Linear Equations

26/11/2021

Linear equation

A linear equation in the variables $x_1, x_2, \ldots \ldots., x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \ a_3x_3 + \cdots \ldots. + a_nx_n = b$$

where b and the coefficients of $x_1, x_2, \ldots \ldots., x_n$ are real or complex numbers

Eg. $7x_1 + 5x_2 - 12x_3 = 4.5$

System of linear equations

A system of linear equations (or a linear system) is a collection of one or more linear equations involving the same variables $x_1, x_2, \ldots \ldots, x_n$

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

Homogeneous linear equations

A system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \; \dots \dots . + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \; \dots \dots . + a_{2n}x_n &= b_2 \\ a_{m1}x_1 + a_{m2}x_2 + \; \dots \dots . + a_{mn}x_n &= b_m \end{aligned}$$

is called homogeneous if $b_1 = b_2 = \cdots = b_m = 0$ and

non-homogeneous, otherwise.

Eg.

$$\begin{aligned} x_1 - x_2 + \quad x_3 &= 8 \\ x_1 \qquad - \quad 4x_3 &= 7 \end{aligned}$$

Solution of the system

- A solution of the system is a list $\{s_1, s_2, \ldots \ldots, s_n\}$ of numbers that makes each equation a true statement when the values $s_1, s_2, \ldots \ldots, s_n$ are substituted for $x_1, x_2, \ldots \ldots, x_n$ respectively.

Eg.,

$$\begin{aligned} x_1 - x_2 + \quad x_3 &= 8 \\ x_1 \qquad - \quad 4x_3 &= 7 \end{aligned}$$

- {11, 4, 1} is a solution of the above equations because, when these values are substituted for $x_1, x_2, \ldots \ldots, x_n$, respectively, the equations simplify to 8 = 8 and 7 = 7

Solution of the system

- The set of all possible solutions is called the solution set of the linear system.
- Two linear systems are called equivalent if they have the same solution set. That is, each solution of the first system is a solution of the second system, and each solution of the second system is a solution of the first.

- A system of linear equations has

1. no solution
2. exactly ones solution
3. infinitely many solutions.

Consistent Solution

- A system of linear equations is said to be consistent if it has either one solution or infinitely many solutions;
- a system is inconsistent if it has no solution.

Augmented Matrix

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{2n} \\ a_{m1} & a_{m2} & a_{m3} & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ x_n \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_m \end{bmatrix}$

Then, the system of equations

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

can be rewritten as Ax = b, where A is called the coefficient matrix and the matrix [A b] is called the augmented matrix.

Eg. Given the system

- $x_1 - 2x_2 + x_3 = 0$
- $2x_2 - 8x_3 = 8$
- $5x_1 - x_3 = 10$

The matrix $\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 8 \\ 5 & 0 & -1 \end{bmatrix}$ is called the coefficient matrix

$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -1 & 10 \end{bmatrix}$ is called the augmented matrix of the system.

Elementary Row Operations on a matrix

1. (Replacement) Replace one row by the sum of itself and a multiple of another row
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply all entries in a row by a non-zero constant.

Elementary Row Operations on a matrix

- Row operations can be applied to any matrix, not merely to one that arises as the augmented matrix of a linear system. Two matrices are called row equivalent if there is a sequence of elementary row operations that transforms one matrix into the other.
- It is important to note that row operations are *reversible*. If two rows are interchanged, they can be returned to their original positions by another interchange. If a row is scaled by a nonzero constant c, then multiplying the new row by 1/c produces the original row.
- A *nonzero* row or column in a matrix means a row or column that contains at least one nonzero entry; a leading entry of a row refers to the leftmost nonzero entry (in a nonzero row).

Row echelon form

A rectangular matrix is in echelon form (or row echelon form) if it has the following three properties:

- All nonzero rows are above any rows of all zeros.
- Each leading entry of a row is in a column to the right of the leading entry of the row above it.
- All entries in a column below a leading entry are zeros.

The following matrices are in echelon form.

eg. $\begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -8 & 8 \\ 0 & 0 & 0 & 10 \end{bmatrix} \quad and \quad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & -5 & 10 \end{bmatrix}$

Reduced row echelon form

If a matrix in echelon form satisfies the following additional conditions, then it is in reduced echelon form (or reduced row echelon form):

- The leading entry in each nonzero row is1
- Each leading 1is the only nonzero entry in its column.

- The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0's below *and above* each leading 1.

Eg. $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad and \quad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -8 & 8 \\ 0 & 0 & 1 & 10 \end{bmatrix}$

Reduced row echelon form

- Any nonzero matrix may be row reduced (that is, transformed by elementary row operations) into more than one matrix in echelon form, using different sequences of row operations. However, the reduced echelon form one obtains from a matrix is unique.

Pivot Positions

A pivot position in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A.

A pivot column is a column of A that contains a pivot position.

Eg, $A = \begin{bmatrix} 0 & 3 & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}$

Here, the entries a_{12} and a_{a23} are pivots and columns 2 and 3 are pivotal columns.

Gauss Elimination Process

We now start with solving a systems of linear equations. The idea is to manipulate the rows of the augmented matrix in place of the linear equations themselves. Since, multiplying a matrix on the left corresponds to row operations, we left multiply by certain matrices to the augmented matrix so that the final matrix is in row echelon form . The process of obtaining the row echelon form of a matrix is called the Gauss Elimination method.

The general Gaussian elimination procedure is applied to the linear systems:

$$R_1 : a_{11}x_1 + a_{12}x_2 + \ \ldots\ldots. + a_{1n}x_n = b_1$$
$$R_2 : a_{21}x_1 + a_{22}x_2 + \ \ldots\ldots. + a_{2n}x_n = b_2$$
$$R_n : a_{n1}x_1 + a_{n2}x_2 + \ \ldots\ldots. + a_{nn}x_n = b_n$$

Form the augmented matrix from the system of equations

The unknowns are eliminated to obtain an upper-triangular matrix.

Gauss Elimination Process

$$R_1 : a_{11}x_1 + a_{12}x_2 + \ \dots\dots + a_{1n}x_n = b_1$$

$$R_2 : a_{21}x_1 + a_{22}x_2 + \ \dots\dots + a_{2n}x_n = b_2$$

To eliminate x_1 from R2, we multiply R1 by (-a_{21}/a_{11}) and obtain

$$-a_{21}x_1 - a_{12}\left(\frac{a_{21}}{a_{11}}\right)x_2 - \ \dots\dots - a_{1n}\left(\frac{a_{21}}{a_{11}}\right)x_n = -b_1\left(\frac{a_{21}}{a_{11}}\right)$$

Adding the above equation to R_2 we obtain

$$\left(a_{22} - a_{12}\frac{a_{21}}{a_{11}}\right)x_2 - \left(a_{23} - a_{13}\frac{a_{21}}{a_{11}}\right)x_3 \dots\dots - \left(a_{2n} - a_{1n}\frac{a_{21}}{a_{11}}\right)x_n$$
$$= b_2 - b_1\left(\frac{a_{21}}{a_{11}}\right)$$

Gauss Elimination Process

R_2 can be rewritten as

$$R_2 : a'_{22}x_2 + \ a'_{23}x_3 + \cdots \ldots . + a'_{2n}x_n = b'_2$$

Where $a'_{22} = \ \left(a_{22} - \ a_{12}\frac{a_{21}}{a_{11}}\right)$ and so on.

Gauss Elimination Process

In a similar fashion, we can eliminate x_1 from the remaining equations and after eliminating x_1 from the last row Rn, we obtain the system

$$R_1 : a_{11}x_1 + a_{12}x_2 + \ a_{13}x_3 + \cdots \ldots. + a_{1n}x_n \ = b_1$$

$$R_2 : \qquad\qquad a'_{22}x_2 + \ a'_{23}x_3 + \cdots \ldots. + a'_{2n}x_n \ = b'_2$$

$$R_n : \qquad\qquad a'_{n2}x_2 + \ a'_{n3}x_3 + \cdots \ldots. + a'_{nn}x_n \ = b'_n$$

Gauss Elimination Process

In the process of obtaining the above system, we have multiplied the first row by (-a_{21}/a_{11}), i.e. we have divided it bya_{11} which is therefore assumed to be nonzero. For this reason, the first row R1 in is called the pivot equation, and a_{11} is called the pivot or pivotal element. The method obviously fails if a_{11}= 0.

-

Gauss Elimination Process

Similarly, we eliminate the variables will be obtain the upper-triangular matrix in the form:

$$R_1 : a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots \ldots . + a_{1n}x_n = b_1$$

$$R_2 : \quad a'_{22}x_2 + a'_{23}x_3 + \cdots \ldots . + a'_{2n}x_n = b'_2$$

$$R_3 : \quad a''_{33}x_3 + \cdots \ldots . + a''_{3n}x_n = b''_3$$

$$R_n : \quad a_{nn}^{(n-1)}x_n = b_n^{(n-1)}$$

where $a_{nn}^{(n-1)}$ indicates the element a_{nn} has changed (n-1) times.

Gauss Elimination Process

From R_n : $a_{nn}^{(n-1)} x_n = b_n^{(n-1)}$

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$$

This is then substituted in the $R_{(n-1)}$ to obtain x_{n-1} and the process is repeated to compute the other unknowns. We have therefore first computed $x_n \; then \; x_{n-1}, \ldots \ldots \; x_2, x_1$ in that order. Due to this reason, the process is called back substitution.

Linear Equations

29/11/2021

Gauss Elimination Process

$$\begin{aligned} x_2 + x_3 &= 2 \\ 2x_1 + 3x_3 &= 5 \\ x_1 + x_2 + x_3 &= 3. \end{aligned}$$

The augmented matrix can be written as

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

Interchange R_2 and R_1 to get

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

Gauss Elimination Process

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

Replace R_3 by $R_3 - {}^1/_2 R_1$ to get

$$\begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 1-({}^1/_2)2 & 1-({}^1/_2)0 & 1-({}^1/_2)3 & 3-({}^1/_2)5 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & -{}^1/_2 & {}^1/_2 \end{bmatrix}$$

Gauss Elimination Process

Replace R_3 by $R_3 - R_2$ to get

$$\begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 0-0 & 1-1 & -({}^{1}/_{2})-1 & ({}^{1}/_{2})-2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -{}^{3}/_{2} & -{}^{3}/_{2} \end{bmatrix}$$

The matrix is in row echelon form. Using the last row we get $x_3 = 1$

Second row of the matrix gives us $x_2 + x_3 = 2$ So, $x_2 = 1$

First row gives us $2x_1 + 3x_3 = 5$ So $x_1 = 1$

Gauss Elimination Process

Eg2.

$$x_1 + 3x_2 + 5x_3 = 14$$
$$2x_1 - x_2 - 3x_3 = 3$$
$$4x_1 + 5x_2 - x_3 = 7$$

Gauss Elimination Process

The augmented matrix can be written as

$$\begin{bmatrix} 1 & 3 & 5 & 14 \\ 2 & -1 & -3 & 3 \\ 4 & 5 & -1 & 7 \end{bmatrix}$$

Replace R_2 by $R_2 - 2R_1$ and R_3 by $R_3 - 4R_1$ to get

$$\begin{bmatrix} 1 & 3 & 5 & 14 \\ 2-2 & -1-2(3) & -3-2(5) & 3-2(14) \\ 4-4 & 5-4(3) & -1-4(5) & 7-4(14) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 3 & 5 & 14 \\ 0 & -7 & -13 & -25 \\ 0 & -7 & -21 & -49 \end{bmatrix}$$

Gauss Elimination Process

Since all the elements in R_2 and R_3 are negative, we multiply throughout by -1

Replace R_2 by (-1)R_2 and R_3 by (-1)R_3 to get

$$\begin{bmatrix} 1 & 3 & 5 & 14 \\ 0 & -7 & -13 & -25 \\ 0 & -7 & -21 & -49 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 3 & 5 & 14 \\ 0 & 7 & 13 & 25 \\ 0 & 7 & 21 & 49 \end{bmatrix}$$

Gauss Elimination Process

Replace R_3 by R_3 - R_2 to get

$$= \begin{bmatrix} 1 & 3 & 5 & 14 \\ 0 & 7 & 13 & 25 \\ 0 & 0 & 8 & 24 \end{bmatrix}$$

Now back substitution gives us

$$\begin{aligned} x_1 + 3x_2 \;+\; 5x_3 &= 14 \\ 7x_2 \;+\; 13x_3 &= 25 \\ 8x_3 &= 24 \end{aligned}$$

$x_1 = 5, \qquad x_2 = -2, \qquad x_3 = 3$

Gauss Elimination Process

Eg3.

$$x_1 + x_2 + x_3 = 3$$
$$x_1 + 2x_2 + 2x_3 = 5$$
$$3x_1 + 4x_2 + 4x_3 = 11$$

Gauss Elimination Process

The augmented matrix can be written as

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 11 \end{bmatrix}$$

Replace R_2 by $R_2 - R_1$ and R_3 by $R_3 - 3R_1$ to get

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1-1 & 2-1 & 2-1 & 5-3 \\ 3-3 & 4-3 & 4-3 & 11-3(3) \end{bmatrix}$$

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

Gauss Elimination Process

Replace R_3 by R_3 - R_2 to get

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

Now back substitution gives us

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

Since there are 3 unknowns but only 2 constrants

The system has infinite number of solutions

Pivoting

- We now come to the important case of the pivot being zero or very close to zero. If the pivot is zero, the entire process fails and if it is close to zero, round-off errors may occur. These problems can be avoided by adopting a procedure called pivoting. If a_{11} is either zero or very small compared to the other coefficients of the equation, then we find the largest available coefficient in the columns below the pivot equation and then interchange the two rows. In this way, we obtain a new pivot equation with a nonzero pivot. Such a process is called partial pivoting, since in this case we search only the columns below for the largest element. If, on the other hand, we search both columns and rows for the largest element, the procedure is called complete pivoting. It is obvious that complete pivoting involves more complexity in computations since interchange of columns means change of order of unknowns which invariably requires more programming effort. In comparison, partial pivoting, i.e. row interchanges, is easily adopted in programming. Due to this reason, complete pivoting is rarely used.

Pivoting

Eg.

$0.0003120x_1 + 0.006032x_2 = 0.003328$

$0.500000x_1 + 0.89420\, x_2 = 0.9471$

The exact solution is x1 = 1 and x2 = 0.5

We first solve the system with pivoting. We write the given system as

$$\begin{bmatrix} 0.500000 & 0.89420 & 0.9471 \\ 0.0003120 & 0.006032 & 0.003328 \end{bmatrix}$$

Pivoting

Replace R_2 by $R_2 - \left(\frac{0.0003120}{0.0000050}\right) R_1$ to get

$$\begin{bmatrix} 0.500000 & 0.89420 & 0.9471 \\ 0 & 0.005474 & 0.002737 \end{bmatrix}$$

Back substitution gives us x1 = 1 and x2 = 0.5

Without pivoting, Gauss elimination gives

$$\begin{bmatrix} 0.0003120 & 0.006032 & 0.003328 \\ 0 & -8.77725 & -5.3300 \end{bmatrix}$$

Back substitution gives us x1 = -1.0803 and x2 = 0.6076

Gauss Seidel

- We shall now describe the iterative or indirect methods, which start from an approximation to the true solution and, if convergent, derive a sequence of closer approximations- the cycle of computations being repeated till the required accuracy is obtained. This means that in a direct method the amount of computation is fixed, while in an iterative method the amount of computation depends on the accuracy required.
- In general, one should prefer a direct method for the solution of a linear system, but in the case of matrices with a large number of zero elements, it will be advantageous to use iterative methods which preserve these elements.

Gauss Seidel

$$R_1 : a_{11}x_1 + a_{12}x_2 + \; \dots\dots . + a_{1n}x_n = b_1$$
$$R_2 : a_{21}x_1 + a_{22}x_2 + \; \dots\dots . + a_{2n}x_n = b_2$$
$$R_n : a_{n1}x_1 + a_{n2}x_2 + \; \dots\dots . + a_{nn}x_n = b_n$$

in which the diagonal elements aii do not vanish. If this is not the case, then the equations should be rearranged so that this condition is satisfied.

Gauss Seidel

- Suppose $x_1^{(1)}, x_2^{(1)}, \dots \dots\ x_n^{(1)}$ are any first approximations to the unknowns $x_1, x_2, \dots \dots ., x_n$
- We rewrite the equations as

$$x_1^{(1)} = \left(\frac{b_1}{a_{11}}\right) - \left(\frac{a_{12}}{a_{11}}\right) x_2 - \left(\frac{a_{13}}{a_{11}}\right) x_3 \dots \dots . - \left(\frac{a_{1n}}{a_{11}}\right) x_n$$

$$x_2^{(1)} = \left(\frac{b_2}{a_{22}}\right) - \left(\frac{a_{21}}{a_{22}}\right) x_1 - \left(\frac{a_{23}}{a_{22}}\right) x_3 \dots \dots . - \left(\frac{a_{2n}}{a_{22}}\right) x_n$$

$$x_3^{(1)} = \left(\frac{b_3}{a_{33}}\right) - \left(\frac{a_{31}}{a_{33}}\right) x_1 - \left(\frac{a_{32}}{a_{33}}\right) x_2 \dots \dots . - \left(\frac{a_{2n}}{a_{33}}\right) x_n$$

$$x_n^{(1)} = \left(\frac{b_n}{a_{nn}}\right) - \left(\frac{a_{n1}}{a_{nn}}\right) x_1 - \left(\frac{a_{n2}}{a_{nn}}\right) x_2 \dots \dots . - \left(\frac{a_{n,n-1}}{a_{nn}}\right) x_{n-1}$$

Gauss Seidel

- We get the second approximations as
- $x_1^{(2)} = \left(\frac{b_1}{a_{11}}\right) - \left(\frac{a_{12}}{a_{11}}\right) x_2^{(1)} - \left(\frac{a_{13}}{a_{11}}\right) x_3^{(1)} \dots \dots . - \left(\frac{a_{1n}}{a_{11}}\right) x_n^{(1)}$
- Since, we already have $x_1^{(2)}$, we can write second estimate of x_2 as
- $x_2^{(2)} = \left(\frac{b_2}{a_{22}}\right) - \left(\frac{a_{21}}{a_{22}}\right) x_1^{(2)} - \left(\frac{a_{23}}{a_{22}}\right) x_3^{(1)} \dots \dots . - \left(\frac{a_{2n}}{a_{22}}\right) x_n^{(1)}$
- Since, we already have$x_1^{(2)}, x_2^{(2)}$, we can write second estimate of x_3 as
- $x_3^{(2)} = \left(\frac{b_3}{a_{33}}\right) - \left(\frac{a_{31}}{a_{33}}\right) x_1^{(2)} - \left(\frac{a_{32}}{a_{33}}\right) x_2^{(2)} \dots \dots . - \left(\frac{a_{2n}}{a_{33}}\right) x_n^{(1)}$
- Since, we already have$x_1^{(2)}, x_2^{(2)}, x_{n-1}^{(2)}$ we can write second estimate of $x_n$as
- $x_n^{(2)} = \left(\frac{b_n}{a_{nn}}\right) - \left(\frac{a_{n1}}{a_{nn}}\right) x_1^{(2)} - \left(\frac{a_{n2}}{a_{nn}}\right) x_2^{(2)} \dots \dots . - \left(\frac{a_{n,n-1}}{a_{nn}}\right) x_{n-1}^{(n-1)}$

Gauss Seidel

- In this manner, we complete the first stage of iteration and the entire process is repeated till the values of $x_1, x_2, \dots \dots .., x_n$ are obtained to the accuracy required.
- It is clear, therefore, that this method uses an improved component as soon as it is available and it is called the method of successive displacements, or the Gauss-Seidel method.

Gauss Seidel

- Diagonal Dominance: The Gauss-Seidel methods converge, for any choice of the first approximation $x_j^{(1)}$) (j = 1, 2, ..., n), if every equation of the system satisfies the condition that the sum of the absolute values of the coefficients $\left(\frac{a_{ij}}{a_{ii}}\right)$ is almost equal to, or in at least one equation less than unity, i.e. provided that

$$\sum_{j=1, j\neq 1}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| \leq 1, (i = 1,2, \dots \dots, n)$$

where the < sign should be valid in the case of 'at least' one equation.