Linear Algebra & Differential Equations

Elif Uskuplu

If you see any mistake, please email me (euskuplu@usc.edu).

2 Week 2

2.1 The general idea behind this week's topics

We will build our first tool to solve a system of linear equations. Before, let us discuss the idea behind the technique. Consider the following system:

$$\begin{aligned} 2x + 3y &= 2 \\ -x + 2y &= 6 \end{aligned}$$

You might be familiar with this from high-school algebra or pre-calculus course that we can make the coefficients simpler to solve these equations. Now, multiply the second equality by 2 and add the first one. Then we will have

$$\begin{aligned} 2x + 3y &= 2 \\ 7y &= 14 \end{aligned}$$

Now, simplify the second equality by dividing with 7, simplify the first equality by dividing with 2 and get

$$\begin{aligned} x + \frac{3}{2}y &= 1 \\ y &= 2 \end{aligned}$$

Then we find $y = 2$, and via back substitution, namely, we replace y with 2 in the first equality, we have $x + 3 = 1$ which means $x = -2$.

Now, consider the augmented matrices of first and last systems:

$$\left[\begin{array}{cc|c} 2 & 3 & 2 \\ -1 & 2 & 6 \end{array}\right] \quad \rightarrow \quad \left[\begin{array}{cc|c} 1 & \frac{3}{2} & 1 \\ 0 & 1 & 2 \end{array}\right]$$

These both represent the same system, but the latter produces solutions easily. In essence, we develop this method to the system with arbitrary size $m \times n$. This row operations on the equations will be row operations on the augmented matrix. The resulting matrix is called **row-echelon matrix**, and its more developed version will be **reduced row-echelon matrix**. Considering also the back substitution technique, we have our first tool to solve linear equations system. This is called **Gauss elimination**, and its develop version is **Gauss-Jordan elimination**.

2.2 Row-echelon matrices & Row operations

Definition 2.1. *An $m \times n$ matrix is called a* ***row-echelon*** *matrix if it satisfies the following three conditions:*

1. *If there are any rows consisting entirely of zeros, they are grouped together at the bottom of the matrix.*
2. *The first nonzero element in any nonzero row is a 1 (called a leading 1).*
3. *The leading 1 of any row below the first row is to the right of the leading 1 of the row above it.*

Example. While $\begin{bmatrix} 1 & -8 & -3 & 7 \\ 0 & 1 & 5 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -3 & -6 & 5 & 7 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ are row-echelon matrices, the matrices $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ are not row-echelon matrices.

Now, for any $m \times n$ matrix A, we have some methods to make A into a row-echelon matrix. There are three operations for this task.

Definition 2.2. ***Elementary Row Operations:***

1. *P_{ij}:* ***P****ermute the ith and jth rows*
2. *$M_i(k)$:* ***M****ultiply every element of the ith row by a nonzero scalar k.*
3. *$A_{ij}(k)$:* ***A****dd to the element of the jth row the scalar k times the elements of the ith row.*

Example. Consider the system

$$\begin{aligned} x + 2y - z &= 4 \\ 2x - y + 3z &= -6 \\ -x + 3y - 2z &= 7 \end{aligned}$$

Then, its augmented matrix is

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 2 & -1 & 3 & | & -6 \\ -1 & 3 & -2 & | & 7 \end{bmatrix}.$$

We want to reduce it into a row-echelon matrix. First, apply $A_{12}(-2)$, namely, add -2 times of the first row to the second row, then we have

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & -5 & 5 & | & -14 \\ -1 & 3 & -2 & | & 7 \end{bmatrix}.$$

Similarly, apply $A_{13}(1)$, and get

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & -5 & 5 & | & -14 \\ 0 & 5 & -3 & | & 11 \end{bmatrix}.$$

Now, apply $M_2(\frac{-1}{5})$ and $M_3(\frac{1}{5})$ and get

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 1 & -\frac{3}{5} & | & \frac{11}{5} \end{bmatrix}.$$

Now, apply $A_{23}(-1)$ and get

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 0 & \frac{2}{5} & | & -\frac{3}{5} \end{bmatrix}.$$

Finally, apply $M_3(\frac{5}{2})$ and get

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}.$$

Now, according to Definition 2.1, the last matrix is a row-echelon matrix.

Definition 2.3. *Let A be an $m \times n$ matrix. Any matrix obtained from A by a finite sequence of elementary row operations is said to be* ***row-equivalent*** *to A.*

Example. From the previous example, we have $\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 2 & -1 & 3 & | & -6 \\ -1 & 3 & -2 & | & 7 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}$ are row-equivalent matrices.

Remark. Using row operations, we can make any matrix into a row-echelon matrix. So we have theorem about this.

Theorem 2.1. *Every matrix is row-equivalent to a row-echelon matrix.*

Let's write the complete algorithm for reducing an $m \times n$ matrix A to row-echelon form. Indeed, the algorithm itself serves as a proof of Theorem 2.1.

1. Start with an $m \times n$ matrix A. If $A = 0$, go to step 7.
2. Determine the leftmost nonzero column (**the pivot column**) and the topmost position in this column (**the pivot position**).
3. Use elementary row operations to put a 1 in the pivot position.
4. Use elementary row operations to put zeros below the pivot position.

5. If there are no more nonzero rows below the pivot position, go to step 7. Otherwise, proceed to step 6.

6. Apply steps 2–5 to the submatrix consisting of the rows below the pivot position.

7. The matrix is now in row-echelon form.

2.3 Reduced row-echelon matrix

Definition 2.4. *An $m \times n$ matrix is called a* ***reduced row-echelon matrix*** *if it satisfies the following conditions:*

1. *It is a row-echelon matrix.*

2. *Any column that contains a leading 1 has zeros everywhere else.*

Example. The matrix $\begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is a reduced row-echelon matrix.

While you may end up with different row-echelon matrices after row operations, **the reduced row-echelon form is unique.** In other words, if you continue to apply more row operations to obtain reduced row-echelon matrix, the resulting matrix will be the unique.

Example. Continue to our previous example in row-echelon section: we have

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}.$$

In order to have reduced form, first we apply $A_{21}(-2)$, namely, we add -2 multiple of the second row to the first row. It gives us

$$\begin{bmatrix} 1 & 0 & 1 & | & -\frac{8}{5} \\ 0 & 1 & -1 & | & \frac{14}{5} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}.$$

Then apply $A_{31}(-1)$ and $A_{32}(1)$ and get

$$\begin{bmatrix} 1 & 0 & 0 & | & -\frac{1}{10} \\ 0 & 1 & 0 & | & \frac{13}{10} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}.$$

And this is a reduced row-echelon matrix.

2.4 Rank of a matrix

Definition 2.5. *The rank of a matrix A, denoted as rank(A), is defined as the number of nonzero rows in any row-echelon form of A.*

This concept is vital in linear algebra and in determining the solution properties of linear systems. Despite the non-uniqueness of a matrix's row-echelon form, all row-echelon forms of A that are row-equivalent have the same number of nonzero rows. The proof is omitted now, and it will be clear after we cover *vector spaces*.

Example. Let's calculate rank(A) where $A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 1 & -2 & 1 & 3 \\ 1 & -5 & 0 & 5 \end{bmatrix}$. After applying $M_1(\frac{1}{2})$, $A_{12}(-1)$, $A_{13}(-1)$, $A_{23}(-3)$, and $M_2(-\frac{2}{3})$, respectively, we get A is row-equivalent to $\begin{bmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Therefore, rank($A$)=2.

Remark. If A is an $m \times n$ matrix, then rank(A) $\leq m$ and rank(A) $\leq n$. This is because the number of nonzero rows in a row-echelon form of A is equal to the number of pivots in a row-echelon form of A, which cannot exceed the number of rows or columns of A, since there can be at most one pivot per row and per column.

2.5 Gauss(-Jordan) elimination

The process of reducing the augmented matrix to row-echelon form and then using back substitution to solve the equivalent system is called **Gaussian elimination**. The particular case of Gaussian elimination that arises when the augmented matrix is reduced to reduced row-echelon form is called **Gauss-Jordan elimination**.

Example. We continue with our main example in the note. Recall that we started with the system

$$\begin{aligned} x + 2y - z &= 4 \\ 2x - y + 3z &= -6 \\ -x + 3y - 2z &= 7. \end{aligned}$$

Its augmented matrix is

$$\begin{bmatrix} 1 & 2 & -1 & | & 4 \\ 2 & -1 & 3 & | & -6 \\ -1 & 3 & -2 & | & 7 \end{bmatrix}.$$

We apply Gauss-Jordan elimination: First, make this a reduced row-echelon matrix. We already have this:

$$\begin{bmatrix} 1 & 0 & 0 & | & -\frac{1}{10} \\ 0 & 1 & 0 & | & \frac{13}{10} \\ 0 & 0 & 1 & | & -\frac{3}{2} \end{bmatrix}.$$

But if we make it a linear equation system again, we have:

$$x = -\frac{1}{10}$$
$$y = \frac{13}{10}$$
$$z = -\frac{3}{2}.$$

This is exactly a solution for the system. The solution can be read off directly as $(-\frac{1}{10}, \frac{13}{10}, -\frac{3}{2})$.

The following theorem gives a general characterization of a system of equations.

Theorem 2.2. *Consider the $m \times n$ linear system $A\mathbf{x} = \mathbf{b}$. Let r denote the rank of A and $r^{\#}$ denote the rank of $A^{\#}$. Then*

1. *If $r < r^{\#}$, the system is inconsistent. There is no solution.*
2. *If $r = r^{\#}$, the system is consistent and*
   (a) *There exists a unique solution if and only if $r^{\#} = n$.*
   (b) *There exists an infinite number of solutions if and only if $r^{\#} < n$.*

2.5.1 Free & bound variables

In the study of linear systems, variables are classified as either **bound** or **free**. Bound variables are those which can be expressed in terms of other variables, while free variables are those that can take any value. This concept is fundamental in understanding the solutions of linear systems.

Example. Consider the following system of linear equations:

$$\begin{aligned} x + 2y - z &= 1 \\ 3x + 8y - 2z &= 12 \\ 4x + 10y - 3z &= 15 \end{aligned}$$

The augmented matrix for this system is:

$$\begin{bmatrix} 1 & 2 & -1 & | & 1 \\ 3 & 8 & -2 & | & 12 \\ 4 & 10 & -3 & | & 15 \end{bmatrix}$$

Applying Gaussian elimination with row operations $A_{12}(-3), A_{13}(-4), A_{23}(-1), M_2(\frac{1}{2})$ respectively, the matrix becomes:

$$\begin{bmatrix} 1 & 2 & -1 & | & 1 \\ 0 & 1 & \frac{1}{2} & | & \frac{9}{2} \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

The system of equations then becomes:

$$x + 2y - z = 1$$
$$y + \frac{1}{2}z = \frac{9}{2}.$$

Now, there are 2 equations but 3 variables. In this case, one of the variables is free, we can parameterize it freely. Let pick z as free. While z is a free variable, x, y are bound variables. Let assume $z = t$, then back substitution gives us:

$$y = \frac{9-t}{2} \qquad x = 2t - 8.$$

Therefore, for any t, we have $(2t - 8, \frac{9-t}{2}, t)$ is a solution. It means that there are infinitely many solutions.

Remark. You might notice that we cannot take x as a free variable in this example because doing so does not give a way to express y and z separately. We could choose y, but there is a convention about picking free variables:

*Choose as free variables those variables that **do not** correspond to a leading 1 in a row-echelon form of $A^{\#}$.*

Example. Determine all values of the constant k for which the following system has **(a)** no solution, **(b)** infinitely many solutions, **(c)** a unique solution.

$$\begin{aligned} x + 2y - z &= 3, \\ 2x + 5y + z &= 7, \\ x + y - k^2 z &= -k. \end{aligned}$$

Its augmented matrix is $\begin{bmatrix} 1 & 2 & -1 & | & 3 \\ 2 & 5 & 1 & | & 7 \\ 1 & 1 & -k^2 & | & -k \end{bmatrix}$. After applying $A_{12}(-2)$, $A_{13}(-1)$, and $A_{23}(1)$, respectively, we get

$$\begin{bmatrix} 1 & 2 & -1 & | & 3 \\ 0 & 1 & 4 & | & 1 \\ 0 & 0 & 5-k^2 & | & 2-k \end{bmatrix}.$$

Now, if k is $\sqrt{5}$ or $-\sqrt{5}$, then we would have $rank(A) = 2$ but $rank(A^{\#}) = 3$. In such a case, there cannot be a solution by Theorem 2.2. Otherwise, observe that $rank(A) = rank(A^{\#}) = 3$. By Theorem 2.2, there is a unique solution.

Special case: homogeneous systems

Recall that the $m \times n$ systems of the form $A\mathbf{x} = \mathbf{0}$ are called homogeneous. Since $\mathbf{x} = \mathbf{0}$ is always a solution for such systems (this is called **trivial solution**), we can conclude that a homogeneous system is always consistent. If $m < n$, then the system has infinitely many solutions. This is because of that $r = r^{\#}$ in a homogeneous system and we have $r^{\#} = r \leq m < n$, so the result follows from Theorem 2.2.