2 Gaussian elimination, systems of equations and matrix inverses

2.1 Matrices and systems of linear equations

Systems of Linear Equations

You have used systems of linear equations in lower level algebra courses, where the definition of systems of linear equations are just a set of two or more equations where more than one variable is combined with first degree polynomials with addition or subtraction. We do not multiply the variables, and the variables can only have up to the first power (nothing squared, cubed, etc) like this:

$$\begin{aligned} 2x + 3y &= 5 \\ -5x - 4y &= 8 \end{aligned}$$

In previous courses, you may have solved these equations with elimination, substitution, or other algebraic methods. In MATH 2400 and in linear algebra in general, we solve these systems by redefining how the equations are written as matrices instead. Equations are written in rows and the variables are the columns. For example, the matrix representation of the above example would be:

$$\begin{bmatrix} 2 & 3 \\ -5 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

Here, we have a 2x2 matrix which represents the coefficients of the system, also known as the **coefficient matrix**, then multiplied by a 2x1 matrix that represents the variables in the equation. This is set equal to a 2x1 matrix where the other side of the system is represented. We will discuss a special case where this is set equal to 0. It is obvious that these matrices represent the above system because if we were to do the matrix multiplication (will discuss later), we would get this:

$$\begin{bmatrix} 2x + 3y \\ -5x - 4y \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

Moreover, when writing this system to actually solve the system, we rewrite the system slightly by **augmenting** the system, and it looks like this:

$$\left[\begin{array}{cc|c} 2 & 3 & 5 \\ -5 & -4 & 8 \end{array}\right]$$

Gaussian elimination

Now that we have a good idea of what a matrix looks like and its conventions, we are going to discuss a process called Gaussian Elimination, which is also know as row reduction. Remember that our rows are the horizontal components of matrices, and each row represents an linear equation in the system. Given a linear system in the form $Ax = b$, our goal is to solve the system and determine the solution, x. In order to do this, we need to eliminate unnecessary information from our matrices in order to get to the true solution of the problem, and to do this, we perform row operations. Row operations can be performed on any size matrix, square or not.

Three Types of Row Operations

Type 1: Switching two rows

This row operation changes the location of two rows in a matrix with each other. It is fairly intuitive and looks like this:

$$\begin{bmatrix} 2 & 3 & 1 \\ -1 & 4 & 5 \\ -3 & 1 & 6 \end{bmatrix} \xrightarrow{R1=R2, R2=R1} \begin{bmatrix} -1 & 4 & 5 \\ 2 & 3 & 1 \\ -3 & 1 & 6 \end{bmatrix}$$

Type 2: Scaling a row

Here, we multiply one row by some scalar, it can be an integer or a fraction, positive or negative.

$$\begin{bmatrix} 2 & 3 & 1 \\ -1 & 4 & 5 \\ -3 & 1 & 6 \end{bmatrix} \xrightarrow{R1=3\times R1} \begin{bmatrix} 6 & 9 & 3 \\ -1 & 4 & 5 \\ -3 & 1 & 6 \end{bmatrix}$$

Type 3: Adding a scaled row to another row

In these operations, the most complicated of the row operations, we add a scaled, or non-scaled, row to another row.

$$\begin{bmatrix} 2 & 3 & 1 \\ -1 & 4 & 5 \\ -3 & 1 & 6 \end{bmatrix} \xrightarrow{R1=R1+3\times R2} \begin{bmatrix} -1 & 15 & 16 \\ -1 & 4 & 5 \\ -3 & 1 & 6 \end{bmatrix}$$

These row operations are relatively straightforward, just requiring some algebraic thinking. However, this is not Gaussian Elimination, we use row operations to perform Gaussian Elimination. Gaussian Elimination essentially reduces the system to its essence (its solution). When we perform Gaussian Elimination withe row operations, our goal is to make the matrix as simple as possible. Using Type 1, 2, 3 operations allow us to do this.

When doing row reduction, either we want the matrix to be in **Row Echelon form**, or more ideally, **Reduced Row Echelon form**.

Row Echelon Form

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

If a column contains a leading entry, in this case $3, 2, 5$, then all entries below it must be 0.

Reduced Row Echelon Form

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

All leading entries must be 1's, and all other entries in a column with a leading 1 must be 0.

Reduced Row Echelon Form is more ideal because it simplifies the system as much as possible and allows for easy solving of the system. Here, since it is a **square** 3×3 matrix, it must only have 1's along the diagonal, with zeros in all other locations. This matrix is called the **identity matrix**. Identity matrices are square matrices with only 1's along their diagonal. For MATH 2400 we will discuss finding the **R**educed **R**ow **E**chelon **F**orm (**RREF**) of a matrix.

Solving for RREF

We will be solving for the RREF of the following augmented system:

$$\left[\begin{array}{ccc|c} 5 & 10 & 5 & 15 \\ 3 & 12 & -9 & 9 \\ 4 & -4 & 12 & 8 \end{array}\right]$$

Step 1: Get the top left = 1

$$\left[\begin{array}{ccc|c} 5 & 10 & 5 & 15 \\ 3 & 12 & -9 & 9 \\ 4 & -4 & 12 & 8 \end{array}\right] \xrightarrow{R1=\frac{1}{5}\times R1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & 12 & -9 & 9 \\ 4 & -4 & 12 & 8 \end{array}\right]$$

Step 2: Use the top left 1 to get zeroes below it

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & 12 & -9 & 9 \\ 4 & -4 & 12 & 8 \end{array}\right] \xrightarrow{R2=R2-3\times R1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 4 & -4 & 12 & 8 \end{array}\right] \xrightarrow{R3=R3-4\times R1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & -12 & 8 & -4 \end{array}\right]$$

Step 3: Try to make middle-bottom = 0

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & -12 & 8 & -4 \end{array}\right] \xrightarrow{R3=R3+2\times R2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & 0 & -16 & -4 \end{array}\right]$$

Step 4: Make = 1 in bottom right corner

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & 0 & -16 & -4 \end{array}\right] \xrightarrow{R3=-\frac{1}{16}\times R3} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

Step 5: Make above that 1 = 0

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & -12 & 0 \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right] \xrightarrow{R2=R2+12\times R3} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

Step 6: Make = 1 in row 2 column 2

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 6 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right] \xrightarrow{R2=\frac{1}{6}\times R2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

Step 7: Make = 0 above our 1 in R3

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right] \xrightarrow{R1=R1-R3} \left[\begin{array}{ccc|c} 1 & 2 & 0 & \frac{11}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

Step 8: Make = 0 above our 1 in R2

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & \frac{11}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right] \xrightarrow{R1=R1-2\times R2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{7}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

The RREF of this example is:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{7}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

We see that in each row, the first entry is a 1. In RREF, the first entry must be 1 or all entries be 0. Further, in each column with a leading 1, there are only zeroes in the other entries in that column. RREF can look very different for different sizes of matrices and different systems:

Examples of RREF

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 6 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \begin{bmatrix} 1 & 0 & 0 & -3 & 4 \\ 0 & 1 & 0 & 4 & 9 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 3 & -4 \\ 0 & 1 & 0 & 5 \end{bmatrix}$$

Overall, Gaussian Elimination is extremely important to be able to solve systems of linear equations. The RREF of the matrix itself tells you the solution to the system, and also so much more, which we will discuss later.

Systems of linear equations

Now that we know how to rewrite a system of linear equations in the form of $Ax = b$ and learned Gaussian Elimination, we can solve these systems!

When you have a system of linear equations, we know that the number of rows in A we start with is the number of equations, and the number of columns in A is the number of variables. We are trying to find x, which are the values of our variables that we can plug into these equations to satisfy them.

x is a $n \times 1$ column vector, and A is a $m \times n$ matrix. Since we know our system is $Ax = b$, it follows that b is a $m \times 1$ column vector.

A is our coefficient matrix because it represents the coefficients of the variables in x, and then b contains what each equation is set equal to on the right hand side. If each of the equations are set $= 0$, then b is full of zeroes, and is called the **zero vector**.

The system is **homogeneous** if $b = 0$. We are now solving $Ax = 0$.

The system is **nonhomogeneous** if $b \neq 0$.

We solve these systems, regardless of homogeneity, by augmenting A with b, performing Gaussian Elimination, and then solving for the basis of the solution set. We can either find no solutions, one solution, or infinitely many solutions. Here is what a sample RREF matrix might look like for each of these cases:

<u>No solutions</u>

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 6 \end{array}\right]$$

<u>One solution</u>

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 3 \end{bmatrix}$$

<u>Infinite solutions</u>

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} t + \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}$$

As we can see, we have a system with **no solutions** when the RREF$(A|b)$ has a row of zeroes that is equal to a nonzero number, in this case 6. Logically, this system must have no solutions because this tells us $0x_1 + 0x_2 + 0x_3 = 6$, which does not make sense.

A system has **one solution** when there is a value for each variable. For square matrices, this is when $RREF(A) = I$.

A system has **infinitely many solutions** when the matrix has at least one zero row at the bottom of its RREF.

Further, when we solve a $Ax = b$ system and find the RREF of $A|b$, each column is associated with a variable in x, and the columns with leading 1's in them in the RREF are called **pivot/basic variables**, and the other variables associated with the other columns are called **free variables**. It follows that # pivot variables + # free variables = # columns.

Let's do a full example:
Solve the following $Ax = b$ (starting with the RREF) :

$$\left[\begin{array}{ccccc|c} 1 & 2 & 3 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Here, we are in RREF because each row starts with a leading 1, and that leading 1 is the only nonzero entry in that column. There are 5 variables in our solution set since there are 5 columns, so we know our solution is going to be in the form of 5×1 column vectors. There are two leading 1's in the spots for x_1 and x_4, so those are our pivot variables, and x_2, x_3, x_5 are our free variables.

Step 1: Extract equations from RREF

$$x_1 + 2x_2 + 3x_3 + 4x_5 = 0$$
$$x_4 + 5x_5 = 6$$

Step 2: Write pivot variables in terms in free variables

$$x_1 = -2x_2 - 3x_3 - 4x_5$$
$$x_4 = 6 - 5x_5$$

Step 3: Form solution skeleton

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} x_2 + \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} x_3 + \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} x_5$$

Since our system is non-homogeneous (not set = 0), we have the first red empty column to account for the non-homogenous vector there. If we were solving a homogeneous system, $Ax = 0$, the red vector would not be there.

Then, we have an empty column vector for each free variable.

Step 4: Plug equations into vectors

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 6 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} -4 \\ 0 \\ 0 \\ -5 \\ 1 \end{bmatrix} x_5$$

In orange is the equation from Step 2 for x_1. In blue is the equation from Step 2 for x_4. In pink, each free variable is set equal to itself (row 2 sets $x_2 = x_2$, row 3 sets $x_3 = x_3$, and row 5 sets $x_5 = x_5$). Then in green, all remaining spots have 0's.

Step 5: (Optional) Rewrite as a set

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 0 \\ -5 \\ 1 \end{bmatrix} \right\}$$

The first column vector is because $b \neq 0$ in $Ax = b$, but the other three vectors associated with the free variables are the solution to $Ax = 0$. The set of solutions to $Ax = b$ is the sum of the homogeneous solutions and the non-homogeneous solutions.