02-680 Module 7

Essentials of Mathematics and Statistics

Aidan Jan

October 2, 2025

Linear Systems of Equations

We can define a **linear system** of n linear equations on m variables as follows:

$$\begin{array}{ccccccccc} C_{11}x_1 & + & C_{12}x_2 & + & \cdots & + & C_{1m}x_m & = & b_1 \\ C_{21}x_1 & + & C_{22}x_2 & + & \cdots & + & C_{2m}x_m & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ C_{n1}x_1 & + & C_{n2}x_2 & + & \cdots & + & C_{nm}x_m & = & b_n \end{array}$$

As an example:

$$\begin{cases} 3z_1 + 2z_2 = -1 \\ z_1 - 5z_2 = 3 \end{cases}$$

$$\begin{bmatrix} 3 & 2 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

If we want to find x (or in our example z), we can say it is

$$x = C^{-1}b$$

where C^{-1} is the inverse matrix of C. It turns out this matrix may not always exist.

- The inverse is the matrix such that $CC^{-1} = C^{-1}C = I_n$ (thus the first condition to the inverse existing is if C is square).
- The inverse exists as long as the determinant of the matrix is not zero.
- Because C^{-1} exists we call C **nonsingular**. (If the inverse does not exist, we would call it **singular**.)

Elementary Operations

The key to solving a **system of linear equations** are elementary transformations that keep the solution set the same, but that transform the equation system into a **simpler form**.

- Exchange of two equations, rows in the matrix representing the system of equations. (Type I, swap rows)
- Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R} \setminus \{0\}$. (Type II, scale a row)
- Addition of two equations (rows). (Type III, add row to another)

Elementary Matrices

An $n \times n$ matrix obtained from I_n by performing a single row operation is called an **elementary** $n \times n$ **matrix**.

- Proposition: Let A be $n \times p$, and assume E is an elementary $n \times n$ matrix. Then EA is the matrix obtained by performing the row operation corresponding to E on A.

Type I (swap rows)

Let

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Here we swapped the second and third rows of E. Then,

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 6 \\ 2 & 5 \end{bmatrix}$$

Notice that E is produced by swapping the 2nd and 3rd rows of I_3, and similarly the result EA is A with the 2nd and 3rd rows swapped.

- In the context of linear systems, doing this refers to swapping the order of the equations. This does not change the location of the intersection of the lines.

Type II (scale a row)

Let

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \quad \text{, where } a \neq 0$$

then,

$$GA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3\alpha & 6\alpha \end{bmatrix}$$

Notice in this case G is obtained from I_3 by multiplying the 3rd row by a and the same thing happens to A in the result.

- In the context of linear systems, a scalar product to a whole equation is still the same line in the vector space (the constant ends up reducing out).

Type III (add row to another)

Let

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \beta & 0 & 1 \end{bmatrix} \quad \text{where } \beta \in \mathbb{R}$$

then

$$GA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \beta & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 + 1\beta & 6 + 4\beta \end{bmatrix}$$

So in this case we add the value of the first row times β to the bottom row.

- In linear systems, finding the (weighted sum) of two equations is like finding a line that's "between" the two. In the plane you can think of it as a rotation around the solution.

Matrix Inverse

All of the elementary operations can be "undone" by finding reversing matrices. In other words, find the inverse.

Existence of the Inverse

- For an inverse to exist, the matrix must be **invertible**.
- We say that A is invertible if there is an $n \times n$ matrix B such that $AB = BA = I_n$.
- In this case, the matrix B is called the **inverse** of A, and we write $B = A^{-1}$
- The matrix is invertible if and only if its determinant is not zero. (and therefore can be expressed as a product of elementary matrices.)

Matrix Inversion Procedure

1. Create a partitioned matrix

$$\left[\begin{array}{cccc|cccc} A_{11} & A_{12} & \cdots & A_{1m} & 1 & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & A_{2m} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} & 0 & 0 & \cdots & 1 \end{array}\right]$$

2. Apply a **row operation** to the matrix created.
3. While the left hand side is not equal to I_n, repeat step 2.

For example, let's start with

$$\left[\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 1 & -5 & 0 & 1 \end{array}\right]$$

Subtract 3 times the second row from the first.

$$\left[\begin{array}{cc|cc} 0 & 17 & 1 & -3 \\ 1 & -5 & 0 & 1 \end{array}\right]$$

Divide the first row by 17.

$$\left[\begin{array}{cc|cc} 0 & 1 & \frac{1}{17} & -\frac{3}{17} \\ 1 & -5 & 0 & 1 \end{array}\right]$$

Swap the first and second rows.

$$\left[\begin{array}{cc|cc} 1 & -5 & 0 & 1 \\ 0 & 1 & \frac{1}{17} & -\frac{3}{17} \end{array}\right]$$

Add 5 times the second row to the first.

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{5}{17} & \frac{2}{17} \\ 0 & 1 & \frac{1}{17} & -\frac{3}{17} \end{array}\right]$$

Note that the right hand side is exactly the inverse of the matrix we saw originally! Note that this is not the only way of finding the inverse.

Invertible Matrix: Equivalence Conditions

$A \in \mathbb{R}^{n \times n}$ is invertible if and only if the following are true:

1. A is invertible ($\exists A^{-1}$ such that $AA^{-1} = A^{-1}A = I$)
2. $Ax = b$ has a unique solution for every $b \in \mathbb{R}^n$
3. $\text{rref}(A) = I$
4. $\det(A) \neq 0$
5. Columns of A are linearly independent

6. Rows of A are linearly independent
7. Columns of A span $\mathbb{R}^n$
8. Rows of A span $\mathbb{R}^n$
9. $\text{rank}(A) = n$
10. 0 is not an eigenvalue of A
11. A is row equivalent to the identity matrix
12. A is a product of elementary matrices.

Algebraic Properties of Invertible Matrix

$$\begin{aligned}
(A^{-1})^{-1} &= A \\
(AB)^{-1} &= B^{-1}A^{-1} \\
(A^T)^{-1} &= (A^{-1})^T \\
\det(A^{-1}) &= 1/\det(A)
\end{aligned}$$

Solving Linear Systems without an Inverse

The procedure above is great if C is nonsingular, but what if it is not? What if C isn't even square (that is, if the system is over- or under-defined?) It turns out we can use a very similar procedure to find a solution (if it exists). This procedure is called **Gauss-Jordan Elimination**. This time we set up a slightly different partitioned matrix:

$$[A|b] = \left[\begin{array}{cccc|c} A_{11} & A_{12} & \cdots & A_{1m} & b_1 \\ A_{21} & A_{22} & \cdots & A_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} & b_n \end{array}\right]$$

This single matrix contains our whole problem, and in fact every time we make an elementary operation on it, it **defines a new problem that has the same solution**. Remember the notes pointed out when talking about elementary operations. That means if we can get to a state where the left hand side is the identity, what we end up with is a system where every equation is simply one variable and a number (in the b column). So we can use, basically, the same procedure as before.

Let's look at our example from earlier, and suppose we set some values the equations are supposed to equal to.

$$\begin{array}{cc|c} 3 & 2 & -1 \\ 1 & -5 & 3 \end{array}$$

Once again, apply a Type III operation with $\beta = -3$ from row 2 to row 1.

$$\begin{array}{cc|c} 0 & 17 & -10 \\ 1 & -5 & 3 \end{array}$$

Apply Type II with $a = \frac{1}{17}$ to row 1.

$$\begin{array}{cc|c} 0 & 1 & -\frac{10}{17} \\ 1 & -5 & 3 \end{array}$$

Apply Type III with $\beta = 5$ from row 1 to row 2.

$$\begin{array}{cc|c} 0 & 1 & -\frac{10}{17} \\ 1 & 0 & \frac{1}{17} \end{array}$$

Apply Type I to swap row 1 and 2.

$$\left.\begin{array}{cc|c} 1 & 0 & \frac{1}{17} \\ 0 & 1 & -\frac{10}{17} \end{array}\right.$$

If we turn that back into a system of equations, we get

$$\begin{aligned} z_1 \quad &= \frac{1}{17} \\ z_2 &= -\frac{10}{17} \end{aligned}$$

Underdetermined and Overdetermined Systems

- An underdetermined system is when you have **fewer equations than unknowns**. Therefore, there is not enough information (constraints) to uniquely determine a single solution.
  - Instead of a point, the solution is a set of infinitely many vectors.
  - Although there is not a single solution, it is possible to determine a plane where all vectors on the plane serve as a solution.
  - We can represent the solution in terms of one or more unknowns (degrees of freedom).
- A determined system is where the number of equations equal the number of unknowns. Most of the time, there will be a single, unique solution. Occasionally we would have two parallel lines, where there is no solution.
- An overdetermined system has more equations than the number of unknowns. As such, there is unlikely to be a solution that satisfies every equation. More on this later.