Introduction to Systems of Equations

Introduction

A *system of linear equations* is a list of m linear equations in a common set of variables $x_1, x_2, \ldots, x_n$.

$$\left\{\begin{array}{c} a_{1,1}\,x_1 + a_{1,2}\,x_2 + \cdots + a_{1,n}\,x_n = b_1 \\ a_{2,1}\,x_1 + a_{2,2}\,x_2 + \cdots + a_{2,n}\,x_n = b_2 \\ \vdots \\ a_{m,1}\,x_1 + a_{m,2}\,x_2 + \cdots + a_{m,n}\,x_n = b_m \end{array}\right\}$$

A *solution* for a set of m linear equations is a set of values for the variables $x_1, x_2, \ldots, x_n$ that solves all m of the equations simultaneously.

If m is not equal to n the situation gets slightly more complicated. For the moment, we will assume that $m = n$. Later we will come back to the case $m \neq n$.

Some systems of equations are much easier to solve than others. The easiest system to solve is the *diagonal system* .

$$\left\{\begin{array}{c} a_{1,1}\,x_1 = b_1 \\ a_{2,2}\,x_2 = b_2 \\ \vdots \\ a_{n,n}\,x_n = b_n \end{array}\right\}$$

which has solution

$$\left\{\begin{array}{c} x_1 = b_1/a_{1,1} \\ x_2 = b_2/a_{2,2} \\ \vdots \\ x_n = b_n/a_{n,n} \end{array}\right\}$$

Here also we set aside for the moment the potential complications caused by one of the coefficients $a_{k,k}$ being zero. We will revisit that issue later.

Another variant that is easy to solve is the *upper-triangular* system.

$$\left\{\begin{array}{c} a_{1,1}\,x_1 + a_{1,2}\,x_2 + a_{1,3}\,x_3 + \cdots + a_{1,n}\,x_n = b_1 \\ a_{2,2}\,x_2 + a_{2,3}\,x_3 + \cdots + a_{2,n}\,x_n = b_2 \\ a_{3,3}\,x_3 + \cdots + a_{3,n}\,x_n = b_3 \\ \vdots \\ a_{n-1,n-1}\,x_{n-1} + a_{n-1,n}\,x_n = b_{n-1} \\ a_{n,n}\,x_n = b_n \end{array}\right\}$$

This system can be solved by the method of *back-substitution*. You begin by solving the last equation

$$x_n = b_n / a_{n,n}$$

and substituting that result back into the equation before that.

$$a_{n-1,n-1}\, x_{n-1} + a_{n-1,n}\, \left(b_n / a_{n,n}\right) = b_{n-1}$$

$$x_{n-1} = \frac{-\frac{a_{n-1,n}\, b_n}{a_{n,n}} + b_{n-1}}{a_{n-1,n-1}}$$

The process repeats until we have solved each of the equations in turn.

A strategy for solving systems of equations

Now that we have seen that there are special systems of equations that are easy to solve, we can pursue a strategy that will allow us to solve a wider range of systems:

> Transform the system to be solved into a system that is both easy to solve and has the same solution as the original.

The trickiest requirement here is the second - we have to be very careful to make sure that any manipulations we perform do not eliminate solutions or introduce new ones.

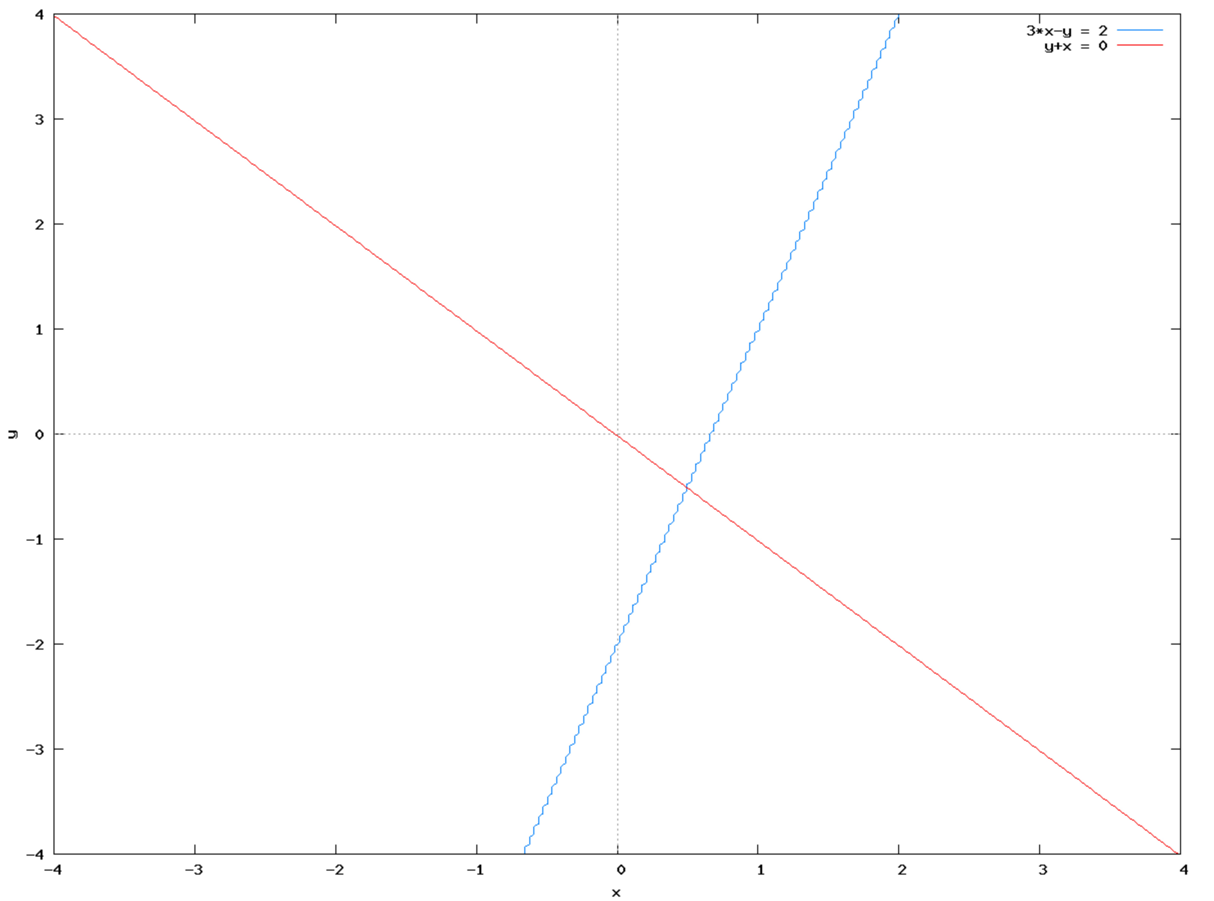
The manipulations we will use to transform systems of equations are called *elementary operations*. There are three elementary operations we will use.

1. Make two equations trade places in the list of equations.
2. Multiply both sides of an equation by a nonzero constant.
3. Replace one of the equations by a linear combination of itself and another equation.

The first two operations clearly will not affect the solutions. To get a better sense for what the third operation does, let's look at a concrete example. Here is a simple system of two equations in two unknowns.

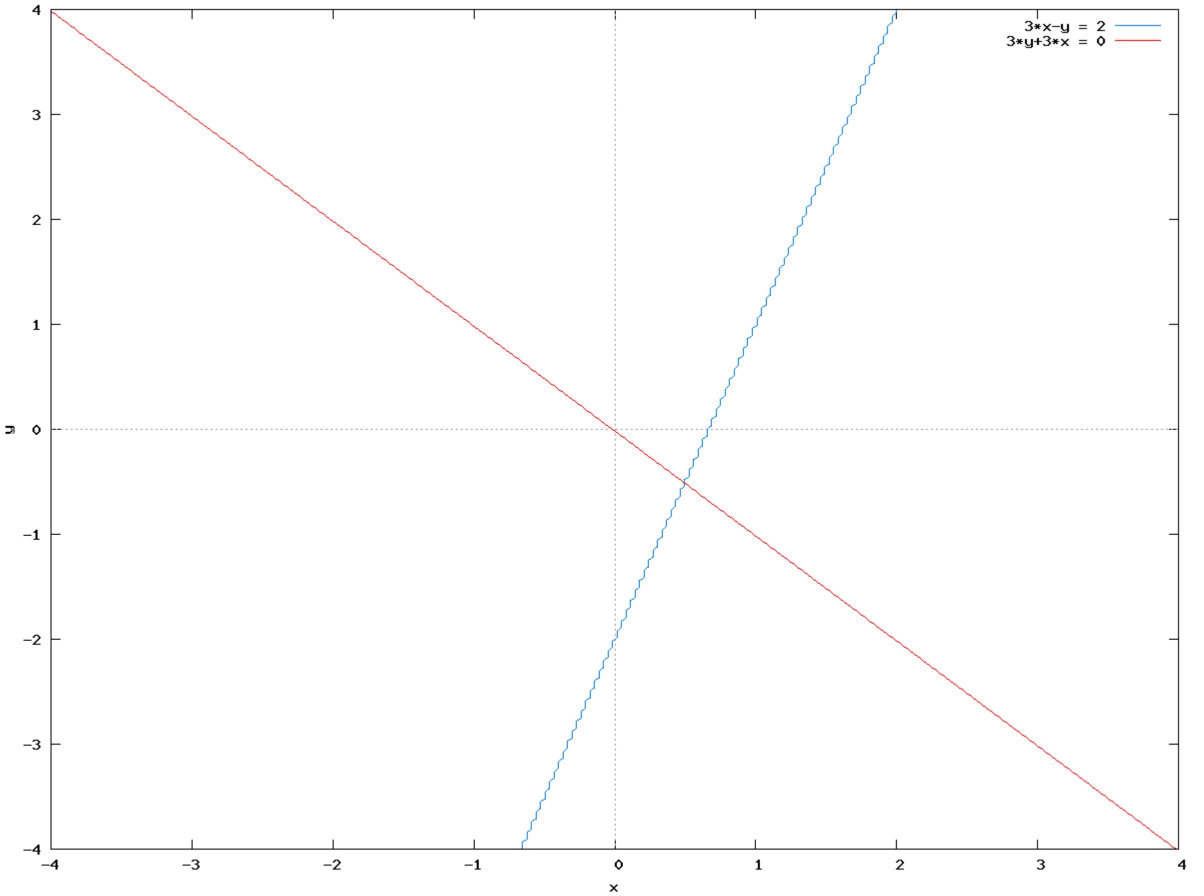
$$\left\{ \begin{array}{c} 3\, x - y = 2 \\ x + y = 0 \end{array} \right\}$$

Here is a plot of these two equations.



$3x - y = 2$
$x + y = 0$

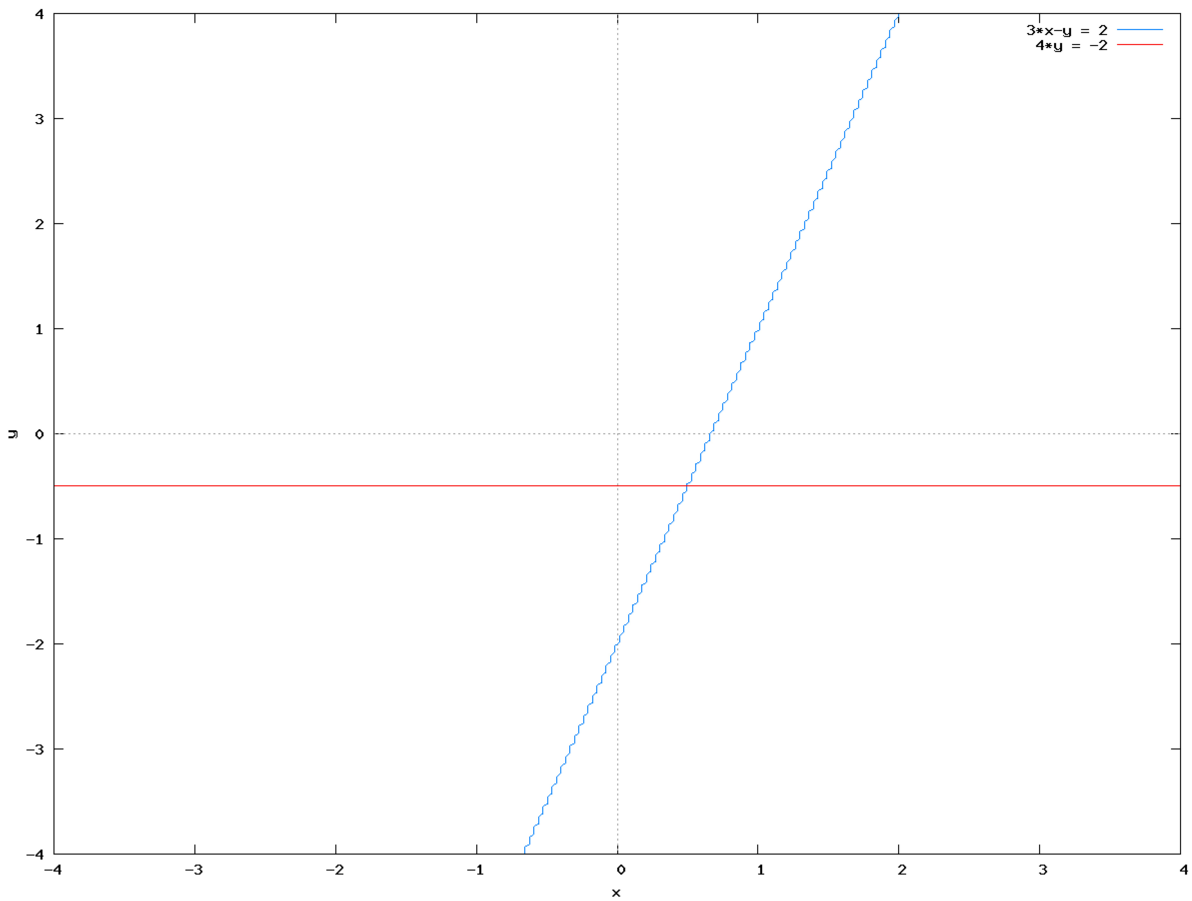
If we replace the second equation by 3 times the original, there is no effect on the system.



$$3x - y = 2$$

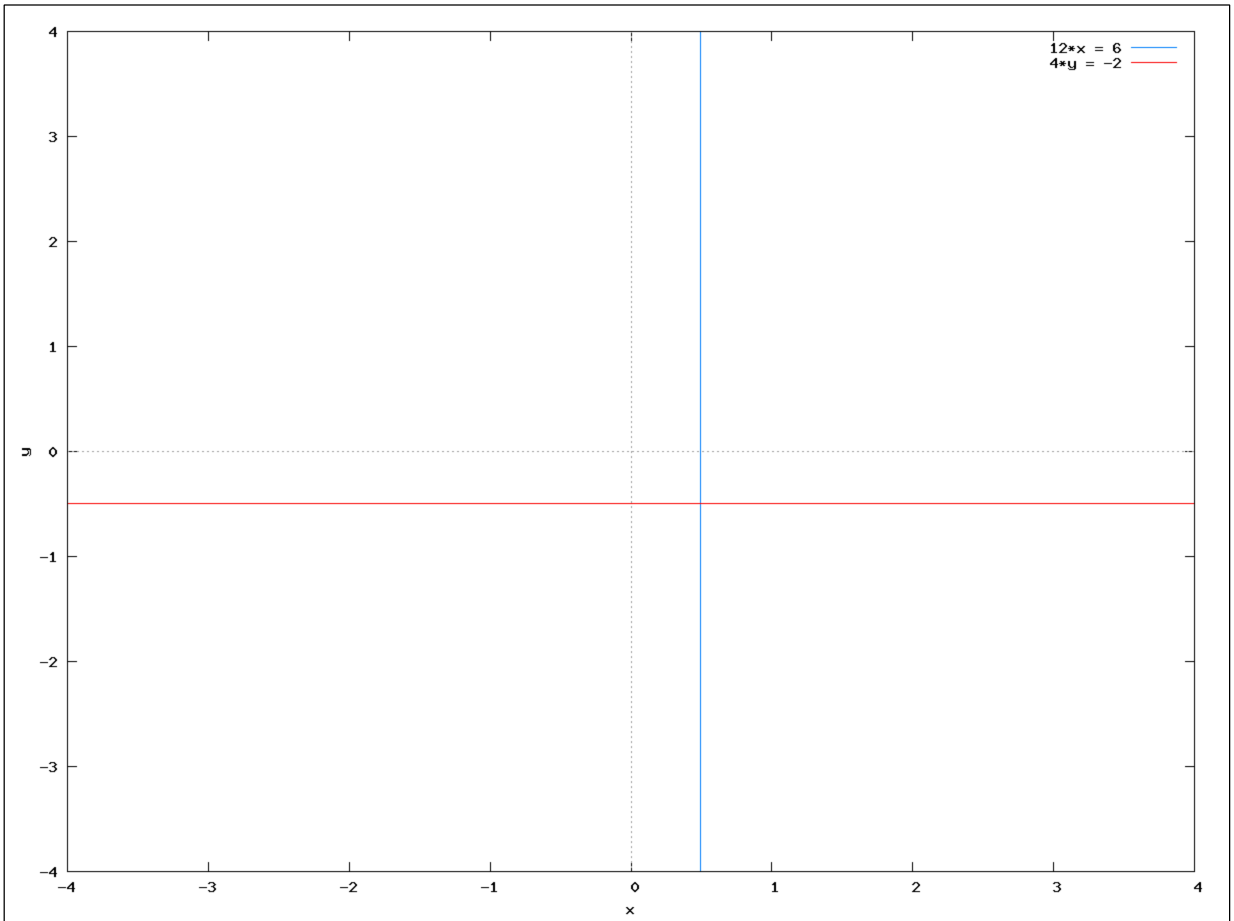
$$3x + 3y = 0$$

If we replace the second equation by the original second equation minus the first, the system of equations changes, but the solution does not.



$3\,x - y = 2$
$4\,y = -2$

Finally, we can replace the first equation with 4 times the first equation plus the second equation. Note that again the point where the two curves intersect does not move.



$12\,x = 6$
$4\,y = -2$

To summarize, we have the following

Theorem The system of equations that results from applying the elementary operations listed above has the same solutions as the original system.

Proof It suffices to show that the system of equations S_2 that results after applying any one of the elementary operations has exactly the same solution as the original system S_1. It is easy to see that this is the case for the first two elementary operations, because they both have a trivial effect on the systems. Consider an operation of type 3. Without loss of generality we can assume that the operation replaces equation 2 with the sum of equation 2 and some multiple of equation 1. Let $\{ x_1, x_2, \dots, x_n \}$ be any solution to system S_1. In particular, this set of values is a solution for the first two equations in S_1.

$$\left\{ \begin{matrix} a_{1,1}\,x_1 + a_{1,2}\,x_2 + \cdots + a_{1,n}\,x_n = b_1 \\ a_{2,1}\,x_1 + a_{2,2}\,x_2 + \cdots + a_{2,n}\,x_n = b_2 \end{matrix} \right\}$$

When we substitute this particular set of values, the first two equations should reduce to

$$\begin{Bmatrix} b_1 = b_1 \\ b_2 = b_2 \end{Bmatrix}$$

The first two equations in S_2 look like

$$\begin{Bmatrix} a_{1,1}\, x_1 + a_{1,2}\, x_2 + \cdots + a_{1,n}\, x_n = b_1 \\ c\, \left(a_{1,1}\, x_1 + a_{1,2}\, x_2 + \cdots + a_{1,n}\, x_n\right) + \left(a_{2,1}\, x_1 + a_{2,2}\, x_2 + \cdots + a_{2,n}\, x_n\right) = c\, b_1 + b_2 \end{Bmatrix}$$

When we substitute this same set of x values into these two equations the left hand sides reduce to

$$\begin{Bmatrix} b_1 = b_1 \\ c\, (b_1) + (b_2) = c\, b_1 + b_2 \end{Bmatrix}$$

Thus this same set of x values also satisfy all the equations in set S_2. The only thing left to prove is that any solution to set S_2 is also a solution to set S_1. This is left as an exercise for the reader.

Gaussian elimination

The Gaussian elimination method is an algorithm for applying elementary operations in a systematic way to transform a system of equations into a diagonal system.

The algorithm starts by writing the coefficients on the left hand side and the constant terms on the right hand side down in a rectangular grid called the *augmented matrix*. For example, the augmented matrix for the system of equations

$$\begin{Bmatrix} x - y + z = 0 \\ 2\, x - 3\, y + 4\, z = -2 \\ -2\, x - y + z = 7 \end{Bmatrix}$$

is

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 2 & -3 & 4 & -2 \\ -2 & -1 & 1 & 7 \end{bmatrix}$$

The algorithm attempt to reduce this matrix to *row-reduced form* through a process called *row reduction*. Reduced row-echelon form satisfies the following requirements.

1. Any rows consisting exclusively of zeros appears at the bottom.
2. The first non-zero entry of each row, called the *leading element* is 1.
3. The leading element of any row appears to the right of leading elements of rows above it.

Note that the augmented matrix for a diagonal system

$$\left\{ \begin{array}{c} x = 2 \\ y = -1 \\ z = 0 \end{array} \right\}$$

is in reduced row-echelon form:

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

Thus, trying to reduce an augmented matrix to this form is equivalent to trying to transform a set of equations into a diagonal set.

An example

Let us now try to transform the augmented matrix

$$A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 2 & -3 & 4 & -2 \\ -2 & -1 & 1 & 7 \end{bmatrix}$$

into reduced row-echelon form. The Gaussian elimination algorithm starts by working down the columns of the matrix turning all the entries below the leading entries to 0. The 1 in the upper left hand corner of the matrix is the leading entry for the first row. Our first task is to turn all the entries below it in the first column to 0. We will do this by replacing rows with linear combinations of existing rows.

In a typical row operation we will be combining a multiple of row i (A_i) with row j (A_j) in an effort to eliminate entry $a_{j,i}$. We will do this by constructing a *multipler* $m_{j,i} = a_{j,i}/a_{i,i}$ and then replacing A_j with A_j - $m_{j,i}$ A_i.

The first entry to eliminate is $a_{2,1}$. Our multiplier is $m_{2,1} = a_{2,1}/a_{1,1}$ = 2/1. We replace row 2 (A_2) with (A_2 - $m_{2,1}$ A_1).

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ -2 & -1 & 1 & 7 \end{bmatrix}$$

Then we form $m_{3,1} = a_{3,1}/a_{1,1}$ = -2/1 and replace A_3 with (A_3 - $m_{3,1}$ A_1).

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & -3 & 3 & 7 \end{bmatrix}$$

Finally we form $m_{3,2} = a_{3,2}/a_{2,2}$ = -3/(-1) and replace A_3 with (A_3 - $m_{3,2}$ A_2).

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & -3 & 13 \end{bmatrix}$$

The matrix is now in *upper-triangular form*. If we were to stop at this point and convert the augmented matrix back into a system of equations, we would have an upper-triangular system of equations that could be solved by back-substitution.

$$\left\{\begin{array}{c} x - y + z = 0 \\ -y + 2\, z = -2 \\ -3\, z = 13 \end{array}\right\}$$

An alternative technique, known as Gauss-Jordan elimination, goes on to eliminate all the entries above the diagonal. The result will be a diagonal system that is trivial to solve.

We will now convert all the entries above the diagonal to 0. We will start from the lower right corner.

We form $m_{2,3} = a_{2,3}/a_{3,3} = 2/(-3)$ and replace A_2 with $(A_2 - m_{2,3}\, A_3)$.

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 20/3 \\ 0 & 0 & -3 & 13 \end{bmatrix}$$

We form $m_{1,3} = a_{1,3}/a_{3,3} = 1/(-3)$ and replace A_1 with $(A_1 - m_{1,3}\, A_3)$.

$$\begin{bmatrix} 1 & -1 & 0 & 13/3 \\ 0 & -1 & 0 & 20/3 \\ 0 & 0 & -3 & 13 \end{bmatrix}$$

Finally, we form $m_{1,2} = a_{1,2}/a_{2,2} = -1/(-1)$ and replace A_1 with $(A_1 - m_{1,3}\, A_3)$.

$$\begin{bmatrix} 1 & 0 & 0 & -7/3 \\ 0 & -1 & 0 & 20/3 \\ 0 & 0 & -3 & 13 \end{bmatrix}$$

Converting this back to a system of equations gives

$$\left\{\begin{array}{c} x = -7/3 \\ -y = 20/3 \\ -3\, z = 13 \end{array}\right\}$$

or equivalently

$$\left\{\begin{array}{c} x = -7/3 \\ y = -20/3 \\ z = -13/3 \end{array}\right\}$$

You can substitute this back into to confirm that this is in fact a solution to the original system of equations.

A second example

Lets apply Gaussian elimination to the system

$$\left\{\begin{array}{c} x_1 + x_2 - x_3 + 2\, x_4 = 1 \\ x_1 + x_2 + x_4 = 2 \\ x_1 + 2\, x_2 - 4\, x_3 = 1 \\ 2\, x_1 + x_2 + 2\, x_3 + 5\, x_4 = 1 \end{array}\right\}$$

The augmented matrix is

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

Eliminating down the first row gives

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

After switching rows 2 and 3 and eliminating down the second column we have

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

Finally, eliminating down the third column produces.

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

At this point we appear to have gotten ourselves into a bind. The last equations says effectively

$$0\,x_1 + 0\,x_2 + 0\,x_3 + 0\,x_4 = -2$$

which is clearly impossible. The only way out of this is to say that this system of equations has no solution.

A third example

Here is a very similar example that turns out differently.

$$\left\{ \begin{array}{c} x_1 + x_2 - x_3 + 2\,x_4 = 1 \\ x_1 + x_2 + x_4 = 0 \\ x_1 + 2\,x_2 - 4\,x_3 = 1 \\ 2\,x_1 + x_2 + 2\,x_3 + 5\,x_4 = 1 \end{array} \right\}$$

Here is the augmented matrix.

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

After Gaussian-Jordan elimination we end up with

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

This is a slightly different situation. In this case, the row of all zeros along the bottom does not hurt anything (and neither does it help). Converting the augmented matrix back to equations gives us.

$$\begin{Bmatrix} x_1 + 6\,x_4 = 3 \\ x_2 - 5\,x_4 = -3 \\ x_3 - x_4 = -1 \\ 0 = 0 \end{Bmatrix}$$

What has happened here is that the system has a *free parameter*. The solutions tell us nothing about the value of x_4 . In fact, x_4 can take any value - it becomes a free parameter. All the other variables can be expressed in terms of that free parameter. To emphasize this point, it is traditional to replace the free parameters with variables whose names (s, t, etc.) suggest that they are parameters that can vary freely. Replacing x_4 with t gives us

$$\begin{Bmatrix} x_1 + 6\,t = 3 \\ x_2 - 5\,t = -3 \\ x_3 - t = -1 \\ 0 = 0 \end{Bmatrix}$$

or equivalently

$$\begin{Bmatrix} x_1 = 3 - 6\,t \\ x_2 = -3 + 5\,t \\ x_3 = -1 + t \\ x_4 = t \end{Bmatrix}$$

which highlights the dependence of the variables on the free parameter t.