Matrices, Row Reduction of Matrices

October 29, 2014

1.2 Row Reduction and Echelon Forms

In the previous section, we saw a procedure for solving systems of equations. It is simple in that it consists of only applying one of three operations on equations, and then rewriting the system with the result in place of one of the other equations. However, when several variables are involved, this can be quite cumbersome. Thus, we will now introduce a simplifying formalism which is particularly convenient for automating the computations involved (say, by using a computer).

The fundamental concept we need is that of a *matrix* (plural: matrices). Simply put, a matrix is simply a collection of numbers, written in the form of a rectangle, usually written with either brackets or parentheses around them. For example, the following are both matrices, and are in fact the same matrix, even though I have written them slightly differently:

$$\begin{bmatrix} 5 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix} \quad \text{and} \quad \begin{pmatrix} 5 & 1 & 2 \\ 1 & 1 & 4 \end{pmatrix}$$

Note that they have to be symmetrical in the following way: you cannot have more numbers in one line than in another line. Aside from that, they can have as many lines as you want, and each line can contain as many numbers as you want. Thus, the following are also perfectly valid matrices (though in this case they are not the same):

$$\begin{bmatrix} 5 & 1 & 2 \\ 1 & 1 & 4 \\ 0 & 0 & 1 \\ 25 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{pmatrix} 5 & 1 & 2 & 0 & 0 \\ 1 & 1 & 4 & 10 & 10 \end{pmatrix}$$

Each number in a matrix is called an *entry*. We refer to all the entries along any horizontal line as a *row*, and the entries along a vertical line are called a *column*. If a matrix has m rows and n columns, we denote its *size* as $m \times n$. Thus, the matrix

$$\begin{bmatrix} 5 & 1 & 2 \\ 1 & 1 & 4 \\ 0 & 0 & 1 \\ 25 & 0 & 0 \end{bmatrix}$$

is a 4×3 matrix (not a "12 matrix").

Matrices and Systems of Linear Equations

Now that we have introduced matrices, we would like to see how they make our life easier in solving systems of equations. The idea is simple: to each system of equations, we can attach a matrix whose entries are simply the coefficients in the equations. This leads us to the following definitions.

Definition 1. Given a system of equations, the *coefficient matrix* for the system is a matrix where each row is determined by one equation in the system, and each column corresponds to one of the variables present in the system, such that each entry is the corresponding coefficient for that equation and that variable.

This definition may sound complicated, but an example will show how simple it really is.

Example 1. The system of equations

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0 \end{aligned}$$

has the corresponding coefficient matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Example 2. The system of equations

$$\begin{aligned} x_1 + x_2 - 2x_3 &= 2 \\ 2x_1 \qquad + \ x_3 &= 0 \end{aligned}$$

has the corresponding coefficient matrix

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

Definition 2. Given a system of equations, the *augmented matrix* for the system is the same as the coefficient matrix for the system, except that we add an extra column which corresponds to the right hand side of the equal signs in the system.

Example 3. The system of equations

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0 \end{aligned}$$

has the corresponding augmented matrix

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

Example 4. The system of equations

$$\begin{aligned} x_1 + x_2 - 2x_3 &= 2 \\ 2x_1 \qquad + \ x_3 &= 0 \end{aligned}$$

has the corresponding augmented matrix

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

Often, the augmented matrix is written with a dashed vertical line in between the last two columns to indicated that the equal signs go in between the two numbers. We will not do this here.

Solving Linear Systems with Matrices

Recall that when we discussed solving systems in the previous lecture, we stated that any system can be solved by repeatedly applying the following three operations:

1. Multiplying or dividing an equation by a non-zero number.

2. Adding or subtracting two equations from each other.

3. Switching two equations.

Each time we apply one of these operations, we obtain a new (but equivalent) system of equations, which will have its own corresponding coefficient and augmented matrices. Thus, instead of discussing operations on an equation or equations, we can talk about operations on a row or rows of a matrix, and equivalent matrices. It should (hopefully) be clear that the three operations we discussed for equations in a system correspond to the following operations on a matrix.

1. Multiplying or dividing a row by a non-zero number.

2. Adding or subtracting two rows from each other.

3. Switching two rows.

In the same way as for equations, we say that two matrices are equivalent if one can be obtained from the other simply by performing any combination of the three operations stated above.

Based on everything we have said so far, you may think that to use a matrix to solve a system, we will simply write the system as a matrix, and then proceed as we learned in the previous section. This is mostly correct, except there are a few modifications we can make which will make the procedure more suited for automation. (This class is meant for computer science students, after all.)

The procedure we will use goes by several names. It is usually referred to as *Gauss-Jordan elimination*, named after Carl Friedrich Gauss (whom you may know of from the many other things he has done) and Wilhelm Jordan (who no one cares about). Another name for the algorithm is simply *row reduction*. I will use the latter. Before we describe the algorithm, more definitions.

Definition 3. The first non-zero entry in a row of a matrix (when read from left to right) is called the *leading entry* of the row.

Definition 4. We say that a matrix is in *echelon form* if:

1. All rows with only zeros are at the bottom.

2. Each leading entry in a row is to the right of the previous leading entry.

3. Each entry below a leading entry is zero.

We say that a matrix is in *reduced echelon form* if it satisfies the additional properties

4. The leading entry in each row is 1.

5. Each leading 1 is the only non-zero entry in its column.

Example 5. The following matrix is in echelon form.

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

while the following matrix is in reduced echelon form

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

We can now state the algorithm known as Gauss-Jordan elimination.

Step 1: Start with the first entry in the first row. If it is 0, switch rows so it is non-zero. Once it is non-zero, we call it the first *pivot*.

Step 2: Use the row with the pivot to put zeros in all entries below the pivot by using combinations of the three allowed row operations on each row.

Step 3: Go down one entry, right one entry. If this entry is zero, switch with a row below it to make it non-zero. Once it is not zero, this is your new pivot. If you cannot make it non-zero, go right one entry and try again. Keep going like this until you find a non-zero entry to be your new pivot.
Step 4: Repeat steps 2 and 3 until you reach the last row.
Step 5: Beginning with the last pivot, create zeros in each entry above the pivot. Divide each row by an appropriate number to make each leading entry 1.

Example 6. Put the matrix

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

in reduced echelon form.

Solution: We make the first pivot the leading 1 in the first row. Then, we put zeros in the entry below the pivot by performing the following computation:

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} R_2 - R_1 \quad \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & -2 \end{bmatrix}$$

We can divide the new R_2 by -2 to get 1 in the leading entry:

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & -2 \end{bmatrix} \ -\frac{1}{2}R_2 \quad \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

Since we are in the last row, we can use the last pivot to put zeros in the previous rows:

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix} R_1 - R_2 \quad \Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

This matrix is in reduced echelon form.

Example 7. Put the matrix

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 1 & 5 \end{bmatrix}$$

in reduced echelon form.

Solution: Since we have a 1 in the top left, we use it as the first pivot. Then, the next step is to use that pivot to put zeros below the pivot. Since the last row has a zero already, we only need to get rid of the 2 in the second row:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 1 & 5 \end{bmatrix} R_2 - 2R_1 \quad \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 5 \end{bmatrix}$$

Next, we go down one, right one from the first pivot, which gives us the zero in the middle of the second row. Since we do not want a zero there, we switch with row 3:

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 5 \end{bmatrix} \text{Switch } R_2 \text{ and } R_3 \quad \Rightarrow \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

Since we now have a 1 in the leading entry, we can go down one, right one entry. But you can see that there is no way to put a 1 in that entry. So we can go to Step 5, returning to the last pivot, which is in the second row, and using that to put zeros in the entries above it:

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix} R_1 - R_2 \ \Rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is in reduced echelon form.