

Revisiting Systems of Linear Equations

Biswajit Sahoo

Almost every reader would have seen systems of linear equations from their high school days. Whether they liked it or not is a separate story. But, in all likelihood, they would have solved these equations by gradually removing variables one by one by substitution. In this way, three equations with three variables (or unknowns) gets transformed to two equations in two variables and one further step of reduction gives us an equation with only one variable which is readily solvable. Then the final solution is obtained by back substituting the obtained value of the variable into remaining equations. This method, in mathematical jargon, is called Gaussian elimination and back substitution.

It turns out (surprisingly) that linear systems form the basis of many interesting engineering applications. Ultimately the problem boils down to solution (or approximate solution) of a system of linear equations. So a thorough understanding of linear systems is essential to appreciate the applications. In this post we will outline all possible cases of finding solutions to linear systems and briefly outline two most important applications.

We will use matrix notation to represent the equations succinctly. It also gives us better insight into their solution. Using matrix notation the system can be represented as

$$\mathbf{Ax} = \mathbf{b}$$

Where \mathbf{A} is the matrix of coefficients of size $(m \times n)$, \mathbf{x} is a vector of variables of size $(n \times 1)$, and \mathbf{b} is a vector of size $(m \times 1)$ representing constant right hand sides. Note that \mathbf{b} can be a vector of all zeros, i.e., $\mathbf{b} = \mathbf{0}$ or it can be any arbitrary vector with some nonzero values, i.e., $\mathbf{b} \neq \mathbf{0}$. The solution(s) of linear systems depend to a large extent on what the right hand side is as we will see shortly.

Apart from notation, we need two other concepts from matrix theory. One is of rank and other is the range space (or column space) of a matrix. Rank ($Rank(\mathbf{A})$) of a matrix (say, \mathbf{A}) is defined as number of independent rows or columns of a matrix. It is a well known result in matrix theory that row rank (number of independent rows) is equal to column rank (number of independent columns) and $Rank(\mathbf{A}) \leq \min(m, n)$.

Range space ($\mathcal{R}(\mathbf{A})$) (in short, Range) of a matrix is the vector space of all possible linear combinations of columns of the matrix. As we take all possible linear combination of columns, it is also known as column space. Readers who are slightly more familiar with linear algebra may know that Range is the span of columns of \mathbf{A} . Zero vector ($\mathbf{0}$) is **always** in the range of \mathbf{A} because if we take linear combination of columns of \mathbf{A} with all coefficients as 0's, we get zero vector. Hence $\mathbf{b} = \mathbf{0} \in \mathcal{R}(\mathbf{A})$ is always true.

Let's now discuss different cases separately and their solutions. We will assume that our system of equations has real entries.

- Case - I: ($m = n$)
 - $Rank(\mathbf{A}) = m$
 - * $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Unique solution (for any \mathbf{b}). For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 8 \\ 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

This system has unique solution.

- * $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: Impossible (This case will never happen because $\text{Rank}(\mathbf{A}) = m$)
- $\text{Rank}(\mathbf{A}) < m$
 - * $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Infinitely many solutions. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 8 \end{bmatrix}$$

This system has infinitely many solutions.

- * $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: No solution. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}$$

This system has no solution.

- Case - II: ($m > n$)

- $\text{Rank}(\mathbf{A}) = n$
 - * $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Unique solution. For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 7 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 11 \end{bmatrix}$$

This system has unique solution.

- * $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: No solution. For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 7 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 11 \end{bmatrix}$$

This system has no solution. But this case is immensely useful from application point of view. Sometimes it is not desirable to obtain the exact solution. Rather an approximate solution suffices for all practical purposes. Finding an approximate solution to an overdetermined system leads to the famous Least Squares problem.

- $\text{Rank}(\mathbf{A}) < n$
 - * $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Infinitely many solutions. For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$$

It has infinitely many solutions.

- * $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: No solution. For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 8 \end{bmatrix}$$

This system has no solution.

- Case - III: ($m < n$)

- $\text{Rank}(\mathbf{A}) = m$:

* $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Infinitely many solutions. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

This system has infinitely many solutions. This case is also used in many applications. As there are infinitely many solutions, a natural choice is to choose the best solution. The qualifier ‘best’ determines what application we have in our mind. If we seek minimum (l_2) norm, we get the so called minimum energy solution, a concept used in signal processing. Yet another concern is to seek for the sparsest solution (a solution with only a few nonzero entries and all other entries being zero). This idea is used in Compressed Sensing, an active research area with many interesting applications.

* $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: Impossible. This case will never happen since $\text{Rank}(\mathbf{A}) = m$.

– $\text{Rank}(\mathbf{A}) < m$

* $\mathbf{b} \in \mathcal{R}(\mathbf{A})$: Infinitely many solutions. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

This system has infinitely many solutions.

* $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$: No solution. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

This system has no solution.

Hope this post gives a clear overview of linear systems of equations. Interested reader may explore further applications. Comments and clarifications are welcome.