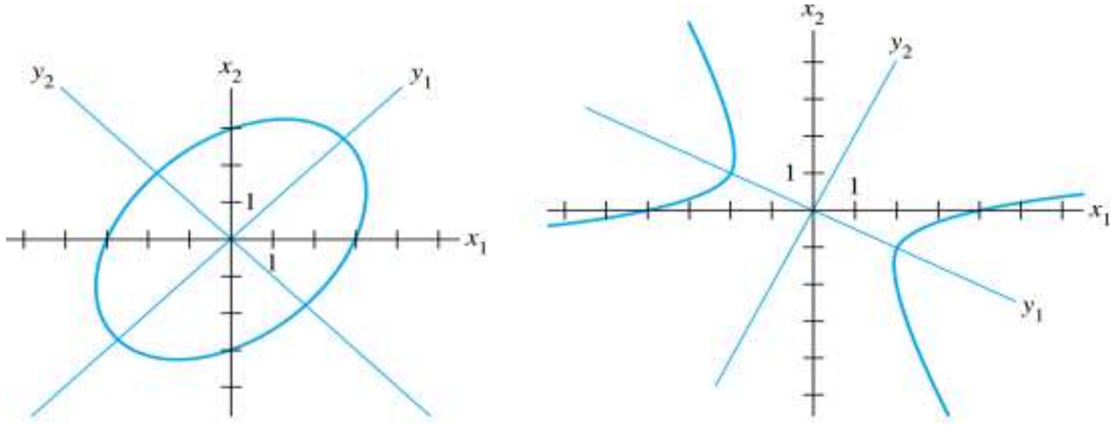
$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \quad a > b > 0$$

hyperbola

An ellipse and a hyperbola in standard position.

A GEOMETRIC VIEW OF PRINCIPAL AXES

- If *A* is not a diagonal matrix, the graph of equation (3) is rotated out of standard position, as in the figure below.



(a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$ (b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

An ellipse and a hyperbola *not* in standard position.

- Finding the *principal axes* (determined by the eigenvectors of *A*) amounts to finding a new coordinate system with respect to which the graph is in standard position.

CLASSIFYING QUADRATIC FORMS

- **Definition:** A quadratic form *Q* is:
  a. **positive definite** if $Q(\mathrm{x}) > 0$ for all $\mathrm{x} \neq 0$,
  b. **negative definite** if $Q(\mathrm{x}) < 0$ for all $\mathrm{x} \neq 0$,
  c. **indefinite** if *Q* (**x**) assumes both positive and negative values.

- Also, *Q* is said to be **positive semidefinite** if $Q(\mathrm{x}) \geq 0$ for all **x**, and

- **negative semidefinite** if $Q(\mathrm{x}) \leq 0$ for all **x**.

QUADRATIC FORMS AND EIGENVALUES

- **Theorem 5:** Let *A* be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}^T A \mathbf{x}$ is:
  a. positive definite if and only if the eigenvalues of *A* are all positive,
  b. negative definite if and only if the eigenvalues of *A* are all negative, or
  c. indefinite if and only if *A* has both positive and negative eigenvalues.

QUADRATIC FORMS AND EIGENVALUES

- **Proof:** By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathrm{x} = P\mathrm{y}$ such that

$$Q(\mathrm{x}) = \mathrm{x}^T A\mathrm{x} = \mathrm{y}^T D\mathrm{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

----(4)

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

- Since P is invertible, there is a one-to-one correspondence between all nonzero **x** and all nonzero **y**.

QUADRATIC FORMS AND EIGENVALUES

- Thus the values of $Q(\mathbf{x})$ for $\mathbf{x} \neq 0$ coincide with the values of the expression on the right side of (4), which is controlled by the signs of the eigenvalues $\lambda_1, \ldots, \lambda_n$, in three ways described in the theorem 5.

MEAN VALUE THEOREM

- **Mean Value Theorem (MVT)**
  - Lagrange’s MVT
  - Rolle’s Theorem
  - Cauchy’s MVT
- **Applications**

Motivation

- **Law of Mean:**

 For a “smooth” curve (a curve which can be drawn in a plane without lifting the pencil on a certain interval) y=f(x) ($a \le x \le b$) it looks evident that at some point c lies between a and b i.e. $a<c<b$, the slope of the tangent f’(c) will be equal to the slope of the chord joining the end points of the curve. That is, for some c lies between a and b ($a<c<b$),

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Physical Interpretation:

The velocity of a particle (matter) is exactly equal to the average speed.

Mean value Theorem (Lagrange's MVT)

- **Statement:**

Let f(x) be any real valued function defined in $a \le x \le b$ such that

(i) f(x) is continuous in $a \le x \le b$

(ii) f(x) is differentiable in $a<x<b$

Then, there exists at least one c in $a<c<b$ such that

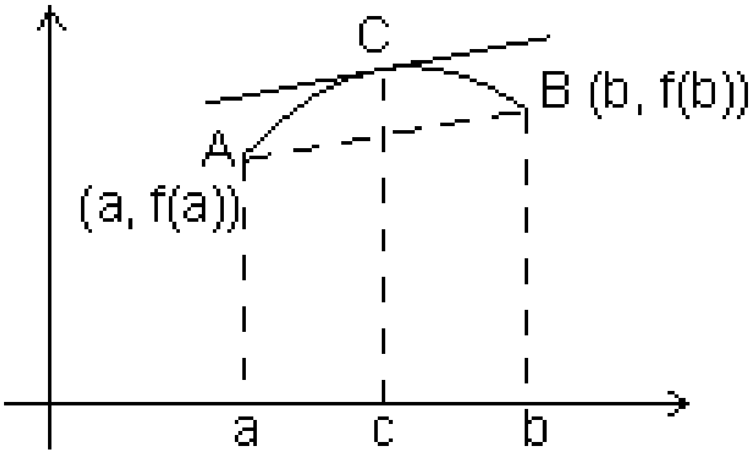
$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Alternative form (by taking $b=a+h$, h: small increment)

$f(a+h)=f(a)+hf'(a+\theta h)$ for some θ, $0< \theta<1$

Geometrical Interpretation of MVT

- For a continuous curve y=f(x) defined in a≤x≤b, the slope of the tangent f'(c) (where c lies between a and b i.e. a<c<b) to the curve is parallel to the slope of the chord joining the end points of the curve.



Special Case of MVT

- **Rolle's Theorem**

Statement:

Let f(x) be any real valued function defined in $a \le x \le b$ such that

(i) f(x) is continuous in $a \le x \le b$

(ii) f(x) is differentiable in $a < x < b$

(iii) $f(a)=f(b)$

Then, there exists at least one c in $a<c<b$ such that $f'(c)=0$.

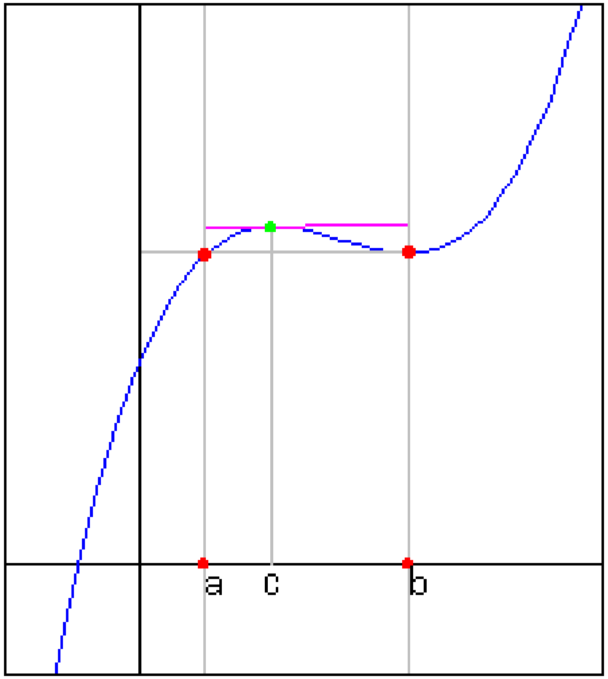
Note: All the conditions of Rolle's theorem are sufficient not necessary.

Counter Example: **i)** $f(x)=2+(x-1)^{2/3}$ **in** $0 \le x \le 2$

ii) $f(x)=|x-1|+|x-2|$ **on** $[-1,3]$

Geometrical Interpretation of Rolle's Theorem

- For a continuous curve $y=f(x)$ defined in $a \le x \le b$, the slope of the tangent $f'(c)$ (where c lies between a and b i.e. $a<c<b$) to the curve joining the two end points a and b is parallel to the x-axis.



General case (Cauchy's MVT)

- **Statement:**

Let f(x) and g(x) be two real valued function defined in a≤x≤b such that

(i) f(x) and g(x) are continuous in a≤x≤b

(ii) f(x) and g(x) are both differentiable in a<x<b

(iii) g'(x)≠o for some a<x<b

Then, there exists at least one c in a<c<b such that

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}$$

Alternative form:

$$\frac{f(a+h)-f(a)}{g(a+h)-g(a)}=\frac{f'(a+\theta h)}{g'(a+\theta h)},(0<\theta<1)$$

Interpretation (Cauchy's MVT)

- Useful generalization of the law of mean by considering a smooth curve in parametric representation x=g(t) and y=f(t) (a≤t≤b).

- The slope of the tangent to the curve at t=c is $\frac{f'(c)}{g'(c)}$

- The generalized law of mean asserts that there is always a value of c in a<c<b, for which the slope of the curve is equal to the slope of the tangent at c.

Observations

Rolle's Theorem

By taking f(x)=x in MVT

↓

MVT (Lagrange's)

By taking g(x)=x in Cauchy's MVT

↓

Cauchy's MVT

Applications

- To estimate some values of trigonometrical function say $\sin 46^{o}$ etc.
- Darboux's theorem: If the interval is an open subset of R and $f:I \longrightarrow R$ is differentiable at every point of I, then the range of an interval f' is an interval (not necessarily an open set).

  [This has the flavour of an "Intermediate Value Theorem" for f', but we are not assuming f' to be continuous].
- L' Hospital's Rule: If $f(x) \longrightarrow 0$, $g(x) \longrightarrow 0$ and $f'(x)/g'(x) \longrightarrow L$ as $x \longrightarrow c$, then $f(x)/g(x) \longrightarrow L$ as $x \longrightarrow c$.

- To deduce the necessary and sufficient condition of monotonic increasing or decreasing function.

For a continuous function $f:[a,b]\rightarrow R$ that is differentiable on (a,b), the following conditions are equivalent:

(i) f is increasing (or decreasing)

(ii) $f'(x)\geq 0$ (or $f'(x)\leq 0$)

➢ Not only the above examples but many more applications can found in different reference books from mathematics.

BETA AND GAMMA FUNCTIONS

Beta function

The first eulerian integral $\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ where m>0, n>0 is called a Beta function and is denoted by B(m,n).

The quantities m and n are positive but not necessarily integers.

Example:-

Properties of Beta Function

$$\mathrm{B}(x,y) = \mathrm{B}(y,x).$$

$$\mathrm{B}(x,y) = \int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}}\,dt, \qquad \mathrm{Re}(x) > 0,\ \mathrm{Re}(y) > 0$$

B(x,y) = B(x,y+1) + B(x+1,y)

xB(x,y +1) =y B(x+1,y)

$$\mathrm{B}(x,y) = 2\int_0^{\pi/2} (\sin\theta)^{2x-1}(\cos\theta)^{2y-1}\,d\theta, \qquad \mathrm{Re}(x) > 0,\ \mathrm{Re}(y) > 0$$

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$$

$$\mathrm{B}(x,y)\cdot\mathrm{B}(x+y,1-y) = \frac{\pi}{x\sin(\pi y)},$$

Gamma function

- The Eulerian integral $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ n>o
is called gamma function and is d ed by
- Example:-

$$\Gamma(x)$$

$$\Gamma(1) = \int_0^\infty t^{1-1} e^{-t} dt$$

$$= \lim_{n\to\infty} \int_0^n e^{-t} dt$$

$$= \lim_{n\to\infty} -e^{-t}\Big|_0^n$$

$$= \lim_{n\to\infty} \frac{-1}{e^n} - \frac{-1}{e^0} = \lim_{n\to\infty} 1 - \frac{1}{e^n} = 1$$

Recurre $\Gamma(x+1) = x\Gamma(x)$ r gamma function

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t} dt$$ Use integration by parts.

$$u = t^x \qquad dv = e^{-t} dt$$

$$du = xt^{x-1} dt \qquad v = -e^{-t}$$

$$\Gamma(x+1) = -t^x e^{-t}\Big|_0^\infty - \int_0^\infty \left(-e^{-t}\right) xt^{x-1} dt$$

$$\Gamma(x+1) = 0 + \int_0^\infty xt^{x-1} e^{-t} dt = x\int_0^\infty t^{x-1} e^{-t} dt$$

$$\therefore \Gamma(x+1) = x\Gamma(x)$$

Relation between gamma and factorial

$$\Gamma(n+1) = n!$$

Other results

$$\Gamma(1/2) = \int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$$

$$\Gamma(3/2) = \frac{1}{2}\Gamma(1/2) = \frac{\sqrt{\pi}}{2}$$

$$\Gamma(5/2) = \frac{3}{2}\Gamma(3/2) = \frac{3}{2}\frac{\sqrt{\pi}}{2} = \frac{3\sqrt{\pi}}{4}$$

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n}\sqrt{\pi} \qquad n = 1, 2, 3, \ldots$$

Relation between beta and gamma function

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx.$$

Setting $x = y + \frac{1}{2}$ gives the more symmetric formula

$$B(a,b) = \int_{-1/2}^{1/2} (\frac{1}{2} + y)^{a-1}(\frac{1}{2} - y)^{b-1}\,dy.$$

Now let $y = \frac{t}{2s}$ to obtain

$$(2s)^{a+b-1}B(a,b) = \int_{-s}^{s} (s+t)^{a-1}(s-t)^{b-1}\,dt.$$

Multiply by e^{-2s} then integrate with respect to s, $0 \le s \le A$, to get

$$B(a,b)\int_0^A e^{-2s}(2s)^{a+b-1}\,ds = \int_0^A \int_{-s}^{s} e^{-2s}(s+t)^{a-1}(s-t)^{b-1}\,dt\,ds.$$

Take the limit as $A \to \infty$ to get

$$\frac{1}{2}\mathrm{B}(a,b)\Gamma(a+b) = \lim_{A\to\infty} \int_0^A \int_{-s}^{s} e^{-2s}(s+t)^{a-1}(s-t)^{b-1}\,dt\,ds.$$

Let $\sigma = s+t,\ \tau = s-t$, so we integrate over

$$R = \{(\sigma,\tau) : \sigma+\tau \le 2A,\ \sigma,\tau \ge 0\}.$$

Since $s = \frac{1}{2}(\sigma+\tau),\ t = \frac{1}{2}(\sigma-\tau)$ the Jacobian determinant of the change of variables is

$$J = \begin{vmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2}$$

so

$$\frac{1}{2}\mathrm{B}(a,b)\Gamma(a+b) = \lim_{A\to\infty} \int\int_R \frac{1}{2} e^{-(\sigma+\tau)}\sigma^{a-1}\tau^{b-1}\,d\tau\,d\sigma.$$

Thus

$$
\begin{aligned}
B(a,b)\Gamma(a+b) &= \int_0^\infty\int_0^\infty e^{-(\sigma+\tau)}\sigma^{a-1}\tau^{b-1}\,d\tau\,d\sigma \\
&= \int_0^\infty\int_0^\infty e^{-\sigma}\sigma^{a-1}e^{-\tau}\tau^{b-1}\,d\tau\,d\sigma \\
&= \left(\int_0^\infty e^{-\sigma}\sigma^{a-1}\,d\sigma\right)\left(\int_0^\infty e^{-\tau}\tau^{b-1}\,d\tau\right).
\end{aligned}
$$

Thus, we have… $$B(x,y)=\frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$$

GAMMA FUNCTIONS

The factorial function

$$\int_0^\infty e^{-ax}dx = -\frac{1}{\alpha}e^{-\alpha x}\Big|_0^\infty = \frac{1}{\alpha} \quad (\alpha > 0)$$

$$\int_0^\infty xe^{-ax}dx = -\frac{1}{\alpha}xe^{-\alpha x}\Big|_0^\infty - \int_0^\infty\left(-\frac{1}{\alpha}e^{-\alpha x}\right)dx = \frac{1}{\alpha^2}.$$

Similarly, $\int_0^\infty x^2e^{-\alpha x}dx = \frac{2}{\alpha^3}$, $\int_0^\infty x^3e^{-\alpha x}dx = \frac{2\cdot 3}{\alpha^4}$

$$\int_0^\infty x^n e^{-\alpha x}dx = \frac{n!}{\alpha^{n+1}} \rightarrow \int_0^\infty x^n e^{-x}dx = n! \ (\alpha = 1)$$

Definition of the gamma function: recursion relation

– Gamma function $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx, \quad p > 0.$

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx = (n-1)!,$$

$$\Gamma(n+1) = \int_0^\infty x^{n} e^{-x} dx = n!.$$

– Recursion relation $\Gamma(p+1) = \int_0^\infty x^{p} e^{-x} dx = p!, \quad p > -1.$

$$\Gamma(p+1) = p\Gamma(p)$$

– Example $\Gamma(9/4) = (5/4)\Gamma(5/4) = (5/4)(1/4)\Gamma(1/4)$

$$so, \ \Gamma(1/4) \div \Gamma(9/4) = 16/5.$$

The Gamma function of negative numbers

$$\Gamma(p) = \frac{1}{p}\Gamma(p+1) \quad (p<0)$$

- Example

$$\Gamma(-0.3) = \frac{1}{-0.3}\Gamma(0.7), \quad \Gamma(-1.3) = \frac{1}{(-0.3)(-1.3)}\Gamma(0.7).$$

$$cf. \quad \Gamma(p) = \frac{1}{p}\Gamma(p+1) \to \infty \quad \text{as} \quad p \to 0.$$

- Using the above relation,

  1) Gamma(p= negative integers) $\to$ infinite.

  2) For p < 0, the sign changes alternatively in the intervals between negative integers

Some important formulas involving gamma functions

– $\Gamma(1/2) = \sqrt{\pi}$

(prove) $\Gamma(1/2) = \int_0^\infty \frac{1}{\sqrt{t}} e^{-t} dt = \int_0^\infty \frac{1}{y} e^{-y^2} 2y dy = 2\int_0^\infty e^{-y^2} dy.$

$$[\Gamma(1/2)]^2 = 4\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dxdy = 4\int_0^{\pi/2} \int_0^\infty e^{-r^2} r dr d\theta = \pi.$$

– $\Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin \pi p}.$

Beta functions

$$B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx, \quad p>0,\ q>0. \quad cf.\ B(p,q)=B(q,p)$$

$$i)\quad B(p,q)=\int_0^a\left(\frac{y}{a}\right)^{p-1}\left(1-\frac{y}{a}\right)^{q-1}\frac{dy}{a}=\frac{1}{a^{p+q-1}}\int_0^a y^{p-1}(a-y)^{q-1}dy.\ (x=y/a)$$

$$ii)\quad B(p,q)=2\int_0^{\pi/2}(\sin\theta)^{2p-1}(\cos\theta)^{2q-1}d\theta. \quad \left(x=\sin^2\theta\right)$$

$$iii)\quad B(p,q)=\int_0^\infty\frac{y^{p-1}dy}{(1+y)^{p+q}}. \quad (x=y/(1+y))$$

Beta functions in terms of gamma functions

$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Prove)

$$\Gamma(p)=\int_0^\infty t^{p-1}e^{-t}dt=2\int_0^\infty y^{2p-1}e^{-y^2}dy,\quad \Gamma(q)=2\int_0^\infty x^{2q-1}e^{-x^2}dx$$

$$\Gamma(p)\Gamma(q)=4\int_0^\infty\int_0^\infty x^{2q-1}y^{2p-1}e^{-(x^2+y^2)}dxdy$$

$$=4\int_0^\infty\int_0^{\pi/2}(r\cos\theta)^{2q-1}(r\sin\theta)^{2p-1}e^{-r^2}rdrd\theta$$

$$=4\int_0^\infty r^{2p+2q-1}e^{-r^2}dr\int_0^{\pi/2}(\cos\theta)^{2q-1}(\sin\theta)^{2p-1}d\theta=\frac{1}{2}\Gamma(p+q)\cdot\frac{1}{2}B(p,q)$$

– Example $I = \int_0^\infty \frac{x^3 dx}{(1+x)^5}$ $cf.$ $B(p,q) = \int_0^\infty \frac{y^{p-1} dy}{(1+y)^{p+q}}.$

$$p + q = 5, \quad p - 1 = 3 \ \rightarrow \ p = 4, \ q = 1.$$

$$\frac{\Gamma(4)\Gamma(1)}{\Gamma(5)} = \frac{3!}{4!} = \frac{1}{4}.$$

LEGENDRE'S DUPLICATION FORMULA

$$\boxed{\Gamma(1+z)\,\Gamma\left(z+\frac{1}{2}\right)=2^{-2z}\sqrt{\pi}\,\Gamma(2z+1)}$$

General proof in §13.3.

Proof for $z=n=1,2,3,\ldots$:

(Case $z=0$ is proved by inspection.)

$$\Gamma\left(n+\frac{1}{2}\right)=\left[\left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\cdots\frac{3}{2}\,\frac{1}{2}\right]\sqrt{\pi}$$

$$\Gamma(z)=(z-1)\,\Gamma(z-1)$$

$$=\frac{1}{2^n}\left[(2n-1)(2n-3)\cdots\right]\sqrt{\pi}\quad=\frac{1}{2^n}(2n-1)!!\,\sqrt{\pi}$$

$$\Gamma(1+n)=n!\;=\frac{(2n)!!}{2^n}\qquad\rightarrow\qquad\Gamma(1+n)\,\Gamma\left(n+\frac{1}{2}\right)=\frac{(2n)!!}{2^{2n}}(2n-1)!!\,\sqrt{\pi}$$

$$=\frac{\Gamma(2n+1)}{2^{2n}}\sqrt{\pi}$$

13.3. THE BETA FUNCTION

Beta Function :
$$\boxed{B(p,q) \equiv \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}} = B(q,p)$$

$$\Gamma(p)\,\Gamma(q) = 4\int_0^{\infty} ds\, e^{-s^2}\, s^{2p-1} \int_0^{\infty} dt\, e^{-t^2}\, t^{2q-1}$$

$$\boxed{\Gamma(z) = 2\int_0^{\infty} ds\, e^{-s^2}\, s^{2z-1}}$$

$$\begin{aligned} s &= r\cos\theta \\ t &= r\sin\theta \end{aligned} \quad \to \quad ds\,dt = \begin{vmatrix} \partial_r s & \partial_\theta s \\ \partial_r t & \partial_\theta t \end{vmatrix} dr\,d\theta = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} dr\,d\theta = r\,dr\,d\theta$$

$$\Gamma(p)\,\Gamma(q) = 4\int_0^{\infty} r\,dr \int_0^{\pi/2} d\theta\, e^{-r^2}\, r^{2(p+q-1)} \cos^{2p-1}\theta \sin^{2q-1}\theta$$

$$= 2\,\Gamma(p+q) \int_0^{\pi/2} d\theta\, \cos^{2p-1}\theta \sin^{2q-1}\theta$$

$$p = m+1$$
$$q = n+1$$

$$\to \quad B(p,q) = 2\int_0^{\pi/2} d\theta\, \cos^{2p-1}\theta \sin^{2q-1}\theta$$

$$\boxed{\frac{m!\,n!}{(m+n+1)!} = 2\int_0^{\pi/2} d\theta\, \cos^{2m+1}\theta \sin^{2n+1}\theta}$$

ALTERNATE FORMS : DEFINITE INTEGRALS

$$B(p+1,q+1) = 2\int_0^{\pi/2} d\theta\ \cos^{2p+1}\theta\ \sin^{2q+1}\theta$$

$t = \cos^2\theta \quad \rightarrow \quad B(p+1,q+1) = \int_0^1 d\cos^2\theta\ \cos^{2p}\theta \sin^{2q}\theta \ = \int_0^1 dt\ t^p\,(1-t)^q$

$t = x^2 \quad \rightarrow \quad B(p+1,q+1) = 2\int_0^1 dx\, x^{2p+1}\left(1-x^2\right)^q$

$t = \dfrac{u}{1+u} \quad \rightarrow \quad dt = \left(\dfrac{1}{1+u} - \dfrac{u}{(1+u)^2}\right) du \ = \dfrac{1}{(1+u)^2}\, du \qquad 1-t = \dfrac{1}{1+u}$

$$B(p+1,q+1) = \int_0^{1/2} du\, \frac{1}{(1+u)^2}\left(\frac{u}{1+u}\right)^p \left(\frac{1}{1+u}\right)^q \ = \int_0^{1/2} du\, \frac{u^p}{(1+u)^{p+q2}}$$

To be used in integral rep. of Bessel (Ex.14.1.17)
& hypergeometric (Ex.18.5.12) functions

DERIVATION: LEGENDRE DUPLICATION FORMULA

$$B(p+1,q+1)=\int_0^1 dt\ t^p (1-t)^q \ =2\int_0^1 dx\, x^{2p+1}(1-x^2)^q$$

$$t=\frac{1}{2}(1+s)$$

$$\rightarrow \quad B\left(z+\frac{1}{2},z+\frac{1}{2}\right)=\int_0^1 dt\ t^{z-1/2}(1-t)^{z-1/2} \ =\left(\frac{1}{2}\right)^{2z}\int_{-1}^1 ds\ (1+s)^{z-1/2}(1-s)^{z-1/2}$$

$$=\left(\frac{1}{2}\right)^{2z}\int_{-1}^1 ds\ (1-s^2)^{z-1/2} \ =\left(\frac{1}{2}\right)^{2z} 2\int_0^1 ds\ (1-s^2)^{z-1/2} \ =\left(\frac{1}{2}\right)^{2z} B\left(\frac{1}{2},z+\frac{1}{2}\right)$$

$$\rightarrow \quad \frac{\Gamma\left(z+\frac{1}{2}\right)\Gamma\left(z+\frac{1}{2}\right)}{\Gamma(2z+1)}=\left(\frac{1}{2}\right)^{2z}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(z+\frac{1}{2}\right)}{\Gamma(z+1)}$$

$$\Gamma(z+1)\Gamma\left(z+\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2^{2z}}\Gamma(2z+1)$$

Double Integrals

The Double Integral over a Rectangle

We start with a function f continuous on a rectangle

$$R : a \le x \le b,\ c \le y \le d$$

We want to define the double integral of f over R:

$$\iint_R f(x, y)\, dx\, dy.$$

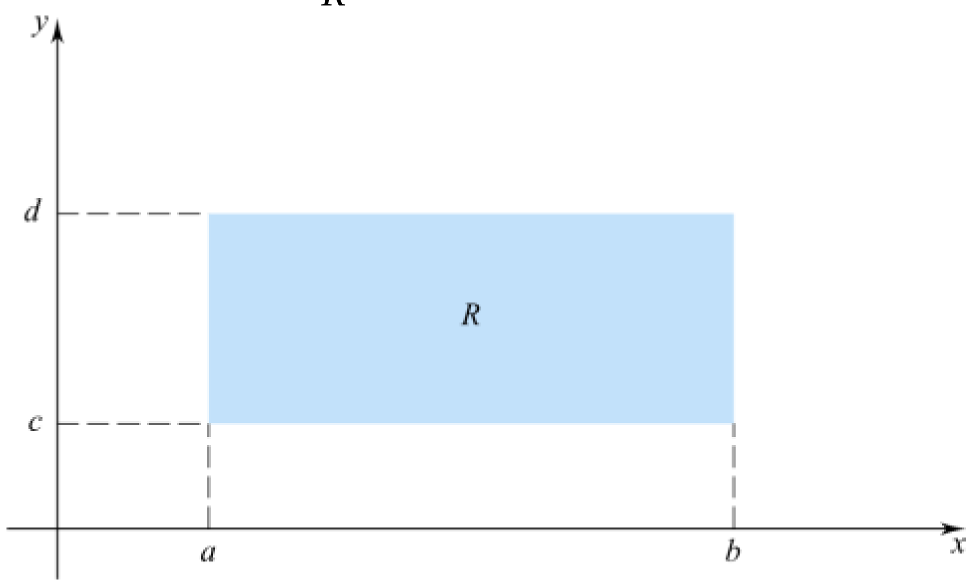


Figure 17.2.1

Let $P_1 = \{x_0, x_1, \ldots, x_m\}$ be a partition of $[a, b]$
and $P_2 = \{y_0, y_1, \ldots, y_n\}$ a partition of $[c, d]$.
Then the set

$$P = P_1 \times P_2 = \{(x_i, y_j) : \; x_i \in P_1, \; y_j \in P_2\}$$

is called a *partition of R*.

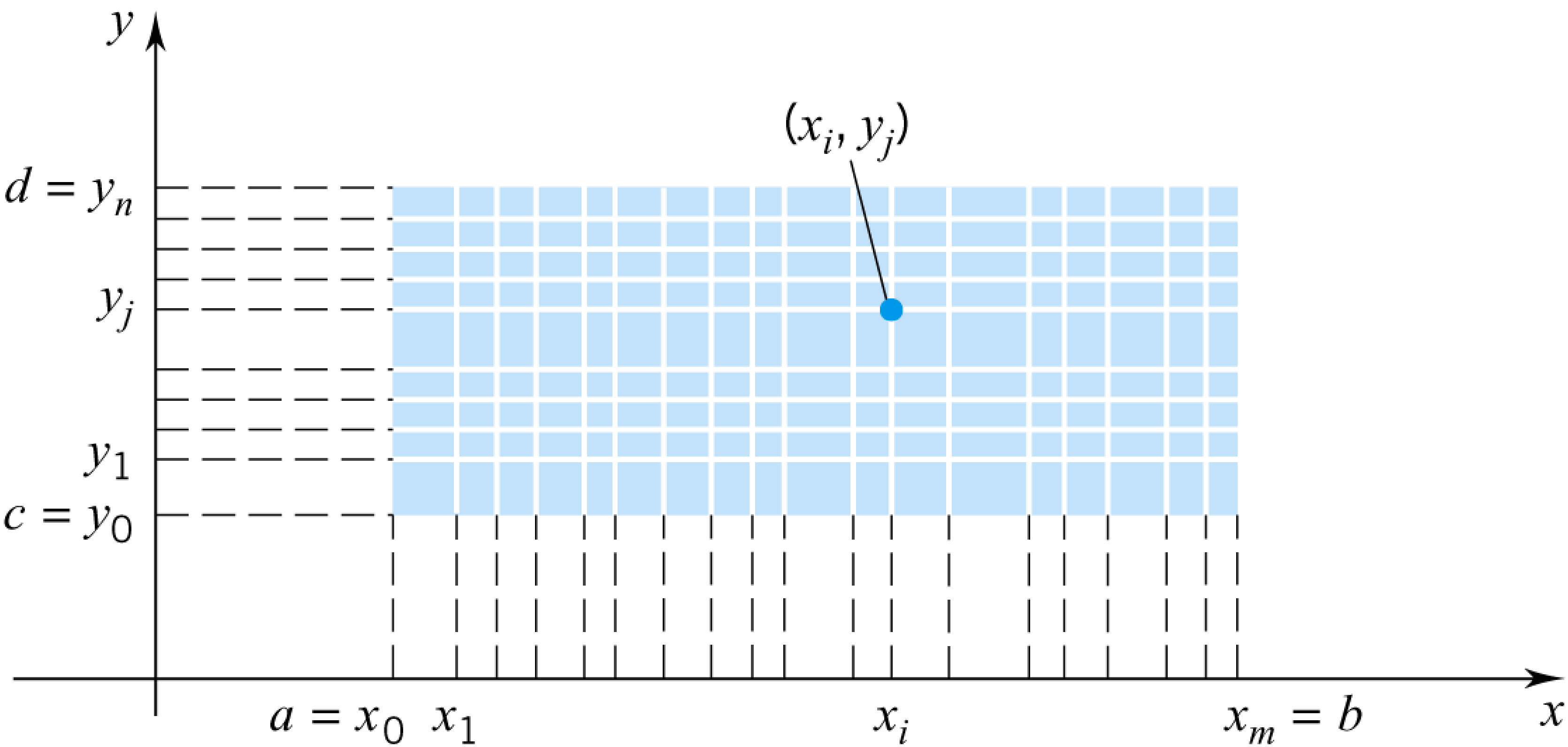


Figure 17.2.2

Double Integrals

The sum of all the products

$$M_{i\,j}\left(\text{area of } R_{i\,j}\right) = M_{i\,j}\left(x_i - x_{i-1}\right)\left(y_j - y_{j-1}\right) = M_{i\,j}\Delta x_i\, \Delta y_j$$

is called the *P upper sum* for f:

(17.2.1)

$$\boxed{U_f(P) = \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}(\text{area of } R_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}\,\Delta x_i\,\Delta y_j.}$$

The sum of all the products

$$m_{ij}\left(\text{area of } R_{ij}\right) = m_{ij}\left(x_i - x_{i-1}\right)\left(y_j - y_{j-1}\right) = m_{ij}\Delta x_i \, \Delta y_j$$

is called the *P lower sum* for *f* :

(17.2.2)

$$\boxed{L_f(P) = \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}(\text{area of } R_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}\,\Delta x_i\,\Delta y_j.}$$

Double Integrals

DEFINITION 17.2.3 THE DOUBLE INTEGRAL OVER A RECTANGLE R

Let f be continuous on a closed rectangle R. The unique number I that satisfies the inequality

$$L_f(P) \le I \le U_f(P) \qquad \text{for all partitions } P \text{ of } R$$

is called the *double integral* of f over R, and is denoted by

$$\iint_R f(x, y)\, dx\, dy.^{\dagger}$$

Double Integrals

The Double Integral as a Volume

If f is continuous and nonnegative on the rectangle R, the equation

$$z = f(x, y)$$

represents a surface that lies above R. In this case the double integral

$$\iint_R f(x)\,dx\,dy$$

gives the volume of the solid that is bounded below by R and bounded above by the surface $z = f(x, y)$

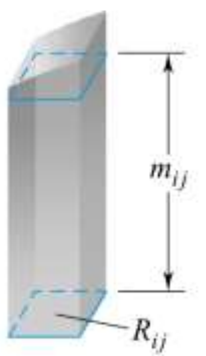


Figure 17.2.6

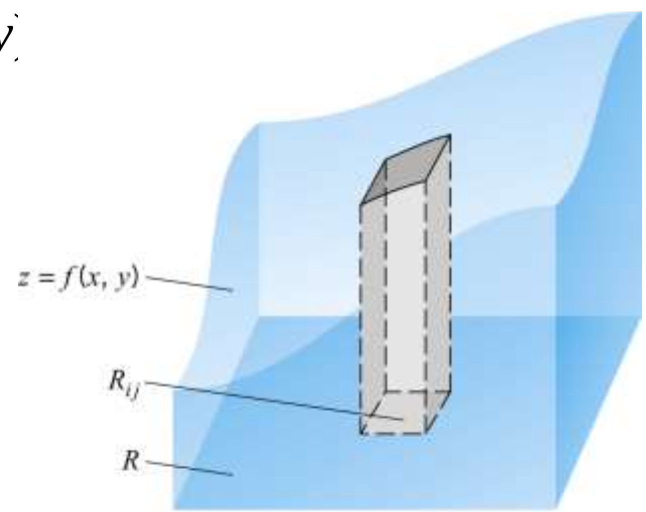


Figure 17.2.5

Double Integrals

Since the choice of a partition P is arbitrary, the volume of T must be the double integral:

(17.2.4)

$$\boxed{\text{volume of } T = \iint_R f(x, y)\, dx\, dy.}$$

The double integral

$$\iint_R 1\, dx\, dy = \iint_R dx\, dy$$

gives the volume of a solid of constant height 1 erected over R. In square units this is just the area of R:

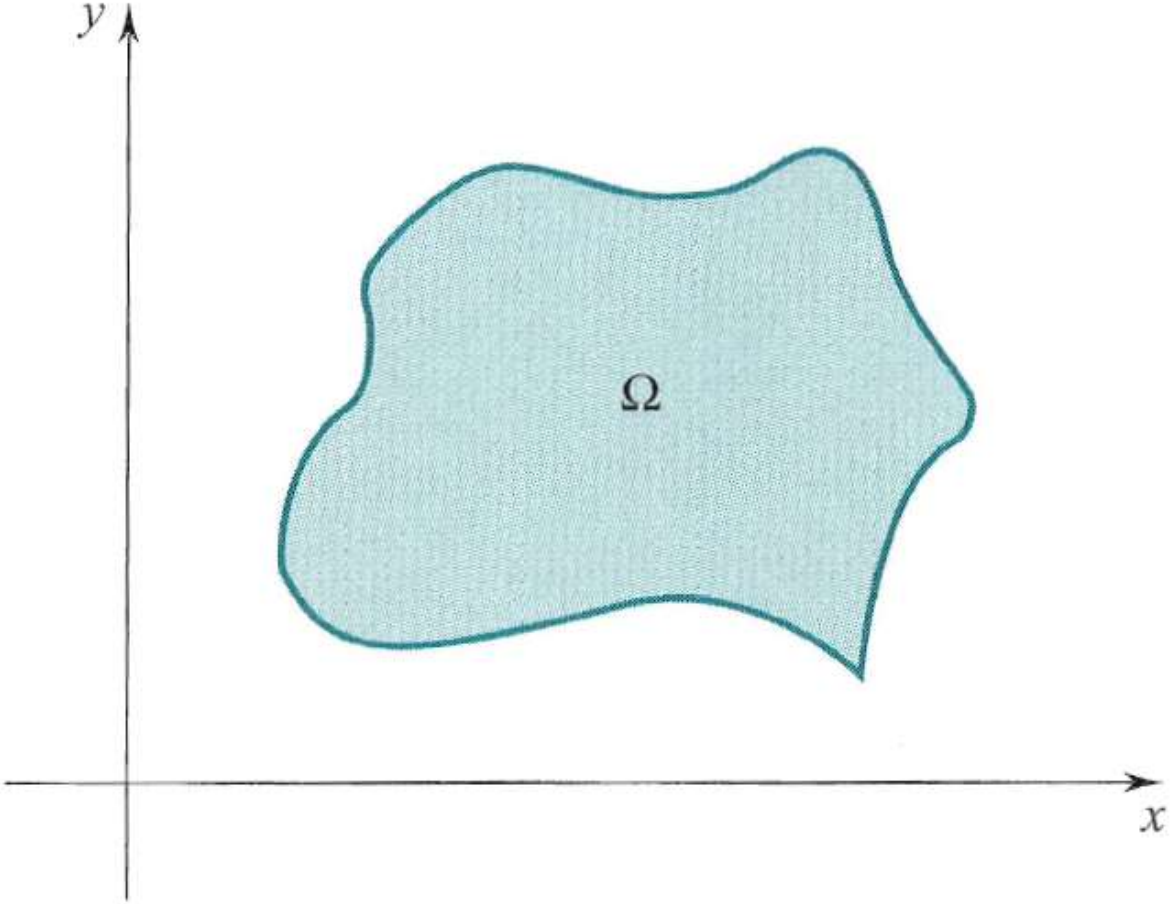
(17.2.5)

$$\boxed{\text{area of } R = \iint_R dx\, dy.}$$

Example 2. Evaluate

$$\iint_R \alpha\, dx\, dy$$

where α is a constant and R is the rectangle $R: \ a \leq x \leq b, c \leq y \leq d$.


y
Ω
x

Double Integrals

The Double Integral over a Region

(17.2.6)

$$\iint_{\Omega} f(x, y)\, dx\, dy = \iint_{R} f(x, y)\, dx\, dy.$$

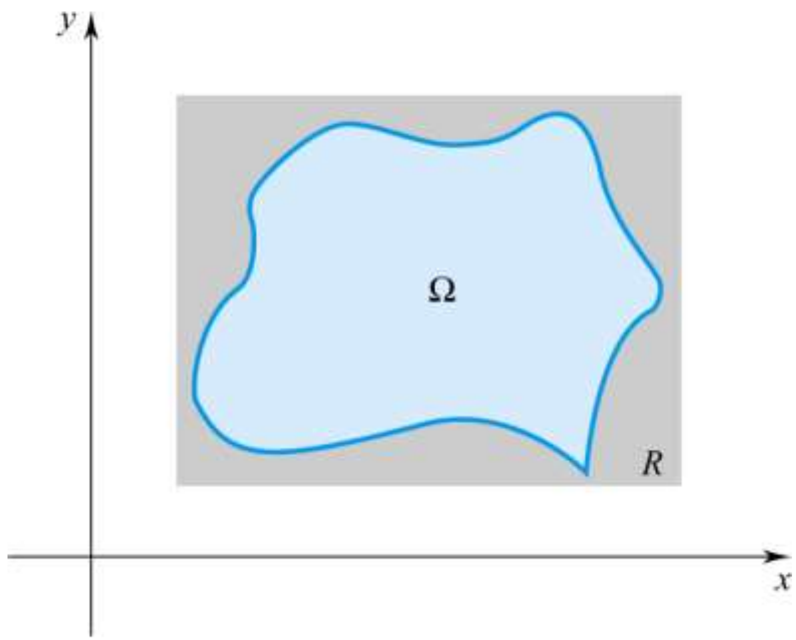


Figure 17.2.12

Double Integrals

If f is continuous and nonnegative over Ω, the extended f is nonnegative on all of R. The *volume of the solid T* bounded above by $z = f(x, y)$ and bounded below by Ω is given by:

(17.2.7)

$$\boxed{\text{volume of } T = \iint_{\Omega} f(x, y)\, dx\, dy.}$$

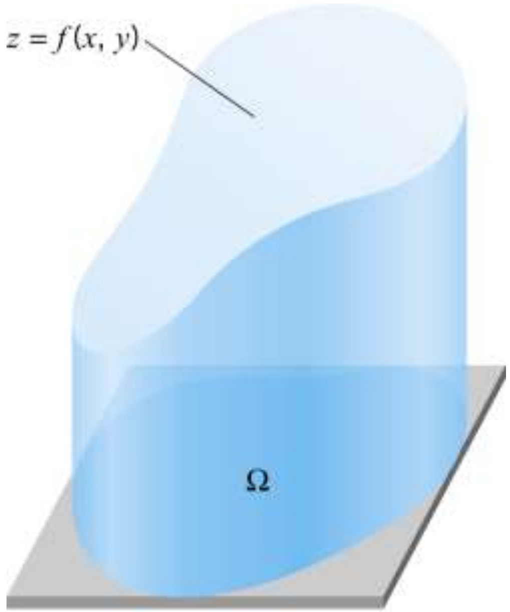


Figure 17.2.13

Double Integrals

Four Elementary Properties of the Double Integral:

I. *Linearity*: The double integral of a linear combination is the linear combination of the double integrals:

$$\iint_{\Omega} \left[\alpha f(x,y) + \beta g(x,y)\right] dx\, dy = \alpha \iint_{\Omega} f(x,y)\, dx\, dy + \beta \iint_{\Omega} g(x,y)\, dx\, dy$$

II. *Order*: The double integral preserves order:

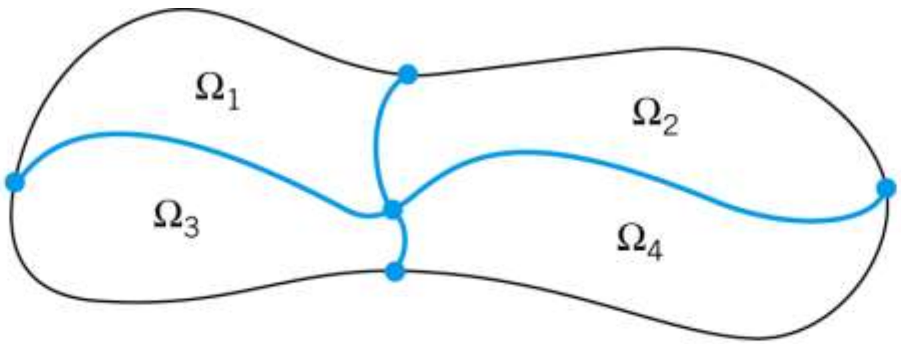
if $f \geq 0$ on Ω, then $\iint_{\Omega} f(x,y)\, dx\, dy \geq 0$

if $f \leq g$ on Ω, then $\iint_{\Omega} f(x,y)\, dx\, dy \leq \iint_{\Omega} g(x,y)\, dx\, dy$

Double Integrals

III. *Additivity*: If Ω is broken up into a finite number of nonoverlapping basic regions $\Omega_1, \ldots, \Omega_n$, then

$$\iint\limits_{\Omega} f(x,y)\,dx\,dy = \iint\limits_{\Omega_1} f(x,y)\,dx\,dy + \cdots + \iint\limits_{\Omega_n} f(x,y)\,dx\,dy$$



$$\iint\limits_{\Omega} f(x,y)\,dx\,dy = \iint\limits_{\Omega_1} f(x,y)\,dx\,dy + \iint\limits_{\Omega_2} f(x,y)\,dx\,dy + \iint\limits_{\Omega_3} f(x,y)\,dx\,dy + \iint\limits_{\Omega_4} f(x,y)\,dx\,dy$$

Figure 17.2.14

Double Integrals

IV. *Mean-value condition*: There is a point (x_0, y_0) in Ω for which

$$\iint_{\Omega} f(x,y)\,dx\,dy = f(x_0, y_0)\cdot(\text{area of }\Omega)$$

We call $f(x_0, y_0)$ *the average value of* f *on* Ω .

(17.2.9)
$$\boxed{\iint_{\Omega} f(x,y)\,dx\,dy = \begin{pmatrix}\text{the average value}\\ \text{of } f \text{ on } \Omega\end{pmatrix}\cdot(\text{area of }\Omega).}$$

THEOREM 17.2.10 THE MEAN-VALUE THEOREM FOR DOUBLE INTEGRALS

Let f and g be functions continuous on a basic region Ω. If g is nonnegative on Ω, then there exists a point (x_0, y_0) in Ω for which

$$\iint_{\Omega} f(x,y)g(x,y)\,dx\,dy = f(x_0,y_0)\iint_{\Omega} g(x,y)\,dx\,dy.^{\dagger}$$

We call $f(x_0, y_0)$ the *g-weighted average of f on* Ω.