Topic 7: Linear Systems & Matrix Inverses

02-680: Essentials of Mathematics and Statistics

September 26, 2024

1 Linear Systems

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

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

As an example:

$$
\begin{array}{ccccc}
3z_1 & + & 2z_2 & = & -1 \\
z_1 & - & 5z_2 & = & 3
\end{array}
$$

If we think of $x = \langle x_1, x_2, \ldots x_m \rangle \in \mathbb{R}^m$, then we can rewrite this system as:

$$
\begin{bmatrix}
C_{11} & C_{12} & \ldots & C_{1m} \\
C_{21} & C_{22} & \ldots & C_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
C_{n1} & C_{n2} & \ldots & C_{nm}
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ \vdots \\ x_n
\end{bmatrix}
=
\begin{bmatrix}
b_1 \\ b_2 \\ \vdots \\ b_n
\end{bmatrix}
$$

or $Cx = b$ for $C \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$ (or really $b \in \mathbb{R}^{n \times 1}$). When you expand the matrix multiplication you recover the original system. So in our example:

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

If we want to find x (or in our example z) we can say its

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

but we never talked about exponentials in matrices, and in fact C^{-1} is actually a special matrix called the ***inverse*** (and it turns out it may not 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). Remember from algebra we needed the number of equations to be equal to (or larger than) the number of free variables to solve a system.

So for C above, the inverse is:

$$C^{-1} = \begin{bmatrix} \frac{5}{17} & \frac{2}{17} \\ \frac{1}{17} & \frac{-3}{17} \end{bmatrix}$$

so

$$CC^{-1} = \begin{bmatrix} 3 & 2 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} \frac{5}{17} & \frac{2}{17} \\ \frac{1}{17} & \frac{-3}{17} \end{bmatrix} = \begin{bmatrix} (3\frac{5}{17} + 2\frac{1}{17}) & (3\frac{2}{17} + 2\frac{-3}{17}) \\ (1\frac{5}{17} + -5\frac{1}{17}) & (1\frac{2}{17} + -5\frac{-3}{17}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Because C^{-1} exists we call C **nonsingular** (if the inverse had *not* existed we would call it **singular**).

If you do the arithmetic:

$$x = \begin{bmatrix} \frac{5}{17} & \frac{2}{17} \\ \frac{1}{17} & \frac{-3}{17} \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} \left(\frac{5}{17} \cdot -1 + \frac{2}{17} \cdot 3\right) \\ \left(\frac{1}{17} \cdot -1 + \frac{-3}{17} \cdot 3\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{17} \\ \frac{-10}{17} \end{bmatrix}$$

This means that solving a system of equations is as simple as finding a matrix inverse!

2 Elementary Operations

We saw in the previous topic that for $A \in \mathbb{R}^{n \times m}$, $I_n A = A$. Elementary operations are simple manipulations of the identity matrix in this multiplication to obtain new matrices. So if

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

then

$$I_3 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

2.1 Type I (swap rows)

Let

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

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, similarly the result EA is A with the 2nd and 3rd row swapped.

> **Connection to linear systems**
>
> This has a correlation in linear systems: swapping the order of the equations. Notice that this does not change the location of the intersection of the lines.

2.2 Type II (scale a row)

Let

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

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 α, and the same thing happens to A in the result.

> **Connection to linear systems**
>
> Again, in linear systems, a scalar product to a whole equation is still the same line in the vector space (the constant ends up reducing out).

2.3 Type III (add row to another)

Let

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

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.

> **Connection to linear systems**
>
> In linear systems, finding the (weighted sum) of two equations is like finding a line thats "between" the two. In the plane you can think of it as a rotation around the solution.

All of these operations can be "undone" by finding reversing matrices $\overline{E}, \overline{G}, \overline{H}$ such that $\overline{E}(EA) = A$, $\overline{G}(GA) = A$, and $\overline{H}(HA) = A$. Try to find them on your own.

The question still remains: how do we find an inverse?!?

3   Finding a Matrix Inverse[*]

The proof of the following is beyond the scope of this course, but can be found in *Hoffman and Kunze* [p.23] (see citation below).

> Matrix $A \in \mathbb{R}^{n \times n}$ is invertible (nonsingular) iff there exists a set of elementary matrices $\{E_1, E_2, \ldots, E_k\}$ such that $A = E_1 E_2 \ldots E_k$. (That is, A is equal to the product of a set of elementary matrices.)

Again the proof is beyond this course but we take the following into account here as well:

> For $A_1, A_2, \ldots, A_n \in \mathbb{R}^{n \times n}$ which are all invertible, the product $A_1 A_2 \ldots A_n$ is also invertible. In addition, it's inverse $(A_1 A_2 \ldots A_{n-1} A_n)^{-1} = A_n^{-1} A_{n-1}^{-1} \ldots A_2^{-1} A_1^{-1}$.

Because A^{-1} is invertible (assuming A is) we can say $A^{-1} = E_k E_{k-1} \ldots E_2 E_1$. And in turn

$$(E_k E_{k-1} \ldots E_2 E_1) A = I_n.$$

Therefore, if we can find a set of elementary operations that convert A into I_n, the result of these operations is equal to A^{-1}.

---

[*]This is not the only method for finding the inverse, but others are not in line with this thread of information.

3.1 Procedure

(1) Create a *partitioned* matrix

$$\left[\; A \;\vdots\; I_n \;\right] = \left[\begin{array}{cccc:cccc} A_{11} & A_{12} & \dots & A_{1m} & 1 & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & A_{2m} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nm} & 0 & 0 & \dots & 1 \end{array}\right]$$

(2) Apply a row operation to the matrix created.

(3) While the *left* hand side is not equal to I_n, repeat step 2.

An Example. So applying the procedure to the example above:

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

Subtract 3 times the second row from the first (that is $\beta = -3$ in Type III).

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

Divide the first row by 17 (this is $\alpha = \frac{1}{17}$ in Type II).

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

Swap the first and second rows (Type I).

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

Add 5 times the second row to the first (Type III with $\beta = 5$).

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

We can now the right hand side is exactly the inverse we saw above!

Note, this is not the only way of finding the inverse, but will be helpful in the next section.

4   Solving Linear Systems without an Inverse

The procedure above is great if C is nonsingular, but what if its 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 time we set up a slightly different partitioned matrix:

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

This single matrix contains our whole problem, and in fact every time we make an elementary operation on it, it defines a new problem thats 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.

Lets look at our example:

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

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

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

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

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

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

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

If we turn that back into a system of equations, its:

$$\begin{array}{rcl} z_2 & = & \frac{-10}{17} \\ z_1 & = & \frac{1}{17} \end{array}$$

Exercise Determine (trough either method above) the solution to the following system:

$$\begin{aligned} -x_1 &+ 3x_2 &= 5 \\ x_1 &- 3x_2 &= 1 \end{aligned}$$

4.1 Example in Under Determined System

Lets look at the following system:

$$\begin{aligned} 2x_1 &- &x_2 &+ 4x_3 &+ &x_4 &- &x_5 &= 4 \\ 3x_1 &+ 2x_2 &- 2x_3 &- 5x_4 &+ 2x_5 &= 9 \\ x_1 &+ &x_2 &+ 6x_3 &+ 2x_4 &+ &x_5 &= 7 \end{aligned}$$

Since the system has only three equations for five dimensions, the solution cannot be a point, but can we determine the hyper-plane (if it exists) of the intersection?

Using the method from above lets set up the problem as a matrix reduction:

$$\left[\begin{array}{ccccc|c} 2 & -1 & 4 & 1 & -1 & 4 \\ 3 & 2 & -2 & -5 & 2 & 9 \\ 1 & 1 & 6 & 2 & 1 & 7 \end{array} \right]$$

Type I: swap rows 1 and 3 (hint puts the leading 1 in the top row):

$$\left[\begin{array}{ccccc|c} 1 & 1 & 6 & 2 & 1 & 7 \\ 3 & 2 & -2 & -5 & 2 & 9 \\ 2 & -1 & 4 & 1 & -1 & 4 \end{array} \right]$$

Type III (x2): add row 1 with $\beta = -3$ to row 2, and add row 1 with $\beta = -2$ to row 3:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 6 & 2 & 1 & 7 \\ 0 & -1 & -20 & -11 & -1 & -12 \\ 0 & -3 & -8 & -3 & -3 & -10 \end{array} \right]$$

Type II: scale row 2 with $\alpha = -1$ (again to get a leading 1):

$$\left[\begin{array}{ccccc|c} 1 & 1 & 6 & 2 & 1 & 7 \\ 0 & 1 & 20 & 11 & 1 & 12 \\ 0 & -3 & -8 & -3 & -3 & -10 \end{array} \right]$$

Type III (x2): add row 2 with $\beta = -1$ to row 1, and add row 2 with $\beta = 3$ to row 3:

$$\left[\begin{array}{ccccc|c} 1 & 0 & -14 & -9 & 0 & -5 \\ 0 & 1 & 20 & 11 & 1 & 12 \\ 0 & 0 & 52 & 30 & 0 & 26 \end{array}\right]$$

Type II: scale row 3 with $\alpha = \frac{1}{52}$ (again to get a leading 1):

$$\left[\begin{array}{ccccc|c} 1 & 0 & -14 & -9 & 0 & -5 \\ 0 & 1 & 20 & 11 & 1 & 12 \\ 0 & 0 & 1 & \frac{30}{52} & 0 & \frac{26}{52} \end{array}\right]$$

Type III (x2): add row 3 with $\beta = 14$ to row 1, and add row 3 with $\beta = -20$ to row 2:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -\frac{48}{52} & 0 & 2 \\ 0 & 1 & 0 & -\frac{28}{56} & 1 & 2 \\ 0 & 0 & 1 & \frac{30}{52} & 0 & \frac{26}{52} \end{array}\right]$$

Put that back into equation form:

$$
\begin{array}{rcl}
x_1 \quad\quad\quad - \quad -\frac{48}{52}x_4 \quad\quad\quad &=& 2 \\
x_2 \quad - \quad \frac{28}{56}x_4 \quad + \quad x_5 &=& 2 \\
x_3 \quad + \quad \frac{30}{52}x_4 \quad\quad\quad &=& \frac{26}{52}
\end{array}
$$

We can solve these equations for x_1, x_2, and x_3 then for any arbitrary x_4 and x_5 we have a valid solution to the system.

Useful References

Isaak and Monougian, "Basic Concepts of Linear Algebra". §1.1-1.6
Hoffman and Kunze, "Linear Algebra, 2e"