Unit 1: Solving Systems

Carter Teplica

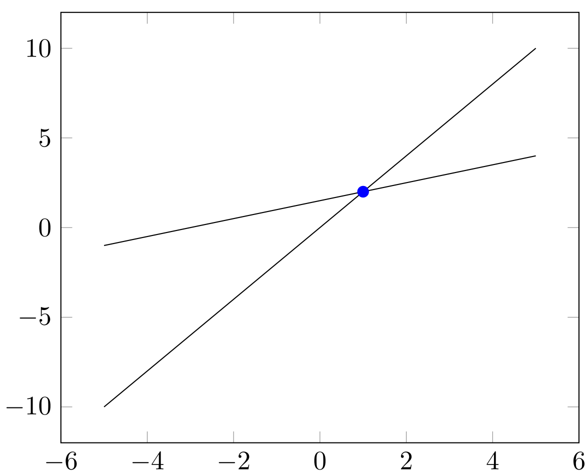
November 5, 2017

1 The Geometry of Linear Equations

Consider the following system of linear equations:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

There are three ways to think of the solution to the system. The row picture views each equation separately; the solution is their point of intersection.



The column picture views the solution (a vector) as the sum of other vectors:

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} x + \begin{bmatrix} -1 \\ 2 \end{bmatrix} y = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

The matrix picture is the most powerful, since we can do linear algebra with it!

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

We often simplify this to $\mathbf{Ax} = \mathbf{b}$.

2 Elimination

Elimination is an essential process in working with matrices. When doing elimination, one replaces each row with a linear combination of rows until the matrix is simpler. This includes multiplication, subtraction, and sometimes exchanging rows.

Consider the system

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

which can be rewritten as

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}.$$

We are trying to get the matrix into reduced row echelon form, meaning that the first nonzero entry in each column (called a *pivot*) is a 1, the rest of each column containing a pivot is filled with zeros, the pivots are in order (diagonally), and any zero rows are at the bottom.

First, we squish the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ together. We'll see later why we can do this.

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Now subtract three times the first row from the second row, so that the second row has a zero in its first column. We can leave the third row alone now, as it already has a zero in the first column.

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

The first column is done! The 1 in the first row and column is a pivot.

Now subtract the second row from the first row and subtract twice the second row from the third row.

$$\begin{bmatrix} 1 & 0 & 3 & -4 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{bmatrix}$$

The second column is done, for now. We'll wait till the end to change that 2 pivot into a 1.

Add $\frac{2}{5}$ of the third row to the second row and subtract $\frac{3}{5}$ of the third row from the first row.

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 5 & -10 \end{bmatrix}$$

We're almost in RREF! Now we can do the final step of dividing the rows to make the pivots equal 1.

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

Remembering what this means in the original equation, we can read off the solution:

$$\begin{aligned} x &= 2 \\ y &= 1 \\ z &= 2 \end{aligned}$$

From the perspective of a system of equations, it makes sense that this works. In each step, you are adding one valid equation to another to get a third valid equation. But why is this valid in terms of linear algebra?

First, note that when we squish two matrices together and left-multiply by something, we get the same thing that we would have gotten had we left-multiplied individually and then squished. Symbolically:

$$\mathbf{A} \times \begin{bmatrix} \mathbf{B} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \times \mathbf{B} & \mathbf{A} \times \mathbf{C} \end{bmatrix}$$

Each step of replacing a row with a linear combination of other rows is a left-multiplication. For instance, to subtract three times the top row from the second row, left-multiply by

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, with each step we are effectively multiplying the left and right sides of the equation by a matrix, which is a perfectly legal thing to do to an equation.

2.1 Factoring into (P)A = LU

When we do Gaussian elimination to get an upper-triangular matrix U we get $EA = U$, where E is the product of all the elimination matrices. (For now, we'll assume there are no row exchanges.) Left-multiplying the left and right sides of this equation by E^{-1} we get $E^{-1}EA = E^{-1}U$, so $A = E^{-1}U$. Let's rename E^{-1} as L, so that $A = LU$.

$A = LU$ is simpler than $EA = U$ because while each row subtraction shows up only once in L, it cam show up multiple times in E.

This almost always works. The only exception is for non-invertible matrices A, in which case we may need to permute the rows to move a zero pivot out of the way, so that

When using elimination to find $A = LU$ we need to do about $\frac{1}{3}n^3$ operations. (A few more operations, on the order of n^2, are needed for the right-hand side b, and on the order of n^2 are needed for the back-substitution process.)

A permutation matrix P is a matrix which permutes the rows of a matrix A when it is multiplied PA. A permutation matrix is just I with the rows reordered; for any permutation matrix, $P^T = P^{-1}$.

3 Matrix Multiplication

There are four ways to view what matrix multiplication does.

First, the cell method. The cell in the ith row and the jth column of $\mathbf{AB}$ is the dot product of the vectors formed by the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$.

Next is the column method. The jth column of $\mathbf{AB}$ is the matrix (cross) product of $\mathbf{A}$ and the jth column of $\mathbf{B}$.

The row method is very similar. The ith row of $\mathbf{AB}$ is the matrix (cross) product of the ith row of $\mathbf{A}$ and $\mathbf{B}$.

Finally, the row-and-column method. $\mathbf{AB}$ is the sum of all the products of the ith row of $\mathbf{A}$ and the ith column of $\mathbf{B}$. For instance,

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix}$$

The inverse of a square matrix $\mathbf{A}$ is a matrix $\mathbf{A}^{-1}$ such that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}.$$

While most square matrices have inverses, some do not. You can solve for the inverse using Gauss-Jordan elimination: simply start with $\begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix}$.

3.1 Properties of Inverses and Transposes

The inverse of a product $\mathbf{AB}$ is $\mathbf{B}^{-1}\mathbf{A}^{-1}$. The transpose of a product $\mathbf{AB}$ is $\mathbf{B}^T\mathbf{A}^T$.

A matrix A is *symmetric* if $A^T = A$. All symmetric matrices are of course square. For any (not necessarily square) matrix R, $R^T R$ is symmetric because $(R^T R)^T = (R)^T (R^T)^T = R^T R$.

4 Vector Spaces

A *vector space* is the set of all the linear combinations of some set of one or more vectors. Every vector space contains the zero vector, since any vector times 0 is 0. *2*