Systems of linear equations

Generalities and Gaussian elimination

Francesco Preta, 07/14/2020

General definition

A system of m linear equation in n **unknown** $x_1, \ldots, x_n$ is an object of the type:

$$\begin{cases} A_{1,1}x_1 + A_{1,2}x_2 + \ldots + A_{1,n}x_n = b_1 \\ A_{2,1}x_1 + A_{2,2}x_2 + \ldots + A_{2,n}x_n = b_2 \\ \ldots \\ A_{m,1}x_1 + A_{m,2}x_2 + \ldots + A_{m,n}x_n = b_m \end{cases}$$

For given **coefficients** $\{A_{ij}\}_{i,j}$ and **constants** $\{b_j\}_j$

Solutions of a system of linear equations

Finding the solution to a system of linear equations is equivalent to finding values $(x_1, \ldots, x_n)$ that satisfy all the equations at once.

We call all the possible values that satisfy a system of linear equations, the **set of solutions** of the system. Two systems are **equivalent** if they have the same set of solutions.

Possible sets of solutions:

We can distinguish three cases:

- **One solution**: there is only one choice of $(x_1, \ldots, x_n)$ satisfying the system of linear equations.
- **Infinite solutions**: there are infinitely many choices of $(x_1, \ldots, x_n)$ satisfying the system of linear equations.
- **No solution:** there is no choice of $(x_1, \ldots, x_n)$ satisfying the system of linear equations.

In the first two cases, we say that the system is **consistent**, while if it has no solutions, the system is said to be **inconsistent**.

Intuition:

Each unknown represents a degree of freedom, while each equation represents an additional requirement to pin down a set of solutions, potentially reducing the degrees of freedom of the system.

If there are too many requirements, then it's possible that no solutions exist.

On the other hand, if the requirements are loose with respect to the degrees of freedom, then there could be an infinite number of solutions.

Intuition (continued)

However, two equations can give the same requirement. Think about the case in which we have equations:

$$x_1 + x_2 = 2$$
$$2x_1 + 2x_2 = 4$$

The second equation does not really give us more information to pin down the solution, as it's a multiple of the first. Therefore more equations don't automatically mean less degrees of freedom. Linear algebra will come to the rescue.

Some more definitions

If we have more unknowns than equations ($n > m$), then the system is said to be **under-determined**.

If there are more equations than unknowns ($m > n$) then the system is **over-determined**.

If $n = m$, the system is called **square**.

Matrix form

A system of linear equations can always be represented in matrix form as$A\mathbf{x} = \mathbf{b}$, where A is the $m \times n$ **coefficient matrix**, $\mathbf{x}$ is the $n \times 1$ vector of **unknowns** and $\mathbf{b}$ is the $m \times 1$ vector of **constants**. The matrix $[A \,|\, \mathbf{b}]$ is called the **augmented matrix** of the system.

Homogenous systems

Notice that if $\mathbf{b} = \mathbf{0}$, then there is always a solution to the system given by $x_i = 0$ for $i = 1, \dots, n$.

In this case, the system is called **homogeneous**.

Example 1:

$$\begin{cases} 3x_1 - 2x_2 = -2 \\ 2x_1 - x_3 = 0 \\ -6x_1 + 4x_2 + 2x_3 = 4 \end{cases}$$

Example 2:

$$\begin{cases} -4x_1 + 2x_3 + x_5 = 2 \\ 2x_1 + 3x_2 - x_4 = 8 \end{cases}$$

Example 3:

$$\begin{cases} 3x_1 + 2x_2 = 0 \\ 2x_1 - x_3 = 1 \\ -6x_1 + 4x_2 - 2x_3 = 0 \\ -2x_2 + x_3 = 1 \end{cases}$$

Back-substitution

Practice by which we use the equations of the systems to find a variable as a function of the others and substitute into the remaining equations to reduce the number of variables.

Ideally, we want to minimize the number of variables present in each equation to understand if the system is consistent or not.

Example 1:

$$\begin{cases} 3x_1 - 2x_2 = -2 \\ 2x_1 - x_3 = 0 \\ -6x_1 + 4x_2 + 2x_3 = 4 \end{cases}$$

Example 1: matrix form

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

Example 2:

$$\begin{cases} -4x_1 + 2x_3 + x_5 = 2 \\ 2x_1 + 3x_2 - x_4 = 8 \end{cases}$$

Example 2: matrix form

$$A' = \begin{bmatrix} -4 & 0 & 2 & 0 & 1 & 2 \\ 2 & 3 & 0 & -1 & 0 & 8 \end{bmatrix}$$

Example 3:

$$
\begin{cases}
3x_1 + 2x_2 = 0 \\
2x_1 - x_3 = 1 \\
-6x_1 + 4x_2 - 2x_3 = 0 \\
-2x_2 + x_3 = 1
\end{cases}
$$

Example 3: matrix form

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

Admissible row operations:

In matrix notation we have used the following row operations to switch between equivalent systems:

- **Replacement**: replace a row by a nontrivial linear combination of itself with another row.
- **Interchange:** interchange two rows.
- **Scaling:** multiply a row by a nonzero real number.

Gaussian elimination algorithm

Goal

We want to write an algorithm that for any given system returns either the set of solution or the information that the system has no solution.

The Gaussian elimination algorithm does so by reducing the augmented matrix A' to a unique form, called **reduced echelon form** on which a decision rule can be applied.

The reduction is done through the admissible row operations previously described.

Definition: Echelon form

We say that a rectangular matrix is in **echelon form** if it has the following properties:

- All nonzero rows are above all the rows of all zeros.
- Each leading term in a row has higher column coefficient than the leading term of any previous row.
- All entries in a column below the leading term of a row are zero.

Examples:

$$\begin{bmatrix}1 & 0 & -1\\0 & -1 & 2\\0 & 0 & 3\end{bmatrix} \quad \begin{bmatrix}1 & 0 & -1\\0 & 0 & 4\\0 & 0 & 0\end{bmatrix} \quad \begin{bmatrix}1 & 0 & -1 & 10\\0 & 0 & 2 & 2\\0 & 0 & 0 & 4\end{bmatrix}$$

Definition: Reduced echelon form

A rectangular matrix is in **reduced echelon form** if it is in echelon form and additionally:

- The leading term of each row is 1.
- Each leading 1 is the only nonzero element of its column.

Examples:

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

Equivalence between matrices

Two systems are equivalent if they have the same set of solutions.

In matrix notation, we say that two matrices are equivalent if they differ from one of the row operations we described before: replacement, interchange or scaling.

Uniqueness of reduced echelon form

Theorem

For every matrix A' there exists a unique equivalent matrix A'_{ref} in reduced echelon form.

Definition: Pivot

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 containing a pivot position.

The algorithm:

1. Choose the leftmost non-zero column: this is a pivot column.
2. Choose a nonzero term in the pivot column to be the pivot and exchange rows till this element is in the pivot position.
3. Use row replacement operations to create all zeros below the pivot.
4. Consider the submatrix formed by the rows below the pivot and repeat the first 3 steps until there is no submatrix left.
5. Beginning with the rightmost pivot, create zeros above each pivot and scale so that the pivot is equal to one.

Application of the algorithm

$$A' = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Applying the algorithm to our previous examples

Example 1:

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

Example 2:

$$A' = \begin{bmatrix} -4 & 0 & 2 & 0 & 1 & 2 \\ 2 & 3 & 0 & -1 & 0 & 8 \end{bmatrix}$$

Example 3:

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