

Solving Linear Systems¹

We will work through several examples. A little later we will show how they are special cases of a general method.

Example 1. Consider the system of equations below. It is equivalent to the matrix equation next to it.

$$\begin{array}{rclcrcl} x & + & y & - & z & = & 0 \\ 2x & & & & + & z & = & 2 \\ x & - & 2y & + & 3z & = & 1 \end{array} \quad \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

We can solve the system by simplifying it and eliminating variables. Each manipulation of the equations is mirrored by a manipulation of the matrix equation. This will be easier to see if we use a device call the **augmented matrix**. Given a matrix equation $A\mathbf{x} = \mathbf{b}$, where A and \mathbf{b} are known and \mathbf{x} is a vector of variables, define the associated augmented matrix to be $[A \mid \mathbf{b}]$. For the system we are studying the augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 2 & 0 & 1 & 2 \\ 1 & -2 & 3 & 1 \end{array} \right].$$

The vertical lines are not really needed, but are often used to remind us that the matrix arose as an augmentation. Each manipulation of the equations below does not effect the solution set. We proceed.

$$\begin{array}{rclcrcl} x & + & y & - & z & = & 0 \\ 2x & & & & + & z & = & 2 \\ x & - & 2y & + & 3z & = & 1 \end{array} \quad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 2 & 0 & 1 & 2 \\ 1 & -2 & 3 & 1 \end{array} \right].$$

- Subtract twice row 1 from row 2. ($R_2 - 2R_1 \mapsto R_2$)

$$\begin{array}{rclcrcl} x & + & y & - & z & = & 0 \\ & & -2y & + & 3z & = & 2 \\ x & - & 2y & + & 3z & = & 1 \end{array} \quad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 3 & 2 \\ 1 & -2 & 3 & 1 \end{array} \right].$$

- Subtract row 1 from row 3. ($R_3 - R_1 \mapsto R_3$)

$$\begin{array}{rclcrcl} x & + & y & - & z & = & 0 \\ & & -2y & + & 3z & = & 2 \\ & & -3y & + & 4z & = & 1 \end{array} \quad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 3 & 2 \\ 0 & -3 & 4 & 1 \end{array} \right].$$

¹©Michael C. Sullivan, June 13, 2008

We will only show the augmented matrices. The reader should write out the corresponding system of equations for each step and use the notation in Example 1 to indicate the corresponding row operations. Row operations are defined formally in the next section.

The first row operation we make is to switch row 1 with row 2 (denote this by $R_1 \leftrightarrow R_2$). This certainly does not change the solution set.

$$\begin{aligned} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 0 \\ 2 & 2 & 0 & -2 & 2 \\ 3 & 4 & 2 & 3 & 0 \end{array} \right] &\rightarrow \left[\begin{array}{cccc|c} 2 & 2 & 0 & -2 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 4 & 2 & 3 & 0 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 4 & 2 & 3 & 0 \end{array} \right] \rightarrow \\ &\left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 6 & -3 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 5 & -3 \end{array} \right] \rightarrow \\ &\left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 5/2 & -3/2 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 5/2 & -3/2 \end{array} \right] \end{aligned}$$

This translates back to,

$$\begin{aligned} w - 2z &= 1 \\ x + z &= 0 \\ y + \frac{5}{2}z &= -\frac{3}{2} \end{aligned}$$

Below we solve for w , x and y in terms of z . We regard z as a **free** or **independent variable**, and the others as **dependent variables**. We have also added the equation $z = z$ below. This certainly does not change the solution set, but including it will help us to understand the *geometric structure* of the solution set.

$$\begin{aligned} w &= 1 + 2z \\ x &= -z \\ y &= -\frac{3}{2} - \frac{5}{2}z \\ z &= z \end{aligned}$$

Rewrite this as a vector equation,

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -\frac{3}{2} \\ 0 \end{bmatrix} + z \begin{bmatrix} 2 \\ -1 \\ -\frac{5}{2} \\ 1 \end{bmatrix}$$

We replace the z on the right-hand side with t to emphasize that it is a free parameter. (This is not mathematically necessary but it has become the custom.)

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -\frac{3}{2} \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -1 \\ -\frac{5}{2} \\ 1 \end{bmatrix}$$

The solution set can now be thought of as a line in \mathbb{R}^4 .

Example 3. Consider the system of equations (expressed as a matrix equation),

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 2 & -4 & -4 \\ 1 & 0 & 2 & 1 & 1 & -2 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$

Starting from the augmented matrix you can use row operations to get the matrix below. Try to do this on your own.

$$\left[\begin{array}{ccccccc|c} 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

We thus have (ignoring $0 = 0$, the equation for the last row),

$$\begin{aligned} a + 2c &= 1 \\ d &= 1 \\ e - 2f - 2g &= 0 \end{aligned}$$

There are three (3) dependent variables and four (4) independent variables. We will solve for a , d , and e in terms of the free variables, b , c , f , and g ; and we add the identity equations for the four free variables.

$$\begin{aligned}
a &= 1 - 2c \\
b &= b \\
c &= c \\
d &= 1 \\
e &= 2f + 2g \\
f &= f \\
g &= g
\end{aligned}$$

In vector form we get

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + f \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} + g \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

The solution set could be thought of as a “4-dimensional plane” sitting in \mathbb{R}^7 .

Remark. The choice as to which variables are free and which are dependent is somewhat arbitrary. In the above example we could replace the equation $a = 1 - 2c$ with $c = 1/2 - a/2$, and replace $c = c$ with $a = a$. Then a would be free and c would be dependent. The solution set is, of course, the same. (Notice, however that we cannot choose to make d a free variable.) It is customary to use the variable corresponding to the pivot columns and the dependent variables. However, and this is extremely important, the number of free variables is not arbitrary, it is always the same no matter how the equations are manipulated. This fact will be proved after we have studied vector spaces. The number of free variables is called the **dimension** of the solution set.

In each of the examples above the original augmented matrix was converted into a special form. This form is called **reduced row echelon form**. We will define it carefully below. Most computer algebra programs, and even many calculators, have a command (often called **rref**) that puts a matrix into reduced row echelon form. It turns out that given an augmented matrix M there is only one matrix N that is in reduced row echelon form and has

the same solution set for its associated system of homogeneous equations. We write $N = \text{rref}(M)$, and think of rref as a function. We will prove this claim later.

Definition .1. A matrix is in **reduced row echelon form** if the following hold.

- The first nonzero entry of each row is a 1, or the row has only zeros - such a 1 is called a **pivot**,
- If a column contains a pivot, all the other entries of that column are zero.
- The rows are ordered so that the column indices of the pivots increase as their row indices increase and any rows without pivots, rows with only zeros, come last.

Check that in each of the examples above the final matrix is in reduced row echelon form; circle the pivots. The number of pivots in $\text{rref}(A)$ is called the **rank** of A . The customary method of choosing the dependent variables when there are solutions is to use the variables whose corresponding columns contain a pivot; the rest are free. Thus, if A has n columns the number of free variable parameterizing the solution set of $A\mathbf{x} = \mathbf{0}$ is n minus the rank of A .

We give two examples where there are no solutions.

Example 4. The linear system $0x = 2$ has no solutions. (Notice that the augmented matrix for $0x = 2$ is $[0|2]$.)

Example 5. Now consider

$$\begin{bmatrix} 2 & 3 \\ 3 & 9/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

We get

$$\left[\begin{array}{cc|c} 2 & 3 & 2 \\ 3 & 9/2 & 3 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 3/2 & 1 \\ 3 & 9/2 & 3 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 3/2 & 1 \\ 0 & 0 & -2 \end{array} \right]$$

But this translates to

$$\begin{aligned} x + 3/2y &= 1 \\ 0x + 0y &= -2 \end{aligned}$$

But there are no values of x and y that will make $0 = -2$. Thus, there are no solutions to the original system. In general, any time an augmented matrix has a row with all zeros except for a nonzero number in the last (right-most) entry, the corresponding system of equations will have no solutions.

This leads to the following definition.

Definition .2. A system of equations given by $A\mathbf{x} = \mathbf{b}$ is **consistent** if $\text{rref}([A|\mathbf{b}])$ does not have a row that has all zeros except for a nonzero entry in the right most place. If there is such a row, the system of equations is **inconsistent**.

Obviously, $A\mathbf{x} = \mathbf{b}$ has solutions if and only if it is consistent. It follows that a system of the form $A\mathbf{x} = \mathbf{0}$ always has at least one solution, namely $\mathbf{x} = \mathbf{0}$! Systems of the form $A\mathbf{x} = \mathbf{0}$ are called **homogeneous** and play a special role in linear algebra.

Theorem .3. Let A be an $m \times n$ matrix, \mathbf{b} be a $1 \times m$ column vector, $\mathbf{0}$ be the $1 \times m$ column vector with all entries zero and \mathbf{x} be a $1 \times n$ column vector of variables.

1. $\text{rank}(A) \leq n$, the number of variables and $\text{rank}(A) \leq m$ the number of rows.
2. If $m = n = \text{rank}(A)$, then there is a unique solution to $A\mathbf{x} = \mathbf{b}$.
3. If $m = n$, $\text{rank}(A) < n$, and $\text{rref}([A|\mathbf{b}])$ is consistent, then there are infinitely many solutions to $A\mathbf{x} = \mathbf{b}$.
4. If $m < n$ (fewer equations than variables) and $\text{rref}([A|\mathbf{b}])$ is consistent, then $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.
5. If $m < n$, then $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions.
6. If $m > n$ (more equations than variables), $\text{rref}([A|\mathbf{b}])$ is consistent and $\text{rank}(A) = n$, then there is a unique solution.
7. If $m > n$, $\text{rref}([A|\mathbf{b}])$ is consistent and $\text{rank}(A) < n$, then there are infinitely many solutions.

Proof. 1. The rank of A is the number of pivots and there is at most one pivot per column by definition. The same holds for the rows.

2. A is a square matrix and the reduced row echelon form of $[A|\mathbf{b}]$ has one pivot for each row. Thus, $\text{rref}([A|\mathbf{b}])$ has the form $[I|\mathbf{b}']$ so $\mathbf{x} = \mathbf{b}'$ in the sole solution.

3. There are $n - \text{rank}(A)$ free variables and no inconsistent rows in $\text{rref}([A|\mathbf{b}])$. Thus there is a parametrization of the solution set with one or more free variable and so the solution set has infinitely many members.

4. The rank of A is less than n . So, just as in 3 there are one or more free variables.

5. Statement 5 follows from Statement 4 since $\text{rref}([A|\mathbf{0}])$ must be consistent.

6. The proof is similar to 2.

7. The proof is similar to 3.

□

Problem 1. Find the solution set of each of the systems of equations below. Do so by setting up the augmented matrix, computing its rref , and then solving the equivalent system. Express your final answers in vector form, as in the examples, indicate which variables you have chosen as free and state the dimension of the solution set.

- a. $3x - y = -10$, $7x + 2y = 7$ and $2x - 5y = -37$. b. $2x - y = 7$
c. $3a - 2b = -8$ and $4a + 5b = -3$. d. $a + b - 3c + d = 0$ and $a - b + c - 2d = 2$.
e. $2w + 2x + y - z = -6$, $w + 3x + 4y = 1$, $6w + 5y + 2z + x = -3$ and $5x - w + 2y - z = 3$.
f. $r - 2g + f - y - 4 = 0$, $2f - 3y + 1 = 3g - 2r$, $3 + 5g + 4y = 3(r + f)$ and $2 - (r + f) = 3 - g - 2y$.

Problem 2. Modify the methods of this section to find the solution set to the system of products below.

$$\begin{aligned} abc &= 1 \\ a^2bc^3 &= 2 \\ bc^2/a &= 8 \end{aligned}$$

Problem 3. Modify the methods of this section to find the solution set to the system of exponential equations below.

$$\begin{aligned} 2^x 3^y 5^z &= 1 \\ 8^x 9^y 5^z &= 1 \\ 4^x 3^y &= 1 \end{aligned}$$

The Gauss-Jordan Algorithm²

In this section we give a formal definition of the **row operations** and present a systematic means to solve systems of linear equations, known as the **Gauss-Jordan Algorithm**.

Definition .4. Given a matrix we define three row operations.

1. Multiply a row by a nonzero constant.
2. Add a multiple of one row to a different row.
3. Switch two rows.

Two $m \times n$ matrices are **row equivalent** if one can be transformed into the other by a finite sequence of row operations. When M and N are row equivalent we may write $M \sim N$.

Theorem .5. *Two systems of linear equations have the same solutions sets if and only if the associated augmented matrices are row equivalent.*

Remark. One direction of the proof is clear since row operations do not change the solution set. The other we postpone until after we have studied vector spaces.

We now give an algorithm, a step by step procedure, that will put any matrix into reduced row echelon form.

Algorithm (Gauss-Jordan Algorithm). Let $A = [a_{ij}]$ be an $m \times n$ matrix. The following steps, repeated as indicated, will produce a matrix in reduced row echelon form that is row equivalent to A .

Step 1: Set i , the row index, and j , the column index, equal to 1.

Step 2: If the j -th column has a nonzero entry with row index at least i go to Step 4.

Step 3: Increase j by 1. If j is now equal to $n + 1$, go to Step 7. Otherwise return to Step 2.

²©Michael C. Sullivan, June 13, 2008

- Step 4: Find the first nonzero entry in the j -th column with row index at least i . Suppose this row index is k . If $k > i$ then switch row i and row k . (Relabel the matrix indices accordingly; the new matrix is still called A .)
- Step 5: Multiply the i -th row by $1/a_{ij}$. (This gives a pivot.)
- Step 6: Use row operation 2 to zero out all the entries in column j below row i . Increase j and i by 1. If $j \leq n$ go to Step 2.
- Step 7: Number the pivots according to the row they are in. Starting with the last and working backwards, use row operation 2, repeatedly, to zero out the column entries above each pivot.
- Step 8: Output the matrix.

We want to apply this to the problem of solving a system of linear equations: given $\mathbf{Ax} = \mathbf{b}$, with \mathbf{A} an $m \times n$ matrix, find all solutions or show that there are none.

Solution: First, recall that n is the number of variables in \mathbf{x} , while m is the number of equations. The rank of \mathbf{A} is the number of equations in the reduced row echelon form system that are not all zero or inconsistent. The equation $\mathbf{Ax} = \mathbf{b}$ can be solved with the following steps.

- Step 1: Set up augmented matrix, $[\mathbf{A} \mid \mathbf{b}]$, and put it into reduced row echelon form with the Gauss-Jordan algorithm. Let $r =$ the rank = number of pivots.
- Step 2: If $r < m$, check for inconsistent equations. If there are any then there are no solutions for \mathbf{x} and so we are done. Otherwise go to Step 3.
- Step 3: If $r < n$, we will get infinitely many solutions. Write them out. The number of free variables = $n - r =$ dimension of solution set and we are done. Otherwise go to Step 4.
- Step 4: It must be that $r = n = m$. Read off the unique solution.

Problem 4. Solve the systems of equations below, using the Gauss-Jordan algorithm. Do not skip any steps; label every step.

$$\begin{array}{l} \text{a. } \begin{bmatrix} 2 & 3 \\ -1 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{b. } \begin{bmatrix} 2 & 3 \\ 16/3 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{c. } \begin{bmatrix} 2 & 1 & 3 \\ 4 & 0 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} \\ \text{d. } \begin{bmatrix} 4 & 0 & 6 \\ 4 & -1 & 1 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} \end{array}$$

Remarks on Equivalence Relations: (Optional Reading). Row equivalence of $m \times n$ matrices is an example of what mathematicians call an **equivalence relation**. We define this notion here. Let S be a set. An equivalence relation on S is a subset R of $S \times S$ that enjoys the following properties (axioms):

- for every $s \in S$, the pair $(s, s) \in R$ (R is reflexive),
- if $(s, t) \in R$, then $(t, s) \in R$ (R is symmetric), and
- if $(r, s) \in R$ and $(s, t) \in R$, then $(r, t) \in R$ (R is transitive).

When $(a, b) \in R$ we often write $a \sim b$, read a is equivalent to b . For example, let $S = \mathbb{Z}$ and say $m \sim n$ if and only if m and n have the same remainder when divided by 7. Thus, $21 \sim 0$, $-5 \sim 47$, $3 \sim 10$, and so on. For each $s \in S$ the **equivalence class** of s , denoted by $[s]$, is defined by $[s] = \{s' \in S \mid s \sim s'\}$. In this example $[3] = \{\dots -11, -4, 3, 10, 17, 24, \dots\}$. There are exactly seven equivalence classes: $[0]$, $[1]$, $[2]$, $[3]$, $[4]$, $[5]$, and $[6]$. It can be shown that equivalence classes never overlap.

Row equivalence is an equivalence relation on the set of matrices of a given size. Equivalence relations arise in many other places in linear algebra and throughout mathematics.

Problem 5. Let S be pairs of integers such that the second member of a pair is not zero; $S = \{(m, n) \mid m, n \in \mathbb{Z}, n \neq 0\}$. Define $(m, n) \sim (m', n')$ if $mn' = m'n$. For example $(2, 3) \sim (8, 12)$. Show this is indeed an equivalence relation. What could this equivalence relation be applied to?