MA106-Linear Algebra

Spring 2011

Anant R. Shastri

Department of Mathematics
Indian Institute of Technology, Bombay

January 19, 2011

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

▶ here

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

here

You are already familiar with how to solve such a system. Essentially, you carry out the following three types of operations on these equations which **do not alter the solutions** and convert the given system into a simpler system by eliminating variables one by one.

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

here

You are already familiar with how to solve such a system. Essentially, you carry out the following three types of operations on these equations which **do not alter the solutions** and convert the given system into a simpler system by eliminating variables one by one.

(1) Interchanging two equations.

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

here

You are already familiar with how to solve such a system. Essentially, you carry out the following three types of operations on these equations which **do not alter the solutions** and convert the given system into a simpler system by eliminating variables one by one.

(1) Interchanging two equations.

(2) Multiplying all the terms of an equation by a nonzero scalar.

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

here

You are already familiar with how to solve such a system. Essentially, you carry out the following three types of operations on these equations which **do not alter the solutions** and convert the given system into a simpler system by eliminating variables one by one.

(1) Interchanging two equations.

(2) Multiplying all the terms of an equation by a nonzero scalar.

(3) Adding to one equation a multiple of another equation.

Gauss Elimination Method

We discuss a widely used method called the Gauss elimination method to solve a system of m linear equations in n unknowns as given by (??).

here

You are already familiar with how to solve such a system. Essentially, you carry out the following three types of operations on these equations which **do not alter the solutions** and convert the given system into a simpler system by eliminating variables one by one.

(1) Interchanging two equations.

(2) Multiplying all the terms of an equation by a nonzero scalar.

(3) Adding to one equation a multiple of another equation.

Our aim is to put this method on a proper foundation and then reap a rich harvest that this method provides.

Gauss Elimination Method

In terms of matrix notation, (5) becomes

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{1}$$

The matrix $A = (a_{ij})$ is called the **coefficient matrix.** By a **solution** of (1) we mean any choice of $x_1, x_2, \dots, x_n$ which satisfies all the equations in it. If each $b_i = 0$ we say that the system is **homogeneous.** Otherwise it is called an **inhomogeneous system.**

GEM

Next we observe that there is no need to keep writing the name of the variables in the computations that follow. This leads us to the following now-standard method of representing a linear system by the matrix

$$\left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right] \tag{2}$$

which is called the *augmented matrix.* Now the above three operations on the equations correspond to certain operations on the row of the augmented matrix. These are called elementary row operations.

Gauss Elimination Method

We illustrate this method by three examples.

Example

A system with a unique solution: *Consider the system :*

$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 6z &= 10. \end{aligned}$$

The matrix form of this system is:

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{3}$$

Gauss Elimination Method

We illustrate this method by three examples.

Example

A system with a unique solution: *Consider the system :*

$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 6z &= 10. \end{aligned}$$

The matrix form of this system is:

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{3}$$

The three basic operations mentioned above will be performed on the rows of the augmented matrix as and when we need.

Gauss Elimination Method

We illustrate this method by three examples.

Example

A system with a unique solution: *Consider the system :*

$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 6z &= 10. \end{aligned}$$

The matrix form of this system is:

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{3}$$

The three basic operations mentioned above will be performed on the rows of the augmented matrix as and when we need.
The first step is to make the $(1,1)^{th}$ entry of the matrix equal to 1.

Gauss Elimination Method

We illustrate this method by three examples.

Example

A system with a unique solution: *Consider the system :*

$$\begin{aligned} 2x - 5y + 4z &= -3 \\ x - 2y + z &= 5 \\ x - 4y + 6z &= 10. \end{aligned}$$

The matrix form of this system is:

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{3}$$

The three basic operations mentioned above will be performed on the rows of the augmented matrix as and when we need.
The first step is to make the $(1,1)^{th}$ entry of the matrix equal to 1.

GEM

We can do this by multiplying all the entries in the first row by $1/2$.

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

The next step is to make all the entries in the first column zero except the $(1,1)^{th}$-entry.

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

The next step is to make all the entries in the first column zero except the $(1,1)^{th}$-entry.To do this, first we add -2 times the first row to the second row.

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

The next step is to make all the entries in the first column zero except the $(1,1)^{th}$-entry.To do this, first we add -2 times the first row to the second row.Then we subtract the first row from the third row.

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

The next step is to make all the entries in the first column zero except the $(1,1)^{th}$-entry.To do this, first we add -2 times the first row to the second row.Then we subtract the first row from the third row.
This yields the matrix :

GEM

We can do this by multiplying all the entries in the first row by $1/2$.
Alternatively, we can interchange the first and the second rows.
We get

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]. \tag{4}$$

The next step is to make all the entries in the first column zero except the $(1,1)^{th}$-entry.To do this, first we add -2 times the first row to the second row.Then we subtract the first row from the third row.
This yields the matrix :

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & -1 & 2 & -13 \\ 0 & -2 & 5 & 5 \end{array}\right]. \tag{5}$$

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.
So far, in effect we have succeeded in eliminaing the first variable from II and III equations.
Now we repeat the process for the smaller matrix.

$$\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right].$$

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.
So far, in effect we have succeeded in eliminaing the first variable from II and III equations.
Now we repeat the process for the smaller matrix.

$$\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right].$$

Multiply the top row by -1, to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ -2 & 5 & 5 \end{array}\right]. \tag{6}$$

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.
So far, in effect we have succeeded in eliminaing the first variable from II and III equations.
Now we repeat the process for the smaller matrix.

$$\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right].$$

Multiply the top row by -1, to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ -2 & 5 & 5 \end{array}\right]. \tag{6}$$

Now add twice the first row to the second to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ 0 & 1 & 31 \end{array}\right]. \tag{7}$$

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.
So far, in effect we have succeeded in eliminaing the first variable from II and III equations.
Now we repeat the process for the smaller matrix.

$$\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right].$$

Multiply the top row by -1, to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ -2 & 5 & 5 \end{array}\right]. \tag{6}$$

Now add twice the first row to the second to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ 0 & 1 & 31 \end{array}\right]. \tag{7}$$

Put back the rows and columns that you have cut out earlier to get

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]. \tag{8}$$

Gauss Elimination Method

This process from (4) to (5) is also called 'sweeping' a column.
Here we are sweeping the first column below the first row.
So far, in effect we have succeeded in eliminaing the first variable from II and III equations.
Now we repeat the process for the smaller matrix.

$$\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right].$$

Multiply the top row by -1, to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ -2 & 5 & 5 \end{array}\right]. \tag{6}$$

Now add twice the first row to the second to get

$$\left[\begin{array}{cc|c} 1 & -2 & 13 \\ 0 & 1 & 31 \end{array}\right]. \tag{7}$$

Put back the rows and columns that you have cut out earlier to get

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]. \tag{8}$$

GEM

The matrix in (8) represents the linear system

$$\begin{aligned} x - 2y + z &= 5 \\ y - 2z &= 13 \\ z &= 31. \end{aligned}$$

GEM

The matrix in (8) represents the linear system

$$\begin{aligned} x - 2y + z &= 5 \\ y - 2z &= 13 \\ z &= 31. \end{aligned}$$

These can be solved successively by **backward substitution.**

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**.

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**. Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0.

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**. Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0. Then we sweep the columns for which the diagonal entries are 1 above these entries.

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**.
Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0.
Then we sweep the columns for which the diagonal entries are 1 above these entries.
This is carried out in the decreasing order of the column numbers.

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**.
Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0.
Then we sweep the columns for which the diagonal entries are 1 above these entries.
This is carried out in the decreasing order of the column numbers.
So in (8), add twice third row to the second and then subtract the third row from the first.

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**.
Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0.
Then we sweep the columns for which the diagonal entries are 1 above these entries.
This is carried out in the decreasing order of the column numbers.
So in (8), add twice third row to the second and then subtract the third row from the first.
Next add twice the second row to the first. This gives

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right] \quad \cdots\cdots \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \end{array}\right]$$

Gauss-Jordan

Alternatively, we can continue the process called the **Gauss-Jordan Process**.
Here at this stage, we first make sure that all diagonal entries are indeed 1 or 0.
Then we sweep the columns for which the diagonal entries are 1 above these entries.
This is carried out in the decreasing order of the column numbers.
So in (8), add twice third row to the second and then subtract the third row from the first.
Next add twice the second row to the first. This gives

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right] \quad \cdots\cdots \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \end{array}\right]$$

The augmented matrix in (**??**) gives the final solution

$$(x, y, z)^T = (124, 75, 31)^T.$$

GEM

Notation

To save space further, we introduce notation for the elementary row operations. Let R_i denote the i^{th} row of a given matrix.

Operation	Notation	Right-multipy by
Multiply R_i by c	cR_i	$(I + (c-1)E_{ii})\cdot$
Add cR_j to R_i	$R_i + cR_j$	$(I + cE_{ij})\cdot$
Interchange R_i and R_j	$R_i \leftrightarrow R_j$	$T_{ij}\cdot$

GEM

Example

A system with infinitely many solutions: *Consider the following system of 3 equations in five unknowns:*

$$\begin{aligned} x - 2y + z - u + v &= 5 \\ 2x - 5y + 4z + u - v &= -3 \\ x - 4y + 6z + 2u - v &= 10 \end{aligned} \tag{9}$$

The corresponding augmented matrix is

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 5 \\ 2 & -5 & 4 & 1 & -1 & -3 \\ 1 & -4 & 6 & 2 & -1 & 10 \end{array}\right] \tag{10}$$

We shall use the notation introduced above for the row operations.

GEM

$$\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 2 & -5 & 4 & 1 & -1 & -3 \\ 1 & -4 & 6 & 2 & -1 & 10 \end{array}\right] \quad \begin{array}{c} R_2 - 2R_1 \\ R_3 - R_1 \\ \longrightarrow \end{array} \quad \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 0 & -1 & 2 & 3 & -3 & -13 \\ 0 & -2 & 5 & 3 & -2 & 5 \end{array}\right]$$

$$\begin{array}{c} R_3 - 2R_2 \\ \longrightarrow \end{array} \quad \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 0 & -1 & 2 & 3 & -3 & -13 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right] \quad \begin{array}{c} -R_2 \\ \longrightarrow \end{array} \quad \left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 0 & 1 & -2 & -3 & 3 & 13 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right]$$

$$\begin{array}{c} R_1 + 2R_2 \\ \longrightarrow \end{array} \quad \left[\begin{array}{ccccc|c} 1 & 0 & -3 & -7 & 7 & 31 \\ 0 & 1 & -2 & -3 & 3 & 13 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right] \quad \begin{array}{c} R_2 + 2R_3 \\ R_1 + 3R_3 \\ \longrightarrow \end{array} \quad \left[\begin{array}{ccccc|c} 1 & 0 & 0 & -16 & 19 & 124 \\ 0 & 1 & 0 & -9 & 11 & 75 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right]$$

GEM

The system of linear equations corresponding to the last augmented matrix is :

GEM

The system of linear equations corresponding to the last augmented matrix is :

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v. \end{aligned}$$

GEM

The system of linear equations corresponding to the last augmented matrix is :

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v. \end{aligned}$$

Assign arbitrary value t_1 to u and t_2 to v to obtain values of x, y, z.

GEM

The system of linear equations corresponding to the last augmented matrix is :

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v. \end{aligned}$$

Assign arbitrary value t_1 to u and t_2 to v to obtain values of x, y, z.
We say that u and v are **independent variables** and x, y, z are **dependent variables**.

GEM

The system of linear equations corresponding to the last augmented matrix is :

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v. \end{aligned}$$

Assign arbitrary value t_1 to u and t_2 to v to obtain values of x, y, z.
We say that u and v are **independent variables** and x, y, z are **dependent variables**.
Thus the set of all solutions $(x, y, z, u, v)^T$ of the system is given in the parametrized form:

GEM

The system of linear equations corresponding to the last augmented matrix is :

$$\begin{aligned} x &= 124 + 16u - 19v \\ y &= 75 + 9u - 11v \\ z &= 31 + 3u - 4v. \end{aligned}$$

Assign arbitrary value t_1 to u and t_2 to v to obtain values of x, y, z.
We say that u and v are **independent variables** and x, y, z are **dependent variables**.
Thus the set of all solutions $(x, y, z, u, v)^T$ of the system is given in the parametrized form:

$$\begin{aligned} & (x, y, z, u, v)^T \\ = \; & (124 + 16t_1 - 19t_2, 75 + 9t_1 - 11t_2, 31 + 3t_1 - 4t_2, t_1, t_2)^T \\ = \; & (124, 75, 31, 0, 0)^T + t_1(16, 9, 3, 1, 0)^T + t_2(-19, -11, -4, 0, 1)^T. \end{aligned}$$

GEM

The above equation gives the general solution to the system.

GEM

The above equation gives the general solution to the system.
The vector $(124, 75, 31, 0, 0)$ is a particular solution of the inhomogeneous system.

GEM

The above equation gives the general solution to the system.
The vector $(124, 75, 31, 0, 0)$ is a particular solution of the inhomogeneous system.
The vectors $\mathbf{v}_1 = (16, 9, 3, 1, 0)$ and $\mathbf{v}_2 = (-19, -11, -4, 0, 1)$ are solutions of the corresponding homogeneous system.

GEM

The above equation gives the general solution to the system.
The vector $(124, 75, 31, 0, 0)$ is a particular solution of the inhomogeneous system.
The vectors $\mathbf{v}_1 = (16, 9, 3, 1, 0)$ and $\mathbf{v}_2 = (-19, -11, -4, 0, 1)$ are solutions of the corresponding homogeneous system.
These two solutions are linearly independent and every other solution of the homogeneous system is a linear combination of these two solutions.

GEM

The above equation gives the general solution to the system.
The vector $(124, 75, 31, 0, 0)$ is a particular solution of the inhomogeneous system.
The vectors $\mathbf{v}_1 = (16, 9, 3, 1, 0)$ and $\mathbf{v}_2 = (-19, -11, -4, 0, 1)$ are solutions of the corresponding homogeneous system.
These two solutions are linearly independent and every other solution of the homogeneous system is a linear combination of these two solutions.
Finally every solution of the inhomogeneous system is the sum of a solution of the homogeneous system and one particular solution of the inhomogeneous system.

GEM

The above equation gives the general solution to the system.
The vector $(124, 75, 31, 0, 0)$ is a particular solution of the inhomogeneous system.
The vectors $\mathbf{v}_1 = (16, 9, 3, 1, 0)$ and $\mathbf{v}_2 = (-19, -11, -4, 0, 1)$ are solutions of the corresponding homogeneous system.
These two solutions are linearly independent and every other solution of the homogeneous system is a linear combination of these two solutions.
Finally every solution of the inhomogeneous system is the sum of a solution of the homogeneous system and one particular solution of the inhomogeneous system.
This illustrates a general principle which we now state and prove .

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.
Hence $\mathbf{r} - \mathbf{c} \in S$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.
Hence $\mathbf{r} - \mathbf{c} \in S$. Thus $\mathbf{r} \in \mathbf{c} + S$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.
Hence $\mathbf{r} - \mathbf{c} \in S$. Thus $\mathbf{r} \in \mathbf{c} + S$.
Conversely let $\mathbf{v} \in S$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.
Hence $\mathbf{r} - \mathbf{c} \in S$. Thus $\mathbf{r} \in \mathbf{c} + S$.
Conversely let $\mathbf{v} \in S$.
Then $A(\mathbf{c} + \mathbf{v}) = A\mathbf{c} + A\mathbf{v} = \mathbf{b} + \mathbf{0} = \mathbf{b}$.

GEM

Theorem

Suppose $A\mathbf{x} = \mathbf{b}$ *is a system of linear equations where* $A = (a_{ij})$ *is an* $m \times n$ *matrix and* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t, \mathbf{b} = (b_1, b_2, \ldots, b_m)^t$. *Suppose* $\mathbf{c} = (c_1, c_2, \ldots, c_n)^t$ *is a solution of* $A\mathbf{x} = \mathbf{b}$ *and* S *is the set of all solutions to the associated homogeneous system* $A\mathbf{x} = 0$. *Then the set of all solutions to* $A\mathbf{x} = \mathbf{b}$ *is*

$$\mathbf{c} + S := \{\mathbf{c} + \mathbf{v} | \mathbf{v} \in S\}.$$

Proof: Let $\mathbf{r} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$.
Then $A(\mathbf{r} - \mathbf{c}) = A\mathbf{r} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = 0$.
Hence $\mathbf{r} - \mathbf{c} \in S$. Thus $\mathbf{r} \in \mathbf{c} + S$.
Conversely let $\mathbf{v} \in S$.
Then $A(\mathbf{c} + \mathbf{v}) = A\mathbf{c} + A\mathbf{v} = \mathbf{b} + \mathbf{0} = \mathbf{b}$.
Hence $\mathbf{c} + \mathbf{v}$ is a solution to $A\mathbf{x} = \mathbf{b}$. ♠

GEM

Example

A system with no solution: *Consider the system*

$$\begin{aligned} x - 5y + 4z &= 3 \\ x - 5y + 3z &= 6 \\ 2x - 10y + 13z &= 5 \end{aligned} \tag{11}$$

GEM

Example

A system with no solution: *Consider the system*

$$\begin{aligned} x - 5y + 4z &= 3 \\ x - 5y + 3z &= 6 \\ 2x - 10y + 13z &= 5 \end{aligned} \tag{11}$$

The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 1 & -5 & 4 & 3 \\ 1 & -5 & 3 & 6 \\ 2 & -10 & 13 & 5 \end{array}\right] \tag{12}$$

GEM

Example

A system with no solution: *Consider the system*

$$\begin{aligned} x - 5y + 4z &= 3 \\ x - 5y + 3z &= 6 \\ 2x - 10y + 13z &= 5 \end{aligned} \tag{11}$$

The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 1 & -5 & 4 & 3 \\ 1 & -5 & 3 & 6 \\ 2 & -10 & 13 & 5 \end{array}\right] \tag{12}$$

Apply Gauss-Elimination Method to (12) to get

$$\left[\begin{array}{ccc|c} 1 & -5 & 4 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 14 \end{array}\right] \tag{13}$$

GEM

Example

A system with no solution: *Consider the system*

$$\begin{aligned} x - 5y + 4z &= 3 \\ x - 5y + 3z &= 6 \\ 2x - 10y + 13z &= 5 \end{aligned} \tag{11}$$

The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 1 & -5 & 4 & 3 \\ 1 & -5 & 3 & 6 \\ 2 & -10 & 13 & 5 \end{array}\right] \tag{12}$$

Apply Gauss-Elimination Method to (12) to get

$$\left[\begin{array}{ccc|c} 1 & -5 & 4 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 14 \end{array}\right] \tag{13}$$

The bottom row corresponds to the equation $0.z = 14$. *Hence the system (11) has no solutions.*

GEM

Observe that we had plenty of freedom to choose the order of equations. This becomes a boon in numerical methods.

GEM

Observe that we had plenty of freedom to choose the order of equations. This becomes a boon in numerical methods.
However, it is also necessary to curtail this freedom in order to bring some uniformity, viz., once a set of equations is presented to us in a particular order, we should change the order of the equations only if it is very much necessary.

GEM

Observe that we had plenty of freedom to choose the order of equations. This becomes a boon in numerical methods.
However, it is also necessary to curtail this freedom in order to bring some uniformity, viz., once a set of equations is presented to us in a particular order, we should change the order of the equations only if it is very much necessary.
For the same reason, making leading entry equal to 1 is not completely necessary at the beginning. Neither there is any need to sweep the columns above a particular row.

GEM

Observe that we had plenty of freedom to choose the order of equations. This becomes a boon in numerical methods.
However, it is also necessary to curtail this freedom in order to bring some uniformity, viz., once a set of equations is presented to us in a particular order, we should change the order of the equations only if it is very much necessary.
For the same reason, making leading entry equal to 1 is not completely necessary at the beginning. Neither there is any need to sweep the columns above a particular row.
With these modification the above method becomes a well defined powerful algorithm thereby providing various advantages. We shall refer to it by Gauss Elimination Method(GEM).

GEM

Observe that we had plenty of freedom to choose the order of equations. This becomes a boon in numerical methods.
However, it is also necessary to curtail this freedom in order to bring some uniformity, viz., once a set of equations is presented to us in a particular order, we should change the order of the equations only if it is very much necessary.
For the same reason, making leading entry equal to 1 is not completely necessary at the beginning. Neither there is any need to sweep the columns above a particular row.
With these modification the above method becomes a well defined powerful algorithm thereby providing various advantages. We shall refer to it by Gauss Elimination Method(GEM).
However, we have also observed that the last two sets of operations have their own advantages. Since there is no ambiguity in them, we shall incorporate them in to a second stage (extended) algorithm called Gauss-Jordan Method (GJEM).
Below we describe these two methods in algorithmic form. Let us begin with some definitions.

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix* (respectively, *reduced echelon matrix*)

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that
(i) $a_{i,j} = 0$ for all $i > r$ and for all j.

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that
(i) $a_{i,j} = 0$ for all $i > r$ and for all j.
(ii) for each $1 \leq i \leq r$, $a_{i,j} = 0$ for $j < k(i)$.

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that
(i) $a_{i,j} = 0$ for all $i > r$ and for all j.
(ii) for each $1 \leq i \leq r$, $a_{i,j} = 0$ for $j < k(i)$.
(iii) for each $1 \leq i \leq r$, $a_{i,k(i)} \neq 0$ (respectively, $= 1$).

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that
(i) $a_{i,j} = 0$ for all $i > r$ and for all j.
(ii) for each $1 \leq i \leq r$, $a_{i,j} = 0$ for $j < k(i)$.
(iii) for each $1 \leq i \leq r$, $a_{i,k(i)} \neq 0$ (respectively, $= 1$).
(iv) for each $1 \leq i \leq r$, $a_{s,k(i)} = 0$, for all $s > i$ (respectively, for $s \neq i$).

GEM

Definition

A $m \times n$ matrix $A = ((a_{i,j}))$ is called *echelon matrix*
(respectively, *reduced echelon matrix*)
if $A = 0$ or if there exists an integer $r,\ 1 \leq r \leq min\{m, n\}$ and integers

$$1 \leq k(1) < k(2) < \cdots < k(r) \leq m$$

such that
(i) $a_{i,j} = 0$ for all $i > r$ and for all j.
(ii) for each $1 \leq i \leq r$, $a_{i,j} = 0$ for $j < k(i)$.
(iii) for each $1 \leq i \leq r$, $a_{i,k(i)} \neq 0$ (respectively, $= 1$).
(iv) for each $1 \leq i \leq r$, $a_{s,k(i)} = 0$, for all $s > i$ (respectively, for $s \neq i$).

$$G(A) = \begin{bmatrix} \phi & * & * & * & * & * \\ 0 & \phi & * & * & * & * \\ 0 & 0 & 0 & \phi & * & * \\ & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad J(A) = \begin{bmatrix} 1 & 0 & * & 0 & * & * \\ 0 & 1 & * & 0 & * & * \\ 0 & 0 & 0 & 1 & * & * \\ & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

GEM

Remark

- (i) Echelon matrices can be defined with respect to columns also. Since by taking transpose, we can easily convert one theory into the other there is no need to study them separately. In the literature there are slightly different conditions in the definition of an echelon matrix and also different names. It would be appropriate to call echelon matrices **Gauss matrices** and the reduced echelon matrices **Gauss-Jordan matrices**.

GEM

Remark

- (i) Echelon matrices can be defined with respect to columns also. Since by taking transpose, we can easily convert one theory into the other there is no need to study them separately. In the literature there are slightly different conditions in the definition of an echelon matrix and also different names. It would be appropriate to call echelon matrices **Gauss matrices** and the reduced echelon matrices **Gauss-Jordan matrices**.
- (ii) The integer r occurring in the definition is quite significant. Let us call it number of steps. The columns $C_{k(s)}$ are called the **step columns.** The $(i, k(i))^{th}$ entries which are non zero in $G(A)$ (and equal to 1 in $J(A)$ are called **pivots.**

GEM

Remark

- (i) Echelon matrices can be defined with respect to columns also. Since by taking transpose, we can easily convert one theory into the other there is no need to study them separately. In the literature there are slightly different conditions in the definition of an echelon matrix and also different names. It would be appropriate to call echelon matrices **Gauss matrices** and the reduced echelon matrices **Gauss-Jordan matrices**.
- (ii) The integer r occurring in the definition is quite significant. Let us call it number of steps. The columns $C_{k(s)}$ are called the **step columns.** The $(i, k(i))^{th}$ entries which are non zero in $G(A)$ (and equal to 1 in $J(A)$ are called **pivots.**
- (iii) It is convenient for later purpose to introduce a notation for the set of indices j such that the j^{th} column of A is not a step. So, let us denote this set by $\rho(A)$ (which is a subset of N).

GEM

Remark

- (i) Echelon matrices can be defined with respect to columns also. Since by taking transpose, we can easily convert one theory into the other there is no need to study them separately. In the literature there are slightly different conditions in the definition of an echelon matrix and also different names. It would be appropriate to call echelon matrices **Gauss matrices** and the reduced echelon matrices **Gauss-Jordan matrices**.
- (ii) The integer r occurring in the definition is quite significant. Let us call it number of steps. The columns $C_{k(s)}$ are called the **step columns.** The $(i, k(i))^{th}$ entries which are non zero in $G(A)$ (and equal to 1 in $J(A)$ are called **pivots.**
- (iii) It is convenient for later purpose to introduce a notation for the set of indices j such that the j^{th} column of A is not a step. So, let us denote this set by $\rho(A)$ (which is a subset of N).
- (iv) We have already seen enough examples of echelon matrices. Indeed in the examples of solutions of systems of equations that we discussed, the end product was always an echelon matrix.

Algorithm for GEM

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X. If you cannot find one, cut down the first column of X and call the new matrix as X and go back to step 1.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X. If you cannot find one, cut down the first column of X and call the new matrix as X and go back to step 1.Otherwise find the first non zero entry in the first column say, $=x_{i,1}$.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X. If you cannot find one, cut down the first column of X and call the new matrix as X and go back to step 1.Otherwise find the first non zero entry in the first column say, $=x_{i,1}$.
Step 3. Swap the first row and i^{th} rows of X. (Swap the corresponding rows of A also.)

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X. If you cannot find one, cut down the first column of X and call the new matrix as X and go back to step 1.Otherwise find the first non zero entry in the first column say, $=x_{i,1}$.
Step 3. Swap the first row and i^{th} rows of X. (Swap the corresponding rows of A also.)
Step 4. For each $j \geq 2$, add $-(x_{j,1}/x_{1,1})R_1$ to R_j of X; perform the corresponding row operations on the matrix A also.

Algorithm for GEM

Theorem

Let $A = ((a_{ij}))$ be a $m \times n$ matrix. The following algorithm will convert A into an echelon matrix Z.
Step 0 Put $X = A$ and $X = ((x_{i,j}))$.
Step 1. If X is the zero matrix or empty then declare $Z = A$ and stop.Else go to step 2.
Step 2. Look for a non zero entry in the first column of X. If you cannot find one, cut down the first column of X and call the new matrix as X and go back to step 1.Otherwise find the first non zero entry in the first column say, $=x_{i,1}$.
Step 3. Swap the first row and i^{th} rows of X. (Swap the corresponding rows of A also.)
Step 4. For each $j \geq 2$, add $-(x_{j,1}/x_{1,1})R_1$ to R_j of X; perform the corresponding row operations on the matrix A also.
Step 5. Cut down both the first column and the first row of X and call the new matrix as X and go back to step 1.

GEM

Theorem

Algorithm for GJEM *Let A be an echelon matrix. The following row operations on A convert it into a row reduced echelon matrix: For $1 \leq i \leq r$ where r is the number of steps in A, let $a_{i,k(i)}$ be the first non zero entry in the i^{th} row. Divide the i^{th} row of A by $a_{i,k(i)}$ and denote the new matrix by $A' = ((a'_{ij}))$. Now for each $2 \leq i \leq r$ add $-a'_{j,k(i)}R_i$ to R_j for all $1 \leq j < i$.*

Proof: The first set of operations make all the pivots equal to 1. Carry out the second set of operations in the decreasing order of i for $2 \leq i \leq r$. This ensures that the only the entries above the i^{th} pivot becomes zero at the i^{th} stage and other pivot columns are not affected. This gives the desired result. ♠

GEM

Remark

To sum up, we see that the algorithms perform a finite sequence of elementary row operations on the given matrix and then successively on the new resulting matrices until the desired result is obtained. Each such operation can be thought of as multiplication on left by a corresponding elementary matrix. Since each elementary matrix is invertible and product of invertible matrices is invertible, we conclude:

Theorem

Given any matrix A there exist an invertible matrix R such that $RA = G(A)$ (respectively, $= J(A)$)) is a echelon (respectively, reduced echelon) matrix.

We shall call $G(A)$ the **Gauss form** of A and $J(A)$ the **Gauss-Jordan form** of A. Observe that by the very definition which is algorithmic $G(A)$ and $J(A)$ are well defined.
Finally we sum up the rule for solving a system of linear equations:

GEM

HOW TO SOLVE A SYSTEM OF LINEAR EQUATIONS $A\mathbf{x} = \mathbf{b}$.

Step 1 Apply JGEM to $[A|\mathbf{b}]$; get Gauss-Jordan form: $J([A|\mathbf{b}]) = [J(A)|\mathbf{b}']$.
Let r be the number of nonzero rows of $J(A)$ and let
$k(1) < \cdots < k(r)$ be the step columns (pivot columns).

Step 3 If $b'_i \neq 0$ for any $i > r$ declare that the system is **inconsistent,** and exit.

Step 4 If $r = n$ (the number of columns of A)
declare that the system has a unique solution
viz., $(x_1, \ldots, x_n) = (b'_1, \ldots, b'_n)$ and exit; if not go to step 5.

Step 5 In $[J(A)|\mathbf{b}']$ delete all the zero-rows.

Step 6 Let the indices corresponding to the non-step columns be
$l(1) < l(2) < \cdots, < l(s)$, $(s + r = n)$.
Keeping the order, transfer all the non-step columns to the right side,
to the right of $\mathbf{b}'$ and change their sign to obtain a matrix
of the form $[I_r|B']$. The general solution is given by putting

$$x_{l(i)} = t_i, 1 \leq i \leq s; \text{ and } (x_{k(1)}, \ldots, x_{k(r)})^T = B'(1, t_1, \ldots, t_s)^T.$$

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.
(b) Suppose now $r = m$. (Observe that this automatically implies that $n \geq m = r$.) Then irrespective of the vector $\mathbf{b}$ we know that $A\mathbf{x} = \mathbf{b}$ will have at least one solution. Indeed, in this case, we get infinitely many solutions unless

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.
(b) Suppose now $r = m$. (Observe that this automatically implies that $n \geq m = r$.) Then irrespective of the vector $\mathbf{b}$ we know that $A\mathbf{x} = \mathbf{b}$ will have at least one solution. Indeed, in this case, we get infinitely many solutions unless
(c) $r = m = n$ when $A\mathbf{x} = \mathbf{b}$ has precisely one solution for all $\mathbf{x} \in \mathbb{R}^m$.

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.
(b) Suppose now $r = m$. (Observe that this automatically implies that $n \geq m = r$.) Then irrespective of the vector $\mathbf{b}$ we know that $A\mathbf{x} = \mathbf{b}$ will have at least one solution. Indeed, in this case, we get infinitely many solutions unless
(c) $r = m = n$ when $A\mathbf{x} = \mathbf{b}$ has precisely one solution for all $\mathbf{x} \in \mathbb{R}^m$.
From GJEM, we know that

$$J(A) = RA$$

where R is an invertible matrix.

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.
(b) Suppose now $r = m$. (Observe that this automatically implies that $n \geq m = r$.) Then irrespective of the vector $\mathbf{b}$ we know that $A\mathbf{x} = \mathbf{b}$ will have at least one solution. Indeed, in this case, we get infinitely many solutions unless
(c) $r = m = n$ when $A\mathbf{x} = \mathbf{b}$ has precisely one solution for all $\mathbf{x} \in \mathbb{R}^m$.
From GJEM, we know that

$$J(A) = RA$$

where R is an invertible matrix.
On the other hand when $r = m = n$ we also that $J(A) = Id_n$.

Significance of the number of Steps in GEM

Let r be the number of steps that occur in the GEM for a given $m \times n$ matrix.
(a) If $r < m$ there is a possibility that the system $A\mathbf{x} = \mathbf{b}$ may not have any solution. This depends on the vector $\mathbf{b}$. Indeed, it is always possible to find such a $\mathbf{b}$.
(b) Suppose now $r = m$. (Observe that this automatically implies that $n \geq m = r$.) Then irrespective of the vector $\mathbf{b}$ we know that $A\mathbf{x} = \mathbf{b}$ will have at least one solution. Indeed, in this case, we get infinitely many solutions unless
(c) $r = m = n$ when $A\mathbf{x} = \mathbf{b}$ has precisely one solution for all $\mathbf{x} \in \mathbb{R}^m$.
From GJEM, we know that

$$J(A) = RA$$

where R is an invertible matrix.
On the other hand when $r = m = n$ we also that $J(A) = Id_n$. Therefore

$$A = (R^{-1}R)A = R^{-1}(RA) = R^{-1}J(A) = R^{-1}.$$

This is the same as saying A is invertible and $A^{-1} = R$.

GEM applied to find the inverse

Given an $n \times n$ matrix A apply GJEM to

$$[A|Id_n].$$

If your end result is

$$[Id_n|R]$$

then declare $R = A^{-1}$. Otherwise declare that A is not invertible.

Definition

The number of steps r in the GEM applied to A is called the row rank of A.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.
It is clear that such a transformation does not change the volume of any solid.
Obviously, this cannot be true of all linear transformations.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.
It is clear that such a transformation does not change the volume of any solid.
Obviously, this cannot be true of all linear transformations. But then how to measure the change in the volume effected by a linear transformation. An answer is in the notion of

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.
It is clear that such a transformation does not change the volume of any solid.
Obviously, this cannot be true of all linear transformations. But then how to measure the change in the volume effected by a linear transformation. An answer is in the notion of
determinants.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.
It is clear that such a transformation does not change the volume of any solid. Obviously, this cannot be true of all linear transformations. But then how to measure the change in the volume effected by a linear transformation. An answer is in the notion of
determinants.
For example consider the map $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by

$$\mathbf{v} \mapsto \lambda \mathbf{v}$$

where $\lambda \in \mathbb{R}$.

Determinants

We are now interested in studying linear transformation from a space to itself:

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

These are also called endomorphisms.
We have seen that a rigid motion which fixes the origin is an invertible linear transformation of a euclidean space.
It is clear that such a transformation does not change the volume of any solid. Obviously, this cannot be true of all linear transformations. But then how to measure the change in the volume effected by a linear transformation. An answer is in the notion of
determinants.
For example consider the map $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by

$$\mathbf{v} \mapsto \lambda \mathbf{v}$$

where $\lambda \in \mathbb{R}$. For $n = 1$ it is clear that a segment of length ℓ in $\mathbb{R}$ (no matter where it is situated in $\mathbb{R}$) is mapped to a segment of length $\ell\lambda$.

Determinants

Likewise it is easily seen that a cube of volume m in $\mathbb{R}^3$ is mapped to a solid of volume $m\lambda^3$.

Determinants

Likewise it is easily seen that a cube of volume m in $\mathbb{R}^3$ is mapped to a solid of volume $m\lambda^3$.
In $\mathbb{R}^n$, the effect on the volume is to multiply by λ^n. The matrix of λ is the diagonal $\mathrm{diag}(\lambda, \lambda, \ldots, \lambda)$.
A little more generally, if we consider the endomorphism which effects a scaling by λ_i in the i^{th} direction, then it is clear that the volume of an n-dimensional cube is changed by a factor of $\lambda_1 \cdots \lambda_n$.

Determinants

Likewise it is easily seen that a cube of volume m in $\mathbb{R}^3$ is mapped to a solid of volume $m\lambda^3$.
In $\mathbb{R}^n$, the effect on the volume is to multiply by λ^n. The matrix of λ is the diagonal $\mathrm{diag}(\lambda, \lambda, \ldots, \lambda)$.
A little more generally, if we consider the endomorphism which effects a scaling by λ_i in the i^{th} direction, then it is clear that the volume of an n-dimensional cube is changed by a factor of $\lambda_1 \cdots \lambda_n$.
These considerations lead us to believe that each linear transformation effects a scaling in the volume. We must hunt for this scaling factor.

Determinants

Example

Consider $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *such that* $T(1,0) = (a,b) =: \mathbf{v}_1$ *and* $T(0,1) = (c,d) = \mathbf{v}_2$.

Determinants

Example

Consider $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *such that* $T(1,0) = (a,b) =: \mathbf{v}_1$ *and* $T(0,1) = (c,d) = \mathbf{v}_2$.

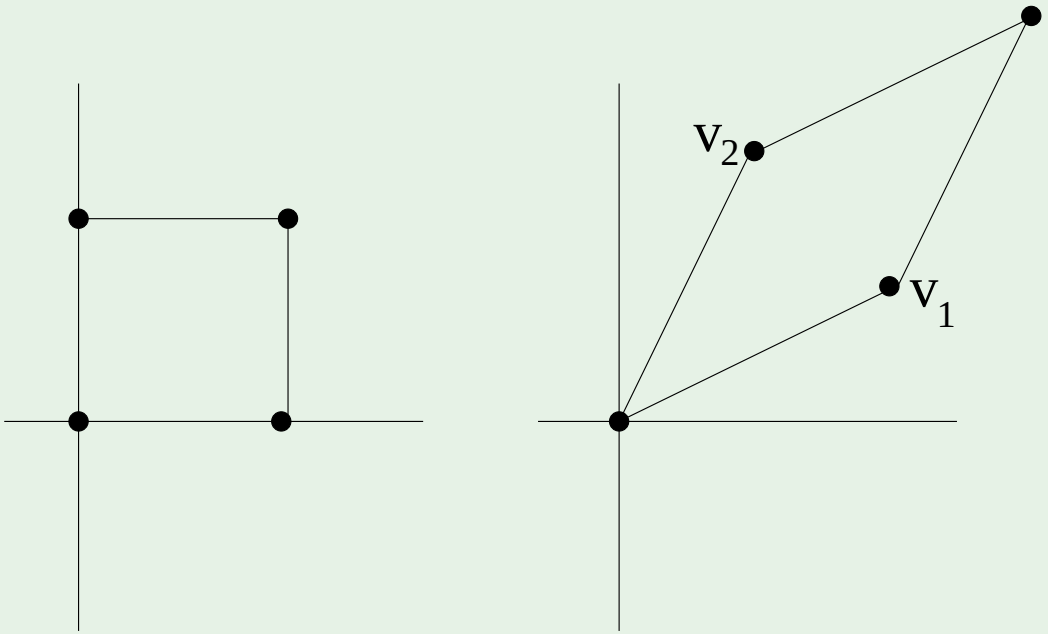
This means that T *maps the unit square in* $\mathbb{R}^2$ *to the parallelogram with its vertices* $(0,0), (a,b), (c,d), (a+c, b+d)$.

Determinants

Example

Consider $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *such that* $T(1,0) = (a,b) =: \mathbf{v}_1$ *and* $T(0,1) = (c,d) = \mathbf{v}_2$.

This means that T *maps the unit square in* $\mathbb{R}^2$ *to the parallelogram with its vertices* $(0,0), (a,b), (c,d), (a+c, b+d)$.



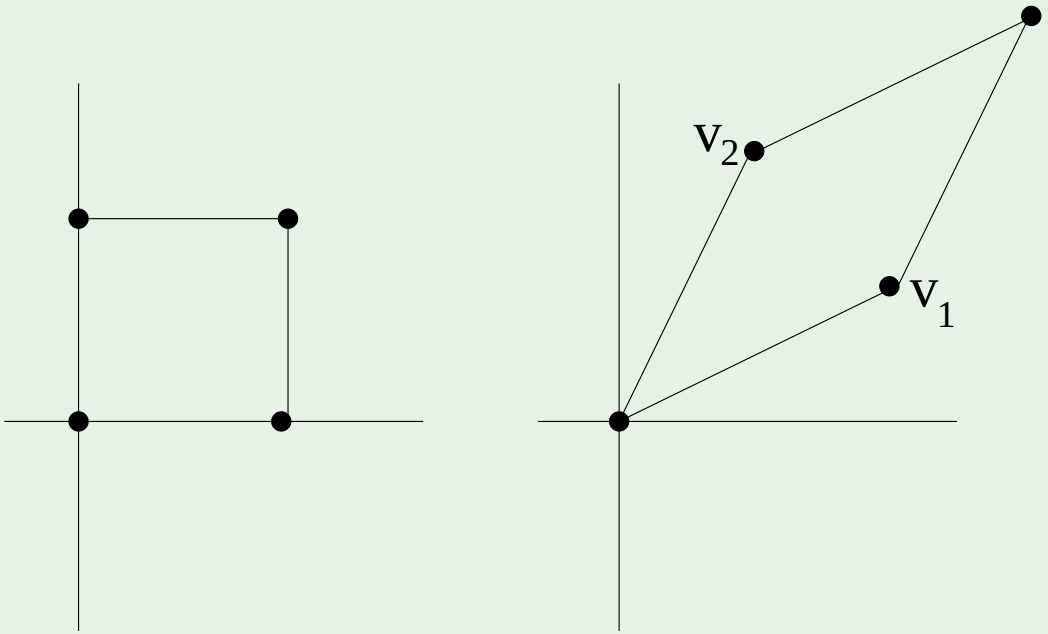
The area of this parallelogram is given by $|ad - bc|$.

Determinants

Example

Consider $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *such that* $T(1,0) = (a,b) =: \mathbf{v}_1$ *and* $T(0,1) = (c,d) = \mathbf{v}_2$.

This means that T *maps the unit square in* $\mathbb{R}^2$ *to the parallelogram with its vertices* $(0,0), (a,b), (c,d), (a+c, b+d)$.



The area of this parallelogram is given by $|ad - bc|$.

In terms of vector product it is $|\mathbf{v}_1 \times \mathbf{v}_2|$.

Determinants

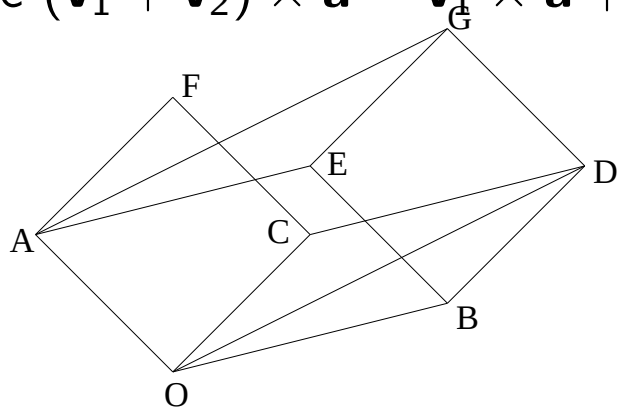
If we do not take the modulus, and only take the value $ad - bc$ it can thought of as a 'signed' area which is allowed to be negative. The vector product itself has the property viz.,

$$\mathbf{v}_1 \times \mathbf{v}_2 = -\mathbf{v}_2 \times \mathbf{v}_1.$$

We also have

$$(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}.$$

The following simple diagram illustrates this and affirms that we are on the right track. Namely, we have $(\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{u} = \mathbf{v}_1 \times \mathbf{u} + \mathbf{v}_2 \times \mathbf{u}$.



OBEA+OCFA=OBEA+BDGE−OBD+AEG=ODGA

Geometric proof of the additivity of the area

(Caution: The above figure is not a 3-dimensional one!)

Determinants

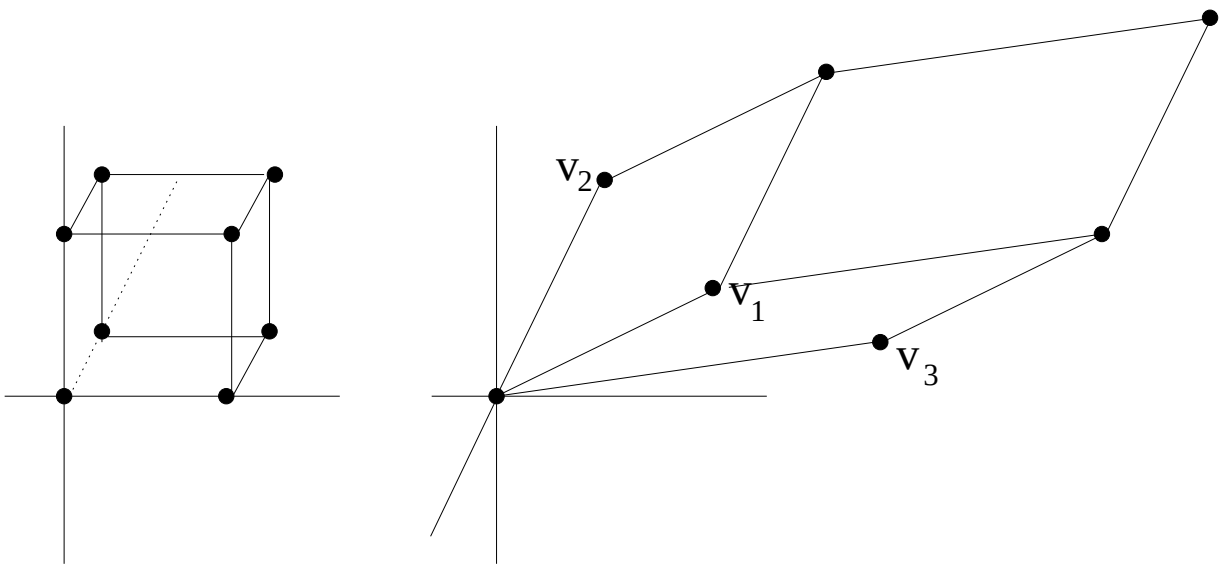
Example

Likewise, if $T' : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *is such that* $T'(\mathbf{e}_i) = \mathbf{v}_i$ *then the volume of the parallelepiped defined by the vectors* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *is given by the modulus of the scalar triple product* $(\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3)$.

Determinants

Example

Likewise, if $T' : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *is such that* $T'(\mathbf{e}_i) = \mathbf{v}_i$ *then the volume of the parallelepiped defined by the vectors* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *is given by the modulus of the scalar triple product* $(\mathbf{v}_1\mathbf{v}_2\mathbf{v}_3)$*.We can even write down the formula for this in terms of coordinates of* $\mathbf{v}_i$*, if we wish.*



Axioms for determinant

What do we do in the general case?

Axioms for determinant

What do we do in the general case? We would like to approach this problem as follows:

Axioms for determinant

What do we do in the general case? We would like to approach this problem as follows:First we ask what are the basic properties of this 'creature' that we are seeking.

Axioms for determinant

What do we do in the general case? We would like to approach this problem as follows:First we ask what are the basic properties of this 'creature' that we are seeking. After listing sufficiently many of these properties in an economic way, we shall launch a search party to look out for the creature.

Axioms for determinant

What do we do in the general case? We would like to approach this problem as follows:First we ask what are the basic properties of this 'creature' that we are seeking. After listing sufficiently many of these properties in an economic way, we shall launch a search party to look out for the creature.
In what follows we shall use the symbol $\mathbb{K}$ to denote either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complx numbers. (Indeed, for many of the concepts and results, we could have taken $\mathbb{K}$ to be just any field.)

Axioms for determinant

What do we do in the general case? We would like to approach this problem as follows:First we ask what are the basic properties of this 'creature' that we are seeking. After listing sufficiently many of these properties in an economic way, we shall launch a search party to look out for the creature.
In what follows we shall use the symbol $\mathbb{K}$ to denote either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complx numbers. (Indeed, for many of the concepts and results, we could have taken $\mathbb{K}$ to be just any field.)
We are looking for a $\mathbb{K}$-valued function defined on the set of $n \times n$ matrices (for each n) say $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ which is likely to represent the volume of the n-dim. parallelepiped defined by the row vectors of a given matrix.

Hunt for determinant

- (i) If two rows of a matrix A are identical then $f(A) = 0$.

Hunt for determinant

- (i) If two rows of a matrix A are identical then $f(A) = 0$.
- (ii) Look at the sign of the area of the parallelogram. Depending upon the order in which we measure it, the sign of the area could be positive or negative. *If we change any two rows of A then the effect on $f(A)$ must be to multiply by -1.*

Hunt for determinant

- (i) If two rows of a matrix A are identical then $f(A) = 0$.
- (ii) Look at the sign of the area of the parallelogram. Depending upon the order in which we measure it, the sign of the area could be positive or negative. *If we change any two rows of A then the effect on $f(A)$ must be to multiply by -1.*
- (iii) We have already observed that if we scale one of the vectors then the entire volume is also scaled by the same factor.
- Likewise the volume is 'additive, in each of the slots,

$$((\mathbf{u}_1 + \mathbf{u}_2)\mathbf{vw}) = (\mathbf{u}_1\mathbf{vw}) + (\mathbf{u}_2\mathbf{vw}).$$

Hunt for determinant

With these things in mind we shall now attack the problem systematically. Given a matrix $A \in M_n(\mathbb{K})$ say,

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix},$$

recall that by arranging the n rows of A one after another we can think of it as an ordered n-tuple of n-vectors: $A = (A_1, \ldots, A_n)$.

Hunt for determinant

With these things in mind we shall now attack the problem systematically. Given a matrix $A \in M_n(\mathbb{K})$ say,

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix},$$

recall that by arranging the n rows of A one after another we can think of it as an ordered n-tuple of n-vectors: $A = (A_1, \ldots, A_n)$. More generally, this method may be used to identify $M_{n,m}(\mathbb{K})$ with $\mathbb{K}^{mn}$. We are going to exploit it now.

Axioms for determinant

Definition

Let $1 \leq i \leq n$. A function $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called *linear in the i^{th} row* if, $f(A_1, \ldots, A_{i-1}, \alpha A + \beta B, A_{i+1}, \ldots, A_n) =$

$$\alpha f(A_1, \ldots, A_{i-1}, A, A_{i+1}, \ldots, A_n) + \beta f(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n) \qquad (14)$$

for all vectors $A_j, A, B \in \mathbb{K}^n$ and all scalars $\alpha, \beta \in \mathbb{K}$. A function f which is linear in i^{th} row for each $1 \leq i \leq n$ will be called **multi-linear.**

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.
Remember that a linear map was completely determined by its values on the standard vectors $\mathbf{e}_i^T.'s$ We are going to exploit this now.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.
Remember that a linear map was completely determined by its values on the standard vectors $\mathbf{e}_i^T.'s$ We are going to exploit this now.
Let $N = \{1, 2, \ldots, n\}$ and let $F(N, N)$ be the set of all maps $\sigma : N \longrightarrow N$.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.
Remember that a linear map was completely determined by its values on the standard vectors $\mathbf{e}_i^T.'s$ We are going to exploit this now.
Let $N = \{1, 2, \ldots, n\}$ and let $F(N, N)$ be the set of all maps $\sigma : N \longrightarrow N$.
(This set can also be thought of as the set of all sequences of length n with values between 1 and n.)

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.
Remember that a linear map was completely determined by its values on the standard vectors $\mathbf{e}_i^T.'s$ We are going to exploit this now.
Let $N = \{1, 2, \ldots, n\}$ and let $F(N, N)$ be the set of all maps $\sigma : N \longrightarrow N$.
(This set can also be thought of as the set of all sequences of length n with values between 1 and n.)
Let E_σ denote the matrix with its i^{th} row equal to $\mathbf{e}_{\sigma(i)}^T$.

Axioms for determinant

Examples of multi-linear maps that we have are the vector product of two vectors in case $n = 2$ and the scalar triple product in case $n = 3$.
We do not know many examples.
This should not deter us from our goal.
Remember that a linear map was completely determined by its values on the standard vectors $\mathbf{e}_i^T.'s$ We are going to exploit this now.
Let $N = \{1, 2, \ldots, n\}$ and let $F(N, N)$ be the set of all maps $\sigma : N \longrightarrow N$.
(This set can also be thought of as the set of all sequences of length n with values between 1 and n.)
Let E_σ denote the matrix with its i^{th} row equal to $\mathbf{e}_{\sigma(i)}^T$.
For example, in this notation if $\sigma = Id$ then $E_\sigma = I_n$.
If $\sigma_1 = (1, 1, 2, 2), \sigma_2 = (3, 2, 1)$ say then

$$E_{\sigma_1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad E_{\sigma_2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Write down a few more examples to get familiar with this notation.

Axioms for determinant

Theorem

Any multi-linear map f is completely determined by its values on $E_\sigma, \sigma \in \Lambda$.

Axioms for determinant

Theorem

Any multi-linear map f is completely determined by its values on $E_\sigma, \sigma \in \Lambda$.

Proof: Let us first illustrate this for $n = 2$. Write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{15}$$

Then $A = (A_1, A_2)$ where $A_1 = (a, b)$ and $A_2 = (c, d)$.

Axioms for determinant

Theorem

Any multi-linear map f is completely determined by its values on $E_\sigma, \sigma \in \Lambda$.

Proof: Let us first illustrate this for $n = 2$. Write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{15}$$

Then $A = (A_1, A_2)$ where $A_1 = (a, b)$ and $A_2 = (c, d)$.
Further $A_1 = a\mathbf{e}_1^T + b\mathbf{e}_2^T$ and $A_2 = c\mathbf{e}_1^T + d\mathbf{e}_2^T$. Therefore

$$\begin{aligned} f(A) &= af(\mathbf{e}_1^T, A_2) + bf(\mathbf{e}_2^T, A_2) \\ &= a(cf(\mathbf{e}_1^T, \mathbf{e}_1^T) + df(\mathbf{e}_1^T, \mathbf{e}_2^T)) + b(cf(\mathbf{e}_2^T, \mathbf{e}_1^T) + df(\mathbf{e}_2^T, \mathbf{e}_2^T)) \\ &= ac\, f(E_{(11)}) + ad\, f(E_{(12)}) + bc\, f(E_{(21)}) + bd\, f(E_{(22)}). \end{aligned}$$

Axioms for determinant

Theorem

Any multi-linear map f is completely determined by its values on $E_\sigma, \sigma \in \Lambda$.

Proof: Let us first illustrate this for $n = 2$. Write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{15}$$

Then $A = (A_1, A_2)$ where $A_1 = (a, b)$ and $A_2 = (c, d)$.
Further $A_1 = a\mathbf{e}_1^T + b\mathbf{e}_2^T$ and $A_2 = c\mathbf{e}_1^T + d\mathbf{e}_2^T$. Therefore

$$\begin{aligned} f(A) &= af(\mathbf{e}_1^T, A_2) + bf(\mathbf{e}_2^T, A_2) \\ &= a(cf(\mathbf{e}_1^T, \mathbf{e}_1^T) + df(\mathbf{e}_1^T, \mathbf{e}_2^T)) + b(cf(\mathbf{e}_2^T, \mathbf{e}_1^T) + df(\mathbf{e}_2^T, \mathbf{e}_2^T)) \\ &= ac\, f(E_{(11)}) + ad\, f(E_{(12)}) + bc\, f(E_{(21)}) + bd\, f(E_{(22)}). \end{aligned}$$

Thus the pattern in the general case is clear. By applying linearity in each slot at a time we obtain,

$$f(A) = \sum_{\sigma \in F(N,N)} a_\sigma f(E_\sigma) \tag{16}$$

Axioms for determinant

Theorem

Any multi-linear map f is completely determined by its values on $E_\sigma, \sigma \in \Lambda$.

Proof: Let us first illustrate this for $n = 2$. Write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{15}$$

Then $A = (A_1, A_2)$ where $A_1 = (a, b)$ and $A_2 = (c, d)$.
Further $A_1 = a\mathbf{e}_1^T + b\mathbf{e}_2^T$ and $A_2 = c\mathbf{e}_1^T + d\mathbf{e}_2^T$. Therefore

$$\begin{aligned} f(A) &= af(\mathbf{e}_1^T, A_2) + bf(\mathbf{e}_2^T, A_2) \\ &= a(cf(\mathbf{e}_1^T, \mathbf{e}_1^T) + df(\mathbf{e}_1^T, \mathbf{e}_2^T)) + b(cf(\mathbf{e}_2^T, \mathbf{e}_1^T) + df(\mathbf{e}_2^T, \mathbf{e}_2^T)) \\ &= ac\, f(E_{(11)}) + ad\, f(E_{(12)}) + bc\, f(E_{(21)}) + bd\, f(E_{(22)}). \end{aligned}$$

Thus the pattern in the general case is clear. By applying linearity in each slot at a time we obtain,

$$f(A) = \sum_{\sigma \in F(N,N)} a_\sigma f(E_\sigma) \tag{16}$$

Axioms for determinant

$$f(A) = \sum_{\sigma} a_{\sigma} f(E_{\sigma}) \tag{17}$$

Axioms for determinant

$$f(A) = \sum_{\sigma} a_\sigma f(E_\sigma) \tag{17}$$

Here the scalars a_σ are given by

$$a_\sigma = a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{18}$$

The verification is straight forward. ♠

Definition

A map $f : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ is called *alternating* if for every pair $(i, j), i \neq j$,

$$f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) = -f(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n). \tag{19}$$

That is, if we interchange the place of any two rows then the sign of the value of the function is changed.

Axioms for determinant

Theorem

A multi-linear map f is alternating iff it vanishes on all matrices with two equal rows.

Axioms for determinant

Theorem

A multi-linear map f is alternating iff it vanishes on all matrices with two equal rows.

Proof: Let f be multi-linear and alternating. If two rows of A are the same then by interchanging them, A does not change. On the other hand the sign of the value of f on A has to change.

Axioms for determinant

Theorem

A multi-linear map f is alternating iff it vanishes on all matrices with two equal rows.

Proof: Let f be multi-linear and alternating. If two rows of A are the same then by interchanging them, A does not change. On the other hand the sign of the value of f on A has to change. This means that $f(A) = -f(A)$ and hence $f(A) = 0$.

Axioms for determinant

Theorem

A multi-linear map f is alternating iff it vanishes on all matrices with two equal rows.

Proof: Let f be multi-linear and alternating. If two rows of A are the same then by interchanging them, A does not change. On the other hand the sign of the value of f on A has to change. This means that $f(A) = -f(A)$ and hence $f(A) = 0$.
To prove the converse, let us consider the 2×2 case.
We assume that f is multi-linear and has the property $f(A, A) = 0$ for all A. We must show that $f(B, C) = -f(C, B)$.

Axioms for determinant

Theorem

A multi-linear map f is alternating iff it vanishes on all matrices with two equal rows.

Proof: Let f be multi-linear and alternating. If two rows of A are the same then by interchanging them, A does not change. On the other hand the sign of the value of f on A has to change. This means that $f(A) = -f(A)$ and hence $f(A) = 0$.
To prove the converse, let us consider the 2×2 case.
We assume that f is multi-linear and has the property $f(A, A) = 0$ for all A. We must show that $f(B, C) = -f(C, B)$. We proceed by taking

$$\begin{aligned} 0 &= f(B + C, B + C) \\ &= = f(B, B + C) + f(C, B + C) \\ &= f(B, B) + f(B, C) + f(C, B) + f(C, C) \\ &= f(B, C) + f(C, B) \end{aligned}$$

Axioms for determinant

In the general case, given a matrix A consider another matrix B whose rows are that of A except in the i^{th} and j^{th} places we have $B_i = B_j = A_i + A_j$. Then $f(B) = 0$.

Axioms for determinant

In the general case, given a matrix A consider another matrix B whose rows are that of A except in the i^{th} and j^{th} places we have $B_i = B_j = A_i + A_j$. Then $f(B) = 0$.
On the other hand by multi-linearity, we have

$$\begin{aligned} & 0 = f(B) \\ = \; & f(A_1, \ldots, A_i, \ldots, A_i + A_j, \ldots, A_n) + f(A_1, \ldots, A_j, \ldots, A_i + A_j, \ldots, A_n) \\ = \; & f(A_1, \ldots, A_i, \ldots, A_i, \ldots, A_n) + f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) \\ & + f(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n) + f(A_1, \ldots, A_j, \ldots, A_j, \ldots, A_n) \end{aligned}$$

Axioms for determinant

In the general case, given a matrix A consider another matrix B whose rows are that of A except in the i^{th} and j^{th} places we have $B_i = B_j = A_i + A_j$. Then $f(B) = 0$.
On the other hand by multi-linearity, we have

$$\begin{aligned} & 0 = f(B) \\ = \; & f(A_1, \ldots, A_i, \ldots, A_i + A_j, \ldots, A_n) + f(A_1, \ldots, A_j, \ldots, A_i + A_j, \ldots, A_n) \\ = \; & f(A_1, \ldots, A_i, \ldots, A_i, \ldots, A_n) + f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) \\ & + f(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n) + f(A_1, \ldots, A_j, \ldots, A_j, \ldots, A_n) \end{aligned}$$

Hence $f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) = -f(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n)$. ♠

Axioms for determinant

Remark

There is more merit in this equivalent formulation of alternating property than meets the eye.

Axioms for determinant

Remark

There is more merit in this equivalent formulation of alternating property than meets the eye. It is the right one for certain other types of fields which we have not considered here.

Axioms for determinant

Remark

There is more merit in this equivalent formulation of alternating property than meets the eye. It is the right one for certain other types of fields which we have not considered here.
As a consequence of this we can relax the condition in the defeinition of alternating property a little bit.

Theorem

A multi-linear map $f : M_n(\mathbb{R}) \to \mathbb{R}$ is alternating iff f simply changes the sign whenever we interchange the position of two consecutive rows.

Axioms for determinant

Remark

There is more merit in this equivalent formulation of alternating property than meets the eye. It is the right one for certain other types of fields which we have not considered here.
As a consequence of this we can relax the condition in the defeinition of alternating property a little bit.

Theorem

A multi-linear map $f : M_n(\mathbb{R}) \to \mathbb{R}$ is alternating iff f simply changes the sign whenever we interchange the position of two consecutive rows.

Proof: Clearly if f is alternating then it does satisfy this weaker condition. We have to prove the converse. So, assuming that f is as given in the theorem, we shall prove that f vanishes on matrices with two rows equal. But then the two rows can be made consecutive, by interchanging the consecutive rows between these two rows and therefore, it suffices to show that f vanishes on matrices which have two consecutive rows equal.

Axioms for determinant

Remark

There is more merit in this equivalent formulation of alternating property than meets the eye. It is the right one for certain other types of fields which we have not considered here.
As a consequence of this we can relax the condition in the defeinition of alternating property a little bit.

Theorem

A multi-linear map $f : M_n(\mathbb{R}) \to \mathbb{R}$ is alternating iff f simply changes the sign whenever we interchange the position of two consecutive rows.

Proof: Clearly if f is alternating then it does satisfy this weaker condition. We have to prove the converse. So, assuming that f is as given in the theorem, we shall prove that f vanishes on matrices with two rows equal. But then the two rows can be made consecutive, by interchanging the consecutive rows between these two rows and therefore, it suffices to show that f vanishes on matrices which have two consecutive rows equal. Now the rest of proof is exactly same as that of the previous theorem.

Axioms for determinant

Theorem

Let f be a multi-linear alternating map on $M_n(\mathbb{K})$. Consider the elementary matrix $E = I + \alpha E_{ij}$. Then for all $A \in M_n$,

$$f(EA) = \begin{cases} f(A) & \text{if } i \neq j \\ (1+\alpha)f(A) & \text{if } i = j \end{cases}$$

Axioms for determinant

Theorem

Let f be a multi-linear alternating map on $M_n(\mathbb{K})$. Consider the elementary matrix $E = I + \alpha E_{ij}$. Then for all $A \in M_n$,

$$f(EA) = \begin{cases} f(A) & \text{if } i \neq j \\ (1+\alpha)f(A) & \text{if } i = j \end{cases}$$

Proof: Let $A = (A_1, \ldots, A_n)$. For $i \neq j$, observe that

$$EA = (A_1, \ldots, A_i + \alpha A_j, \ldots, A_j, \ldots, A_n).$$

Axioms for determinant

Theorem

Let f be a multi-linear alternating map on $M_n(\mathbb{K})$. Consider the elementary matrix $E = I + \alpha E_{ij}$. Then for all $A \in M_n$,

$$f(EA) = \begin{cases} f(A) & \text{if } i \neq j \\ (1+\alpha)f(A) & \text{if } i = j \end{cases}$$

Proof: Let $A = (A_1, \ldots, A_n)$. For $i \neq j$, observe that

$$EA = (A_1, \ldots, A_i + \alpha A_j, \ldots, A_j, \ldots, A_n).$$

Therefore

$$\begin{aligned} f(EA) &= f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) + \alpha f(A_1, \ldots, A_j, \ldots, A_j, \ldots, A_n) \\ &= f(A). \end{aligned}$$

Axioms for determinant

Theorem

Let f be a multi-linear alternating map on $M_n(\mathbb{K})$. Consider the elementary matrix $E = I + \alpha E_{ij}$. Then for all $A \in M_n$,

$$f(EA) = \begin{cases} f(A) & \text{if } i \neq j \\ (1+\alpha)f(A) & \text{if } i = j \end{cases}$$

Proof: Let $A = (A_1, \ldots, A_n)$. For $i \neq j$, observe that

$$EA = (A_1, \ldots, A_i + \alpha A_j, \ldots, A_j, \ldots, A_n).$$

Therefore

$$\begin{aligned} f(EA) &= f(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) + \alpha f(A_1, \ldots, A_j, \ldots, A_j, \ldots, A_n) \\ &= f(A). \end{aligned}$$

For $i = j$, we have $EA = (A_1, \ldots, A_{i-1}, (1+\alpha)A_i, A_{i+1}, \ldots, A_n)$. Therefore, $f(EA) = (1+\alpha)f(A)$, by linearity in the i^{th} slot. ♠

Axioms for determinant

Remark

In the case of linear maps we were free to choose the value of the function on $\mathbf{e}_i's$.

Axioms for determinant

Remark

In the case of linear maps we were free to choose the value of the function on $\mathbf{e}_i's$. Same is true for a multi-linear map, viz., we are free to choose its value on E_{ij} and then extend it by the formula (17).

Axioms for determinant

Remark

In the case of linear maps we were free to choose the value of the function on $\mathbf{e}_i's$. Same is true for a multi-linear map, viz., we are free to choose its value on E_{ij} and then extend it by the formula (17). However, in addition, if f is alternating also, then we are forced to choose its value on E_σ to be zero for all those σ which are not one-to-one.

Axioms for determinant

Remark

In the case of linear maps we were free to choose the value of the function on $\mathbf{e}_i's$. Same is true for a multi-linear map, viz., we are free to choose its value on E_{ij} and then extend it by the formula (17). However, in addition, if f is alternating also, then we are forced to choose its value on E_σ to be zero for all those σ which are not one-to-one.

Definition

Let $\Sigma(n)$ denote those elements $\sigma \in F(N, N)$ which are one-to-one. These are called permutations on n letters. If a permutation changes its value only on two indices then it is called a transposition.

Axioms for determinant

Given $\sigma \in \Sigma(n)$, let us observe what GEM does on E_σ.

Axioms for determinant

Given $\sigma \in \Sigma(n)$, let us observe what GEM does on E_σ.
It is clear that GEM simply consists of a sequence of change of rows and we obtain $G(E_\sigma) = J(E_\sigma) = I_n$. Observe that a change of two rows of E_σ is effected by composing σ with a corresponding transposition.

Axioms for determinant

Given $\sigma \in \Sigma(n)$, let us observe what GEM does on E_σ.
It is clear that GEM simply consists of a sequence of change of rows and we obtain $G(E_\sigma) = J(E_\sigma) = I_n$. Observe that a change of two rows of E_σ is effected by composing σ with a corresponding transposition.
Thus GEM yields a sequence of transpositions $\tau_1, \ldots, \tau_k$ such that

$$E(\tau_k \circ \cdots \circ \tau_1 \circ \sigma) = I_n.$$

Axioms for determinant

Given $\sigma \in \Sigma(n)$, let us observe what GEM does on E_σ.
It is clear that GEM simply consists of a sequence of change of rows and we obtain $G(E_\sigma) = J(E_\sigma) = I_n$. Observe that a change of two rows of E_σ is effected by composing σ with a corresponding transposition.
Thus GEM yields a sequence of transpositions $\tau_1, \ldots, \tau_k$ such that

$$E(\tau_k \circ \cdots \circ \tau_1 \circ \sigma) = I_n.$$

This means that $\tau_k \circ \cdots \tau_1 \circ \sigma = Id$. Hence $\sigma = \tau_1 \circ \cdots \circ \tau_k$. Thus we have proved:

Lemma

Every permutation is a (finite) composite of transpositions.

Axioms for determinant

Given $\sigma \in \Sigma(n)$, let us observe what GEM does on E_σ.
It is clear that GEM simply consists of a sequence of change of rows and we obtain $G(E_\sigma) = J(E_\sigma) = I_n$. Observe that a change of two rows of E_σ is effected by composing σ with a corresponding transposition.
Thus GEM yields a sequence of transpositions $\tau_1, \ldots, \tau_k$ such that

$$E(\tau_k \circ \cdots \circ \tau_1 \circ \sigma) = I_n.$$

This means that $\tau_k \circ \cdots \tau_1 \circ \sigma = Id$. Hence $\sigma = \tau_1 \circ \cdots \circ \tau_k$. Thus we have proved:

Lemma

Every permutation is a (finite) composite of transpositions.

Definition

Let $e(\sigma)$ denote the number of row changes in the GEM applied to E_σ to bring it to $G(E_\sigma) = I_n$. Let $sgn(\sigma) = (-1)^{e(\sigma)}$. We call this the signature of σ.

Axioms for determinant

Applying repeatedly the alternating property, it follows that

$$f(E_\sigma) = sgn(\sigma) f(I_n). \tag{20}$$

Axioms for determinant

Applying repeatedly the alternating property, it follows that

$$f(E_\sigma) = sgn(\sigma) f(I_n). \tag{20}$$

Combining

$$(17), \qquad f(A) = \sum_\sigma a_\sigma f(E_\sigma)$$

and

$$(18), \qquad a_\sigma = a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

with (20), we get

$$f(A) = \left(\sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_\sigma \right) f(I_n). \tag{21}$$

Existence

Theorem

For every $t \in \mathbb{R}$, there exists a unique $f_t : M_n(\mathbb{R}) \to \mathbb{R}$ which is row-wise multi-linear, alternating and such that $f_t(I_n) = t$. Moreover, for any such f_t, we have $f_t(E_\sigma) = sgn(\sigma)t$.

Existence

Theorem

For every $t \in \mathbb{R}$, there exists a unique $f_t : M_n(\mathbb{R}) \to \mathbb{R}$ which is row-wise multi-linear, alternating and such that $f_t(I_n) = t$. Moreover, for any such f_t, we have $f_t(E_\sigma) = sgn(\sigma)t$.

Note that so far what we have already proved the uniqueness of such a function and the formula $f_t(E_\sigma) = sgn(\sigma)t$.

Existence

Theorem

For every $t \in \mathbb{R}$, there exists a unique $f_t : M_n(\mathbb{R}) \to \mathbb{R}$ which is row-wise multi-linear, alternating and such that $f_t(I_n) = t$. Moreover, for any such f_t, we have $f_t(E_\sigma) = sgn(\sigma)t$.

Note that so far what we have already proved the uniqueness of such a function and the formula $f_t(E_\sigma) = sgn(\sigma)t$.

Existence

To show that such a function exists, we just define f_t by the formula

$$f_t(A) = \left(\sum_{\sigma \in \Sigma} sgn(\sigma) a_\sigma \right) t. \tag{22}$$

To verify that f_t is multi-linear is straight forward.
Assuming that formula (22) has alternating property on the set $\{E_\sigma, \sigma \in \Sigma\}$, it follows easily that the same holds for all matrices. Therefore, it is enough to prove:

Existence

To show that such a function exists, we just define f_t by the formula

$$f_t(A) = \left(\sum_{\sigma \in \Sigma} sgn(\sigma) a_\sigma \right) t. \tag{22}$$

To verify that f_t is multi-linear is straight forward.
Assuming that formula (22) has alternating property on the set $\{E_\sigma, \sigma \in \Sigma\}$, it follows easily that the same holds for all matrices. Therefore, it is enough to prove:

Lemma

Let $\sigma, \tau \in \Sigma(n)$, where τ is a transposition. Then

$$sgn(\sigma) = -sgn(\tau \circ \sigma). \tag{23}$$

Existence

Proof: We prove it by induction on n. If $n = 2$ this is obvious. (For $n = 3$ this can be verified straight forward way, by considering various cases. Try this as an exercise before reading the rest of the proof which will help you to understand the proof better. However, this is not a logical necessity.)
Inductively, let the statement (23) hold for $n - 1$. Let $A = E_\sigma,\ B = E_{\tau\circ\sigma}$.

Existence

Proof: We prove it by induction on n. If $n = 2$ this is obvious. (For $n = 3$ this can be verified straight forward way, by considering various cases. Try this as an exercise before reading the rest of the proof which will help you to understand the proof better. However, this is not a logical necessity.)
Inductively, let the statement (23) hold for $n-1$. Let $A = E_\sigma,\ B = E_{\tau\circ\sigma}$.
Suppose τ interchanges i and j, $i \neq j$. Notice that B is obtained from A by interchanging the i^{th} and j^{th} row.
We look at the first step of GEM performed on both A and B cut down the first row and first column from both of them compare the results on either side, use induction to obtain the equality (23) for $n-1$ and finally relate it to (23) for n.

Existence

Suppose GEM applied to A interchanges 1^{st} row with k^{th} row. (We do not exclude the case when $k = 1$; it merely means that the first step of GEM does not do anything on A.)
The rest proof is best understood by glancing at the table below which is indeed a schematic representation of the proof.

Existence

Case 1: Let $\{1, k\} = \{i, j\}$. Here first step of GEM applied to A is the same as τ applied to A and hence both produce B. Clearly, the first step of GEM applied to B does nothing and all the remaining steps will coincide for both. Hence (23) gets verified here.

Case 2: Suppose the set $\{1, k, i, j\}$ has three elements. Observe that during the first step of GEM on A or B none of the other rows are affected. (Thus by simply ignoring these rows and the corresponding columns, this reduces to the case when $n = 3$ for which you have already verified (23). However, we shall write down a detailed proof.)

Existence

We consider first the subcase, where $k = 1$. Then the first step of GEM does not produce any change in A or B and we simply pass to the next stage. The equation (23) for n follows from the same for $n - 1$ by induction.

Existence

Next, suppose k is equal to one of the indices $\{i, j\}$ say for definiteness, $k = i$. This means that the non zero entry in the first column of A is in i^{th} place and hence the same for B is j^{th} place. Therefore the first step of GEM for A swaps rows $1 \leftrightarrow i$, where as the same for B swaps rows $1 \leftrightarrow j$. Therefore after cutting down the first row and first column, the two matrices are again related by the transposition $\tau' = (i-1, j-1)$. Again the equation (23) for n is obtained by multiplying by -1 the same equation for $n-1$.

Existence

Case 3: Suppose $\{1, k, i, j\}$ has four elements. Under this condition, it is clear that the first step of GEM on both A and B is to interchange the first and k^{th} rows. After this the two $(n-1) \times (n-1)$ matrices A', B' are again related by a transposition. Therefore, by induction, (23) is true for A', B' and hence after mutliplying it by -1 the same holds for A, B. ♠

Existence of determinant

Case	Stage	A		B	
(a)	Initial stage	Y ———	(1)	Y ———	(1)
		.....		.....	
		R ———	(i)	P ———	(i)
		.....		.....	
		P ———	(j)	R ———	(j)
		.....		.....	
	After step 1	Y ———	(1=k)	Y ———	(1=k)
		.....		.....	
		R ———	(i)	P ———	(i)
		.....		.....	
		P ———	(j)	R ———	(j)
		.....		.....	
(b)	Initial stage	Y ———	(1)	Y ———	(1)
		.....		.....	
		R ———	(i=k)	P ———	(i)
		.....		.....	
		P ———	(j)	R ———	(j)
		.....		.....	
	After Step 1	R ———	(1)	R ———	(1)
		.....		.....	
		Y ———	(i)	P ———	(i)
		.....		.....	
		P ———	(j)	Y ———	(j)
		.....		.....	

Axioms for determinant

Case 3		
Stage	A	B
Initial stage	Y ——— (1) . . . . . . R ——— (i) . . . . . . G ——— (k) . . . . . . P ——— (j)	Y ——— (1) . . . . . . P ——— (i) . . . . . . G ——— (k) . . . . . . R ——— (j) . . . . . .
After Step 1	G ——— (1) . . . . . . R ——— (i) . . . . . . Y ——— (k) . . . . . . P ——— (j)	G ——— (1) . . . . . . P ——— (i) . . . . . . Y ——— (k) . . . . . . R ——— (j) . . . . . .

Existence of determinant

Definition

A multi-linear alternating map $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called the **determinant function** if $f(I_n) = 1$.

Existence of determinant

Definition

A multi-linear alternating map $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called the **determinant function** if $f(I_n) = 1$.

Theorem

There exist a unique determinant function for each $n \geq 1$. This is given by the Expansion formula

$$\det A = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_{\sigma} = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{24}$$

Remark

The condition in the above definition is called normalization. Thus it follows that viz., a multi-linear, alternating, normalised map is unique.

Existence of determinant

Definition

A multi-linear alternating map $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called the **determinant function** if $f(I_n) = 1$.

Theorem

There exist a unique determinant function for each $n \geq 1$. This is given by the Expansion formula

$$\det A = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_\sigma = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{24}$$

Remark

The condition in the above definition is called normalization. Thus it follows that viz., a multi-linear, alternating, normalised map is unique.
It is denoted by $\det$. Often one writes $\det A$ or $|A|$ to denote the determinant of a square matrix A.

Existence of determinant

Definition

A multi-linear alternating map $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called the **determinant function** if $f(I_n) = 1$.

Theorem

There exist a unique determinant function for each $n \geq 1$. This is given by the Expansion formula

$$\det A = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_\sigma = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{24}$$

Remark

The condition in the above definition is called normalization. Thus it follows that viz., a multi-linear, alternating, normalised map is unique.
It is denoted by $\det$. Often one writes $\det A$ or $|A|$ to denote the determinant of a square matrix A.
Clearly it coincides with the concept that we already have in case $n = 1, 2$.

Existence of determinant

Definition

A multi-linear alternating map $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$ is called the **determinant function** if $f(I_n) = 1$.

Theorem

There exist a unique determinant function for each $n \geq 1$. This is given by the Expansion formula

$$\det A = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_\sigma = \sum_{\sigma \in \Sigma(n)} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{24}$$

Remark

The condition in the above definition is called normalization. Thus it follows that viz., a multi-linear, alternating, normalised map is unique.
It is denoted by $\det$. Often one writes $\det A$ or $|A|$ to denote the determinant of a square matrix A.
Clearly it coincides with the concept that we already have in case $n = 1, 2$.
We shall now establish other properties of $\det$ which we are familiar with.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.

Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.

Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$. By theorem 8, it follows that $\det(A) = (-1)^{e(A)} \det(G(A))$.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.

Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$. By theorem 8, it follows that $\det(A) = (-1)^{e(A)} \det(G(A))$. Now look at the diagonal entries of $G(A)$. While arriving at $J(A)$ from $G(A)$, we first divide each row by the corresponding diagonal entry c_i and then sweep each column.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.

First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.

Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$. By theorem 8, it follows that $\det(A) = (-1)^{e(A)} \det(G(A))$. Now look at the diagonal entries of $G(A)$. While arriving at $J(A)$ from $G(A)$, we first divide each row by the corresponding diagonal entry c_i and then sweep each column. Since the sweeping of columns does not change the determinant, it follows that

$$\det(G(A)) = c_1 c_2 \cdots c_n. \tag{25}$$

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.
First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.
Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$. By theorem 8, it follows that $\det(A) = (-1)^{e(A)} \det(G(A))$. Now look at the diagonal entries of $G(A)$. While arriving at $J(A)$ from $G(A)$, we first divide each row by the corresponding diagonal entry c_i and then sweep each column. Since the sweeping of columns does not change the determinant, it follows that

$$\det(G(A)) = c_1 c_2 \cdots c_n. \tag{25}$$

Combining these two we have

$$\det A = (-1)^{e(A)} c_1 \cdots c_n. \tag{26}$$

where c_i are the diagonal entries of $G(A)$.

Computing the determinant

Let us now see how GEM can be applied to compute the determinant of a $n \times n$ matrix.
First consider the case when the row rank of A is equal to n. Then we know that $J(A) = I_n$.
Suppose the number of change of rows performed to arrive at $G(A)$ is $e(A)$. By theorem 8, it follows that $\det(A) = (-1)^{e(A)} \det(G(A))$. Now look at the diagonal entries of $G(A)$. While arriving at $J(A)$ from $G(A)$, we first divide each row by the corresponding diagonal entry c_i and then sweep each column. Since the sweeping of columns does not change the determinant, it follows that

$$\det(G(A)) = c_1 c_2 \cdots c_n. \tag{25}$$

Combining these two we have

$$\det A = (-1)^{e(A)} c_1 \cdots c_n. \tag{26}$$

where c_i are the diagonal entries of $G(A)$. Now if the row rank of A is $< n$ then clearly some row of $G(A)$ is zero and hence one of the $c_i = 0$. Also, $\det G(A) = 0$ and since $\det A = \pm \det G(A)$, it follows that $\det A = 0$. So, the formula (26) is still valid.

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence* $\det A$ *is the product of the diagonal entries.*

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$.*

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$. Thus M can written in the form*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \tag{27}$$

where A, B, D are respectively of type $k \times k, k \times (n-k), (n-k) \times (n-k)$.

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$. Thus M can written in the form*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \tag{27}$$

where A, B, D are respectively of type $k \times k, k \times (n-k), (n-k) \times (n-k)$. Such a matrix is called a block matrix.

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$. Thus M can written in the form*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \tag{27}$$

where A, B, D are respectively of type $k \times k, k \times (n-k), (n-k) \times (n-k)$. Such a matrix is called a block matrix.
Let us compute the determinant of such a matrix. Let us apply GEM to M. Observe that unless A has become upper triangular, the process will never involve the rows numbered $> k$.

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$. Thus M can written in the form*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \tag{27}$$

where A, B, D are respectively of type $k \times k, k \times (n-k), (n-k) \times (n-k)$. Such a matrix is called a block matrix.
Let us compute the determinant of such a matrix. Let us apply GEM to M. Observe that unless A has become upper triangular, the process will never involve the rows numbered $> k$. Suppose the number of steps in $G(A)$ is less than k. Then there will be a row of zeros in $G(A)$ and hence $G(M)$ will have a zero on the diagonal.

Computing determinant

Example

- *(i) In the special case when A is upper triangular, we have $G(A) = A$. Hence $\det A$ is the product of the diagonal entries.*
- *(ii) Consider an $n \times n$ matrix M for which $a_{ij} = 0, k+1 \leq i \leq n, 1 \leq j \leq k$. Thus M can written in the form*

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} \tag{27}$$

where A, B, D are respectively of type $k \times k, k \times (n-k), (n-k) \times (n-k)$. Such a matrix is called a block matrix.
Let us compute the determinant of such a matrix. Let us apply GEM to M. Observe that unless A has become upper triangular, the process will never involve the rows numbered $> k$. Suppose the number of steps in $G(A)$ is less than k. Then there will be a row of zeros in $G(A)$ and hence $G(M)$ will have a zero on the diagonal. It follows that in this case $\det A = 0 = \det M$.

Computing determinant

Example

(ii) continued:
Now suppose the number of steps in $G(A)$ is k.

Computing determinant

Example

(ii) continued:
Now suppose the number of steps in $G(A)$ is k. Then GEM applied to M will first make A into an upper triangular matrix and the rest of the operation on M will not involve the first k rows at all.

Computing determinant

Example

(ii) continued:
Now suppose the number of steps in $G(A)$ is k. Then GEM applied to M will first make A into an upper triangular matrix and the rest of the operation on M will not involve the first k rows at all. Therefore, these operations can be thought of as two separate sets of row operations on A and D which correspond to operations of GEM on them.

Computing determinant

Example

(ii) continued:
Now suppose the number of steps in $G(A)$ is k. Then GEM applied to M will first make A into an upper triangular matrix and the rest of the operation on M will not involve the first k rows at all. Therefore, these operations can be thought of as two separate sets of row operations on A and D which correspond to operations of GEM on them. Thus we conclude that the diagonal entries of $G(M)$ are those of $G(A)$ and $G(D)$ whereas the number of row swaps for M will be the sum of these numbers for A and D separately. Thus we have proved:

Theorem

For a block matrix

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

we have $\det M = (\det A)(\det D)$.

Computing Determinants

Formula (23) has some other interesting applications.

Corollary

Let a permutation σ be the composite of s transpositions. Then $sgn(\sigma) = (-1)^s$.

Computing Determinants

Formula (23) has some other interesting applications.

Corollary

Let a permutation σ be the composite of s transpositions. Then $sgn(\sigma) = (-1)^s$.

Proof: Follows by induction on s by successive application of (23). ♠

Computing Determinants

Formula (23) has some other interesting applications.

Corollary

Let a permutation σ be the composite of s transpositions. Then $sgn(\sigma) = (-1)^s$.

Proof: Follows by induction on s by successive application of (23). ♠

Corollary

For any two permutations, $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$. In particular $sgn(\sigma) = sgn(\sigma^{-1})$.

Computing Determinants

Formula (23) has some other interesting applications.

Corollary

Let a permutation σ be the composite of s transpositions. Then $sgn(\sigma) = (-1)^s$.

Proof: Follows by induction on s by successive application of (23). ♠

Corollary

For any two permutations, $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$. In particular $sgn(\sigma) = sgn(\sigma^{-1})$.

Proof: First part is clear. Now observe that for any transposition τ, we have $\tau^{-1} = \tau$. Therefore, if $\sigma = \tau_1 \circ \cdots \circ \tau_k$ then $\sigma^{-1} = \tau_k \circ \cdots \circ \tau_1$. Now apply the previous lemma. ♠

The Transpose

The entire discussion so far was with respect to the rows.

The Transpose

The entire discussion so far was with respect to the rows. We could have carried the whole thing with respect to rows.

The Transpose

The entire discussion so far was with respect to the rows. We could have carried the whole thing with respect to rows. We then get another normalised alternating multi-linear map with respect to the columns.

The Transpose

The entire discussion so far was with respect to the rows. We could have carried the whole thing with respect to rows. We then get another normalised alternating multi-linear map with respect to the columns. Let us call it δ.

The Transpose

The entire discussion so far was with respect to the rows. We could have carried the whole thing with respect to rows. We then get another normalised alternating multi-linear map with respect to the columns. Let us call it δ. The following theorem tell you that $\delta = \det$.

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Put $\tau = \sigma^{-1}$. Then as σ runs over all possible permutations of $\{1, \ldots, n\}$ so does τ.

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Put $\tau = \sigma^{-1}$. Then as σ runs over all possible permutations of $\{1, \ldots, n\}$ so does τ. On the other hand for each i we have $i = \tau(\sigma(i))$ and for each fixed σ as i runs over $\{1, \ldots, n\}$ so does $\sigma(i)$.

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Put $\tau = \sigma^{-1}$. Then as σ runs over all possible permutations of $\{1, \ldots, n\}$ so does τ. On the other hand for each i we have $i = \tau(\sigma(i))$ and for each fixed σ as i runs over $\{1, \ldots, n\}$ so does $\sigma(i)$. Therefore

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\tau(1)} \cdots a_{n\tau(n)}.$$

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Put $\tau = \sigma^{-1}$. Then as σ runs over all possible permutations of $\{1, \ldots, n\}$ so does τ. On the other hand for each i we have $i = \tau(\sigma(i))$ and for each fixed σ as i runs over $\{1, \ldots, n\}$ so does $\sigma(i)$. Therefore

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\tau(1)} \cdots a_{n\tau(n)}.$$

Since we have seen that $sgn(\sigma) = sgn(\sigma^{-1})$, the above summation is the same as the rhs of (24). ♠

Det A^T =Det A

Theorem

For any square matrix A we have $\det A^T = \det A$.

Proof: Putting $B = A^T$ and observing that $b_{ij} = a_{ji}$ from (24) it follows that

$$\det A^T = \sum_{\sigma} sgn(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Put $\tau = \sigma^{-1}$. Then as σ runs over all possible permutations of $\{1, \ldots, n\}$ so does τ. On the other hand for each i we have $i = \tau(\sigma(i))$ and for each fixed σ as i runs over $\{1, \ldots, n\}$ so does $\sigma(i)$. Therefore

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\tau(1)} \cdots a_{n\tau(n)}.$$

Since we have seen that $sgn(\sigma) = sgn(\sigma^{-1})$, the above summation is the same as the rhs of (24). ♠

Remark

This result is extremely useful. For, all the properties of the determinant function that we formulated or verified for rows will now be available for columns also.

Product formula

(**Skip the proof**)

Theorem

Product formula for determinant *For any two $n \times n$ matrices A and B, we have* $\det(AB) = (\det A)(\det B)$.

Product formula

(**Skip the proof**)

Theorem

Product formula for determinant *For any two $n \times n$ matrices A and B, we have* $\det(AB) = (\det A)(\det B)$.

Proof: Fix B and consider the map $f(A) = \det(AB)$. Then $f(I_n) = \det B$. If we show that f is multi-linear and alternating then by the uniqueness of such maps the theorem would follow. Observe that the i^{th} row of AB is nothing but A_iB. Hence if two rows of A are equal then so are the corresponding two rows of AB. Hence $f(A) = \det(AB) = 0$. This proves alternating property. Now to check multi-linearity:

$$\begin{aligned} & f(A_1, \ldots, aA_i + a'A_i', \ldots, A_n) \\ = \; & \det(A_1B, \ldots, (aA_i + a'A_i')B, \ldots, A_nB) \\ = \; & a\det(A_1B, \ldots, A_iB, \ldots, A_nB) + a'\det(A_1B, \ldots, A_i'B, \ldots, A_nB) \\ = \; & a\,f(A_1, \ldots, A_i, \ldots, A_n) + a'f(A_1, \ldots, A_i', \ldots, A_n) \end{aligned}$$

This proves the multi-linearity of f. ♠

Laplace determinant

(**Skip the proofs**)

Definition

Let A_{ij} denote submatrix (minor) of A obtained by cutting down the i^{th} row and j^{th} column. We define the $(i,j)^{th}$ cofactor of A by $cof_{i,j}(A) = (-1)^{i+j} \det A_{ij}$.

For example

$$A_{32} = \begin{bmatrix} * & \times & * & * \\ * & \times & * & * \\ \times & \times & \times & \times \\ * & \times & * & * \end{bmatrix}$$

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A_{ij}^{\#}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$.

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A^{\#}_{ij}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$. In order to compute the determinant of $A^{\#}_{ij}$, we can sweep the j^{th} column and assume that $a_{kj} = 0$ for $k \neq i$.

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A^{\#}_{ij}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$. In order to compute the determinant of $A^{\#}_{ij}$, we can sweep the j^{th} column and assume that $a_{kj} = 0$ for $k \neq i$. Then by a sequence of interchanging the i^{th} row with the row just above it we bring this row to the 1^{st} place.

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A^{\#}_{ij}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$. In order to compute the determinant of $A^{\#}_{ij}$, we can sweep the j^{th} column and assume that $a_{kj} = 0$ for $k \neq i$. Then by a sequence of interchanging the i^{th} row with the row just above it we bring this row to the 1^{st} place. Now the j^{th} column is equal to $(1, 0, \ldots, 0)^T$. Mean while the determinant has changed its sign by $(-1)^{i-1}$.

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A^{\#}_{ij}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$. In order to compute the determinant of $A^{\#}_{ij}$, we can sweep the j^{th} column and assume that $a_{kj} = 0$ for $k \neq i$. Then by a sequence of interchanging the i^{th} row with the row just above it we bring this row to the 1^{st} place. Now the j^{th} column is equal to $(1, 0, \ldots, 0)^T$. Mean while the determinant has changed its sign by $(-1)^{i-1}$.
  Likewise we now swap this column with the columns left to it and bring it to the first place.

Expansion by minors

Remark

- (i) Consider a matrix $A = ((a_{ij}))$ in which $a_{1j} = a_{i1} = 0$ for all $i, j > 1$ and $a_{11} = 1$. Then check that $\det A = cof_{11}(A) = \det A_{11}$. (This is an easy application of GEM).
- (ii) Let now for any square matrix A and (i, j), let us introduce a temporary notation $A^{\#}_{ij}$ to denote the matrix obtained by replacing the i^{th} row by $\mathbf{e}_j$. In order to compute the determinant of $A^{\#}_{ij}$, we can sweep the j^{th} column and assume that $a_{kj} = 0$ for $k \neq i$. Then by a sequence of interchanging the i^{th} row with the row just above it we bring this row to the 1^{st} place. Now the j^{th} column is equal to $(1, 0, \ldots, 0)^T$. Mean while the determinant has changed its sign by $(-1)^{i-1}$.
  Likewise we now swap this column with the columns left to it and bring it to the first place. The determinant now changed its sign by $(-1)^{i+j-2} = (-1)^{i+j}$. Thus if we denote this new matrix by B then $\det B = (-1)^{i+j} \det A^{\#}_{ij}$. Also $B_{11} = A_{i,j}$ Hence by (i) above we have $\det B = \det B_{11} = \det A_{ij}$.

Expansion by minors

Thus we have proved:

Lemma

For any square matrix A we have,

$$\det A^{\#}_{ij} = cof_{ij}(A). \tag{28}$$

Expansion by minors

Thus we have proved:

Lemma

For any square matrix A we have,

$$\det A^{\#}_{ij} = cof_{ij}(A). \tag{28}$$

The following theorem gives an inductive formula to obtain the determinant of a matrix in terms of its entries.

Expansion by minors

Theorem

Laplace Expansion Formula For Determinant: *For any $n \times n$ matrix A we have*

$$\det A = \sum_{j=1}^{n} a_{ij} cof_{ij} A, \ \forall \ 1 \leq i \leq n. \tag{29}$$

Expansion by minors

Theorem

Laplace Expansion Formula For Determinant: *For any $n \times n$ matrix A we have*

$$\det A = \sum_{j=1}^{n} a_{ij} cof_{ij} A, \ \forall \ 1 \leq i \leq n. \tag{29}$$

Proof: Using linearity in the i^{th} row we expand the determinant function. Thus

Expansion by minors

Theorem

Laplace Expansion Formula For Determinant: *For any $n \times n$ matrix A we have*

$$\det A = \sum_{j=1}^{n} a_{ij}\, cof_{ij} A, \ \forall\ 1 \leq i \leq n. \tag{29}$$

Proof: Using linearity in the i^{th} row we expand the determinant function. Thus

$$\begin{aligned} \det A &= \det(A_1, \ldots, A_i, \ldots A_n) \\ &= \det(A_1, \ldots, \textstyle\sum_j a_{ij}\mathbf{e}_j, \ldots, A_n) \\ &= \textstyle\sum_j a_{ij} \det(A_1, \ldots, \mathbf{e}_j, \ldots, A_n) \\ &= \textstyle\sum_j a_{ij} \det A_{ij}^{\#} = \sum_j a_{ij}\, cof_{ij}(A). \end{aligned}$$

the last equality is taken from (28).

Expansion by minors

Theorem

Laplace Expansion Formula For Determinant: *For any $n \times n$ matrix A we have*

$$\det A = \sum_{j=1}^{n} a_{ij} cof_{ij} A, \ \forall \ 1 \leq i \leq n. \tag{29}$$

Proof: Using linearity in the i^{th} row we expand the determinant function. Thus

$$\begin{aligned} \det A &= \det(A_1, \ldots, A_i, \ldots A_n) \\ &= \det(A_1, \ldots, \textstyle\sum_j a_{ij}\mathbf{e}_j, \ldots, A_n) \\ &= \textstyle\sum_j a_{ij} \det(A_1, \ldots, \mathbf{e}_j, \ldots, A_n) \\ &= \textstyle\sum_j a_{ij} \det A_{ij}^{\#} = \sum_j a_{ij}\, cof_{ij}(A). \end{aligned}$$

the last equality is taken from (28).

Remark

Laplace formula can be used to give an inductive proof of the existence of the determinant. We shall not discuss this any further.

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row.

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.

On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.

On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

$$0 = \det(M) = \sum_{k=1}^{n} m_{ik}\, cof_{jk}(M).$$

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.

On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

$$0 = \det(M) = \sum_{k=1}^{n} m_{ik}\, cof_{jk}(M).$$

This is true for every $i \neq j$.

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.

On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

$$0 = \det(M) = \sum_{k=1}^{n} m_{ik}\, cof_{jk}(M).$$

This is true for every $i \neq j$.

Combining it with (29) we obtain

$$\sum_{k=1}^{n} a_{ik}\, cof_{jk}(A) = \begin{cases} \det A, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.
On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

$$0 = \det(M) = \sum_{k=1}^{n} m_{ik}\, cof_{jk}(M).$$

This is true for every $i \neq j$.
Combining it with (29) we obtain

$$\sum_{k=1}^{n} a_{ik}\, cof_{jk}(A) = \begin{cases} \det A, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Similarly by considering the transpose we have

Cramer's Rule

Given any square matrix M, and $i \neq j$, let $M = M_{ij}$ denote the matrix obtained by replacing the j^{th} row of M by its i^{th} row. Since two rows of M_{ij} are equal we know that $\det M = \det M_{ij} = 0$.
On the other hand (29) applied to M_{ij} and taking the expansion along the j^{th} row we get

$$0 = \det(M) = \sum_{k=1}^{n} m_{ik}\, cof_{jk}(M).$$

This is true for every $i \neq j$.
Combining it with (29) we obtain

$$\sum_{k=1}^{n} a_{ik}\, cof_{jk}(A) = \begin{cases} \det A, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Similarly by considering the transpose we have

$$\sum_{k=1}^{n} a_{ki}\, cof_{kj}(A) = \begin{cases} \det A, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \tag{31}$$

Cramer's rule: The Adjoint of a matrix

Given a $n \times n$ matrix A we define the adjoint A to be the $n \times n$ matrix $adj\ A = (m_{ij})$ where

$$m_{ij} = cof_{ji}(A).$$

Cramer's rule: The Adjoint of a matrix

Given a $n \times n$ matrix A we define the adjoint A to be the $n \times n$ matrix $adj\ A = (m_{ij})$ where

$$m_{ij} = cof_{ji}(A).$$

Then

Theorem

$(adj\ A)A = A(adj\ A) = diag(\det A, \ldots, \det A)$

Check that this is just a reformulation of (30).

Cramer's rule: The Adjoint of a matrix

Given a $n \times n$ matrix A we define the adjoint A to be the $n \times n$ matrix $adj\ A = (m_{ij})$ where

$$m_{ij} = cof_{ji}(A).$$

Then

Theorem

$(adj\ A)A = A(adj\ A) = diag(\det A, \ldots, \det A)$

Check that this is just a reformulation of (30).
In particular if $\det A \neq 0$, then it follows that A is invertible with the inverse given by

Theorem

$A^{-1} = (\det A)^{-1}(adj\ A)$.

Cramer's rule

As a simple minded application we obtain:

Cramer's rule

As a simple minded application we obtain:

Theorem

Cramer's Rule: *Given a system of linear equation* $A\mathbf{x} = \mathbf{b}$ *with an invertible coefficient matrix, let* M_j *be the matrix obtained from* A *by replacing the* j^{th} *column by the column* $\mathbf{b}$. *Then*

$$x_i = \frac{\det M_i}{\det A} \tag{32}$$

Cramer's rule

As a simple minded application we obtain:

Theorem

Cramer's Rule: *Given a system of linear equation* $A\mathbf{x} = \mathbf{b}$ *with an invertible coefficient matrix, let* M_j *be the matrix obtained from* A *by replacing the* j^{th} *column by the column* $\mathbf{b}$. *Then*

$$x_i = \frac{\det M_i}{\det A} \tag{32}$$

Proof: We have $\mathbf{x} = A^{-1}\mathbf{b} = (\det A)^{-1}(adj\ A)\mathbf{b}$. Equating the i^{th} entry on either side, we get

Cramer's rule

As a simple minded application we obtain:

Theorem

Cramer's Rule: *Given a system of linear equation* $A\mathbf{x} = \mathbf{b}$ *with an invertible coefficient matrix, let* M_j *be the matrix obtained from* A *by replacing the* j^{th} *column by the column* $\mathbf{b}$. *Then*

$$x_i = \frac{\det M_i}{\det A} \tag{32}$$

Proof: We have $\mathbf{x} = A^{-1}\mathbf{b} = (\det A)^{-1}(adj\, A)\mathbf{b}$. Equating the i^{th} entry on either side, we get

$$x_i == (\det A)^{-1}(\sum_{j=1}^{n} m_{ij} b_j) = (\det A)^{-1} \sum_{j=1}^{n} cof_{ji}(A) b_j = (\det A)^{-1}(\det M_i)$$

the last equality is obtained by taking the expansion of $\det M_i$ along the i^{th} column. ♠

Cramer's rule

As a simple minded application we obtain:

Theorem

Cramer's Rule: *Given a system of linear equation* $A\mathbf{x} = \mathbf{b}$ *with an invertible coefficient matrix, let* M_j *be the matrix obtained from* A *by replacing the* j^{th} *column by the column* $\mathbf{b}$. *Then*

$$x_i = \frac{\det M_i}{\det A} \tag{32}$$

Proof: We have $\mathbf{x} = A^{-1}\mathbf{b} = (\det A)^{-1}(adj\ A)\mathbf{b}$. Equating the i^{th} entry on either side, we get

$$x_i == (\det A)^{-1}(\sum_{j=1}^{n} m_{ij}b_j) = (\det A)^{-1}\sum_{j=1}^{n} cof_{ji}(A)b_j = (\det A)^{-1}(\det M_i)$$

the last equality is obtained by taking the expansion of $\det M_i$ along the i^{th} column. ♠

Remark

Cramer's rule is not good for computational purposes as it involves computation of determinants which required huge number of operations, as compared to GEM and certain modifications of GEM.