# Systems of Linear Equations

# Gaussian Elimination

*Hashim Mokhtar, PhD   2010, 2011*

It is quite hard to solve non-linear systems of equations, while linear systems are quite easy to study. There are numerical techniques which help to approximate nonlinear systems with linear ones in the hope that the solutions of the linear systems are close enough to the solutions of the nonlinear systems. We will not discuss this here. Instead, we will focus our attention on linear systems.

For the sake of simplicity, we will restrict ourselves to three, at most four, unknowns. The reader interested in the case of more unknowns may easily extend the following ideas.

**Definition.** The equation

$$a\,x + b\,y + c\,z + d\,w = h$$

where $a$, $b$, $c$, $d$, and $h$ are known numbers, while $x$, $y$, $z$, and $w$ are unknown numbers, is called a **linear equation**. If $h$ =0, the linear equation is said to be **homogeneous**. A **linear system** is a set of linear equations and a **homogeneous linear system** is a set of homogeneous linear equations.

For example,

$$\begin{cases} 2x - 3y &= 1 \\ x + 3y &= -2 \end{cases}$$

and

$$\begin{cases} x + y - z &= 1 \\ x + 3y + 3z &= -2 \end{cases}$$

are linear systems, while

$$\begin{cases} 2x - 3y^2 &= -1 \\ x + y + z &= 1 \end{cases}$$

is a nonlinear system (because of $y^2$). The system

$$\begin{cases} 2x - 3y - 3z + w = 0 \\ x + 3y \phantom{- 3z + w} = 0 \\ x - y + \phantom{3z +} w = 0 \end{cases}$$

is an homogeneous linear system.

**Matrix Representation of a Linear System**

Matrices are helpful in rewriting a linear system in a very simple form. The algebraic properties of matrices may then be used to solve systems. First, consider the linear system

$$\left\{\begin{array}{lcl} ax+by+cz+dw & = & e \\ fx+gy+hz+iw & = & j \\ kx+ly+mz+nw & = & p \\ qx+ry+sz+tw & = & u \end{array}\right. .$$

Set the matrices

$$A=\begin{pmatrix} a & b & c & d \\ f & g & h & i \\ k & l & m & n \\ q & r & s & t \end{pmatrix},\ C=\begin{pmatrix} e \\ j \\ p \\ u \end{pmatrix},\ \text{and}\ X=\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

Using matrix multiplications, we can rewrite the linear system above as the matrix equation

$$A\cdot X = C.$$

As you can see this is far nicer than the equations. But sometimes it is worth to solve the system directly without going through the matrix form. The matrix *A* is called the matrix coefficient of the linear system. The matrix *C* is called the **non-homogeneous term**. When $C=\mathcal{O}$, the linear system is homogeneous. The matrix *X* is the unknown matrix. Its entries are the unknowns of the linear system. The **augmented matrix** associated with the system is the matrix [*A*|*C*], where

$$[A|C] == \left(\begin{array}{cccc|c} a & b & c & d & e \\ f & g & h & i & j \\ k & l & m & n & p \\ q & r & s & t & u \end{array}\right).$$

In general if the linear system has n equations with m unknowns, then the matrix coefficient will be a nxm matrix and the augmented matrix an nx(m+1) matrix. Now we turn our attention to the solutions of a system.

**Definition.** Two linear systems with n unknowns are said to be **equivalent** if and only if they have the same set of solutions.

This definition is important since the idea behind solving a system is to find an equivalent system which is easy to solve. You may wonder how we will come up with such system? Easy, we do that through **elementary operations**. Indeed, it is clear that if we interchange two equations, the new system is still equivalent to the old one. If we multiply an equation with a nonzero number, we obtain a new system still equivalent to old one. And finally replacing one equation with the sum of two equations, we again obtain an equivalent system. These operations are called **elementary operations** on systems. Let us see how it works in a particular case.

**Example.** Consider the linear system

$$\left\{\begin{array}{rcl} x+y+z & = & 0 \\ x-2y+2z & = & 4 \\ x+2y-z & = & 2 \end{array}\right.$$

The idea is to keep the first equation and work on the last two. In doing that, we will try to kill one of the unknowns and solve for the other two. For example, if we keep the first and second equation, and subtract the first one from the last one, we get the equivalent system

$$\left\{\begin{array}{rrcl} x+ & y+z & = & 0 \\ x- & 2y+2z & = & 4 \\ & y-2z & = & 2 \end{array}\right.$$

Next we keep the first and the last equation, and we subtract the first from the second. We get the equivalent system

$$\left\{\begin{array}{rrcl} x+ & y+z & = & 0 \\ - & 3y+z & = & 4 \\ & y-2z & = & 2 \end{array}\right.$$

Now we focus on the second and the third equation. We repeat the same procedure. Try to kill one of the two unknowns (*y* or *z*). Indeed, we keep the first and second equation, and we add the second to the third after multiplying it by 3. We get

$$\left\{\begin{array}{rrrcl} x+ & y+ & z & = & 0 \\ - & 3y+ & z & = & 4 \\ & - & 5z & = & 10 \end{array}\right.$$

This obviously implies $z = -2$. From the second equation, we get $y = -2$, and finally from the first equation we get $x = 4$. Therefore the linear system has one solution

$$x = 4,\ y = -2,\ z = -2.$$

Going from the last equation to the first while solving for the unknowns is called **back-solving**.

Keep in mind that linear systems for which the matrix coefficient is upper-triangular are easy to solve. This is particularly true, if the matrix is in echelon form. So the trick is to perform elementary operations to transform the initial linear system into another one for which the coefficient matrix is in echelon form.
Using our knowledge about matrices, is there anyway we can rewrite what we did above in matrix form which will make our notation (or representation) easier? Indeed, consider the augmented matrix

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 1 & -2 & 2 & 4 \\ 1 & 2 & -1 & 2 \end{array}\right).$$

Let us perform some elementary row operations on this matrix. Indeed, if we keep the first and second row, and subtract the first one from the last one we get

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 1 & -2 & 2 & 4 \\ 0 & 1 & -2 & 2 \end{array}\right).$$

Next we keep the first and the last rows, and we subtract the first from the second. We get

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & 4 \\ 0 & 1 & -2 & 2 \end{array}\right).$$

Then we keep the first and second row, and we add the second to the third after multiplying it by 3 to get

$$\left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & 4 \\ 0 & 0 & -5 & 10 \end{array}\right).$$

This is a triangular matrix which is not in echelon form. The linear system for which this matrix is an augmented one is

$$\left\{\begin{array}{rrrcr} x+ & y+ & z & = & 0 \\ - & 3y+ & z & = & 4 \\ & - & 5z & = & 10 \end{array}\right.$$

As you can see we obtained the same system as before. In fact, we followed the same elementary operations performed above. In every step the new matrix was exactly the augmented matrix associated to the new system. This shows that instead of writing the systems over and over again, it is easy to play around with the elementary row operations and once we obtain a triangular matrix, write the associated linear system and then solve it. This is known as **Gaussian Elimination**. Let us summarize the procedure:

**Gaussian Elimination.** Consider a linear system.

**1.** Construct the augmented matrix for the system;
**2.** Use elementary row operations to transform the augmented matrix into a triangular one;
**3.** Write down the new linear system for which the triangular matrix is the associated augmented matrix;
**4.** Solve the new system. You may need to assign some parametric values to some unknowns, and then apply the method of back substitution to solve the new system.

**Example.** Solve the following system via Gaussian elimination

$$\left\{\begin{array}{rrrrrr} 2x- & 3y-z+ & 2w & +3v & = & 4 \\ 4x- & 4y-z+ & 4w & +11v & = & 4 \\ 2x- & 5y-2z+ & 2w & -v & = & 9 \\ & 2y+z & & +4v & = & -5 \end{array}\right.$$

The augmented matrix is

$$\left(\begin{array}{rrrrr|r} 2 & -3 & -1 & 2 & 3 & 4 \\ 4 & -4 & -1 & 4 & 11 & 4 \\ 2 & -5 & -2 & 2 & -1 & 9 \\ 0 & 2 & 1 & 0 & 4 & -5 \end{array}\right).$$

We use elementary row operations to transform this matrix into a triangular one. We keep the first row and use it to produce all zeros elsewhere in the first column. We have

$$\left(\begin{array}{rrrrr|r} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & -2 & -1 & 0 & -4 & 5 \\ 0 & 2 & 1 & 0 & 4 & -5 \end{array}\right).$$

Next we keep the first and second row and try to have zeros in the second column. We get

$$\left(\begin{array}{rrrrr|r} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \end{array}\right).$$

Next we keep the first three rows. We add the last one to the third to get

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

This is a triangular matrix. Its associated system is

$$\begin{cases} 2x - 3y - z + 2w + 3v = 4 \\ 2y + z + 5v = -4 \\ v = 1 \end{cases}$$

Clearly we have $v = 1$. Set $z=s$ and $w=t$, then we have

$$y = -2 - \frac{1}{2}z - \frac{5}{2}v = -\frac{9}{2} - \frac{1}{2}s.$$

The first equation implies

$$x = 2 + \frac{3}{2}y + \frac{1}{2}z - w - \frac{3}{2}v.$$

Using algebraic manipulations, we get

$$x = -\frac{25}{4} - \frac{1}{4}s - t.$$

Putting all the stuff together, we have

$$\begin{pmatrix} x \\ y \\ z \\ w \\ v \end{pmatrix} = \begin{pmatrix} -\frac{25}{4} - \frac{1}{4}s - t \\ -\frac{9}{2} - \frac{1}{2}s \\ s \\ t \\ 1 \end{pmatrix}.$$

**Example.** Use Gaussian elimination to solve the linear system

$$\begin{cases} x - y &= 4 \\ 2x - 2y &= -4 \end{cases}$$

The associated augmented matrix is

$$\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 2 & -2 & -4 \end{array}\right).$$

We keep the first row and subtract the first row multiplied by 2 from the second row. We get

$$\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 0 & 0 & -12 \end{array}\right).$$

This is a triangular matrix. The associated system is

$$\begin{cases} x - y &= 4 \\ \phantom{x - {}} 0 &= -12 \end{cases}$$

Clearly the second equation implies that this system has no solution. Therefore this linear system has no solution.

**Definition.** A linear system is called **inconsistent or over-determined** if it does not have a solution. In other words, the set of solutions is empty. Otherwise the linear system is called **consistent**.

Following the example above, we see that if we perform elementary row operations on the augmented matrix of the system and get a matrix with one of its rows equal to

$(0, 0, \cdots, 0, c)$

, where $c \neq 0$, then the system is inconsistent.

# Examples

Let's start simple, and work our way up to messier examples.

- **Solve the following system of equations.**

$$\begin{aligned} 5x + 4y - z &= 0 \\ 10y - 3z &= 11 \\ z &= 3 \end{aligned}$$

It's fairly easy to see how to proceed in this case. I'll just back-substitute the $z$-value from the third equation into the second equation, solve the result for $y$, and then plug $z$ and $y$ into the first equation and solve the result for $x$.

$10y - 3(3) = 11$
$10y - 9 = 11$
$10y = 20$
$y = 2$

$5x + 4(2) - (3) = 0$
$5x + 8 - 3 = 0$
$5x + 5 = 0$
$5x = -5$
$x = -1$

Then the solution is **$(x, y, z) = (-1, 2, 3)$.**

The reason this system was easy to solve is that the system was "triangular"; this refers to the equations having the form of a triangle, because of the lower equations containing only the later variables.

$$\begin{aligned} 5x + 4y - z &= 0 \\ 10y - 3z &= 11 \\ z &= 3 \end{aligned}$$

The point is that, in this format, the system is simple to solve. And Gaussian elimination is the method we'll use to convert systems to this upper triangular form, using the row operations we learned when we did the addition method.

- **Solve the following system of equations using Gaussian elimination.**

$$\begin{aligned} -3x + 2y - 6z &= 6 \\ 5x + 7y - 5z &= 6 \\ x + 4y - 2z &= 8 \end{aligned}$$

No equation is solved for a variable, so I'll have to do the multiplication-and-addition thing to simplify this system. In order to keep track of my work, I'll write down each step as I go. But I'll do my computations on scratch paper. Here is how I did it:

The first thing to do is to get rid of the leading $x$-terms in two of the rows. For now, I'll just look at which rows will be easy to clear out; I can switch rows later to get the system into "upper triangular" form. There is no rule that says I have to use the $x$-term from the first row, and, in this case, I think it will be simpler to use the $x$-term from the third row, since its coefficient is simply "$1$". So I'll multiply the third row by $3$, and add it to the first row. I do the computations on scratch paper:

$$\begin{array}{r} -3x + 2y - 6z = 6 \\ \underline{3x + 12y - 6z = 24} \\ 14y - 12z = 30 \end{array}$$

...and then I write down the results:

$$\begin{bmatrix} -3x + 2y - 6z = 6 \\ 5x + 7y - 5z = 6 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{3R_3 + R_1} \begin{bmatrix} 14y - 12z = 30 \\ 5x + 7y - 5z = 6 \\ x + 4y - 2z = 8 \end{bmatrix}$$

(When we were solving two-variable systems, we could multiply a row, rewriting the system off to the side, and then add down. There is no space for this in a three-variable system, which is why we need the scratch paper.)

Warning: Since I didn't actually do anything to the third row, I copied it down, unchanged, into the new matrix of equations. I *used* the third row, but I didn't actually *change* it. Don't confuse "using" with "changing".

To get smaller numbers for coefficients, I'll multiply the first row by one-half:

$$\begin{bmatrix} 14y - 12z = 30 \\ 5x + 7y - 5z = 6 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 7y - 6z = 15 \\ 5x + 7y - 5z = 6 \\ x + 4y - 2z = 8 \end{bmatrix}$$

Now I'll multiply the third row by $-5$ and add this to the second row. I do my work on scratch paper:

$$\begin{array}{r} 5x + 7y - 5z = 6 \\ \underline{-5x - 20y + 10z = -40} \\ -13y + 5z = -34 \end{array}$$

...and then I write down the results:

$$\begin{bmatrix} 7y - 6z = 15 \\ 5x + 7y - 5z = 6 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{-5R_3 + R_2} \begin{bmatrix} 7y - 6z = 15 \\ -13y + 5z = -34 \\ x + 4y - 2z = 8 \end{bmatrix}$$

I didn't do anything with the first row, so I copied it down unchanged. I worked with the third row, but I only worked *on* the second row, so the second row is updated and the third row is copied over unchanged.

Okay, now the $x$-column is cleared out except for the leading term in the third row. So next I have to work on the $y$-column.

Warning: Since the third equation has an $x$-term, I cannot use it on either of the other two equations any more (or I'll undo my progress). I can work *on* the equation, but not *with* it.

If I add twice the first row to the second row, this will give me a leading $1$ in the second row. I won't have gotten rid of the leading $y$-term in the second row, but I will have converted it (without getting involved in fractions) to a form that is simpler to deal with. (You should keep an eye out for this sort of simplification.) First I do the scratch work:

$$\begin{array}{r} -13y + 5z = -34 \\ \underline{14y - 12z = 30} \\ y - 7z = -4 \end{array}$$

...and then I write down the results:

$$\begin{bmatrix} 7y - 6z = 15 \\ -13y + 5z = -34 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{2R_1 + R_2} \begin{bmatrix} 7y - 6z = 15 \\ y - 7z = -4 \\ x + 4y - 2z = 8 \end{bmatrix}$$

Now I can use the second row to clear out the $y$-term in the first row. I'll multiply the second row by $-7$ and add. First I do the scratch work:

$$\begin{array}{r} 7y - 6z = 15 \\ \underline{-7y + 49z = 28} \\ 43z = 43 \end{array}$$

...and then I write down the results:

$$\begin{bmatrix} 7y - 6z = 15 \\ y - 7z = -4 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{-7R_2 + R_1} \begin{bmatrix} 43z = 43 \\ y - 7z = -4 \\ x + 4y - 2z = 8 \end{bmatrix}$$

I can tell what $z$ is now, but, just to be thorough, I'll divide the first row by $43$. Then I'll rearrange the rows to put them in upper-triangular form:

$$\begin{bmatrix} 43z = 43 \\ y - 7z = -4 \\ x + 4y - 2z = 8 \end{bmatrix} \xrightarrow{\frac{1}{43}R_1} \begin{bmatrix} z = 1 \\ y - 7z = -4 \\ x + 4y - 2z = 8 \end{bmatrix} \longrightarrow \begin{bmatrix} x + 4y - 2z = 8 \\ y - 7z = -4 \\ z = 1 \end{bmatrix}$$

Now I can start the process of back-solving:

$y - 7(1) = -4$
$y - 7 = -4$
$y = 3$

$x + 4(3) - 2(1) = 8$
$x + 12 - 2 = 8$
$x + 10 = 8$
$x = -2$

Then the solution is **$(x, y, z) = (-2, 3, 1)$**.

Note: There is nothing sacred about the steps I used in solving the above system; there was nothing special about how I solved this system. You could work in a different order or simplify different rows, and still come up with the correct answer. These systems are sufficiently complicated that there is unlikely to be one right way of computing the answer. So don't stress over "how did she know to do that next?", because there is no rule. I just did whatever struck my fancy; I did whatever seemed simplest or whatever came to mind first. Don't worry if you would have used completely different steps. As long as each step along the way is correct, you'll come up with the same answer.

---

In the above example, I could have gone further in my computations and been more thorough-going in my row operations, clearing out all the $y$-terms other than that in the second row and all the $z$-terms other than that in the first row. This is what the process would then have looked like:

$$\begin{bmatrix} -3x+2y-6z=6 \\ 5x+7y-5z=6 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{3R_3+R_1} \begin{bmatrix} 14y-12z=30 \\ 5x+7y-5z=6 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1}$$

$$\begin{bmatrix} 7y-6z=15 \\ 5x+7y-5z=6 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{-5R_3+R_2} \begin{bmatrix} 7y-6z=15 \\ -13y+5z=-34 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{2R_1+R_2}$$

$$\begin{bmatrix} 7y-6z=15 \\ y-7z=-4 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{-7R_2+R_1} \begin{bmatrix} 43z=43 \\ y-7z=-4 \\ x+4y-2z=8 \end{bmatrix} \xrightarrow{-4R_2+R_3}$$

$$\begin{bmatrix} 43z=43 \\ y-7z=-4 \\ x+26z=24 \end{bmatrix} \xrightarrow{\frac{1}{43}R_1} \begin{bmatrix} z=1 \\ y-7z=-4 \\ x+26z=24 \end{bmatrix} \begin{matrix} \xrightarrow{7R_1+R_2} \\ \xrightarrow{-26R_1+R_3} \end{matrix}$$

$$\begin{bmatrix} z=1 \\ y=3 \\ x=-2 \end{bmatrix} \longrightarrow \begin{bmatrix} x=-2 \\ y=3 \\ z=1 \end{bmatrix}$$

This way, I can just read off the values of $x$, $y$, and $z$, and I don't have to bother with the back-substitution. This more-complete method of solving is called "Gauss-Jordan elimination" (with the equations ending up in what is called "reduced-row-echelon form"). Many texts only go as far as Gaussian elimination, but I've always found it easier to continue on and do Gauss-Jordan.

Note that I did two row operations at once in that last step before switching the rows. As long as I'm not working *with* and working *on* the same row in the same step, this is okay. In this case, I was working *with* the first row and working *on* the second and third rows.