

2 - Solving Linear Equations

Linear algebra developed from studying methods for solving systems of linear equations. We say an equation in the variables $x_1, x_2, x_3, \dots, x_n$ is **linear** if it consists of a linear combination of the x_i equaling a constant.

Example

The following equations are linear:

$$4x - 2y + z = 25$$

$$\frac{1}{2}a + 13.49b + \pi c - d = 0$$

$$\sqrt{17}x_1 + 86^3x_2 + (5.2 - 1/\pi)x_3 = 31$$

Here are some nonlinear equations:

$$x^2 + y = 0$$

$$6a + 3\sqrt{b} + 1/c = \sqrt[3]{47}$$

$$(3x + 9y)(z - 2w) = 73$$

A **system of linear equations** is simply a list or set of linear equations. A **solution** of the system consists of values for the variables which is a solution of each of the equations.

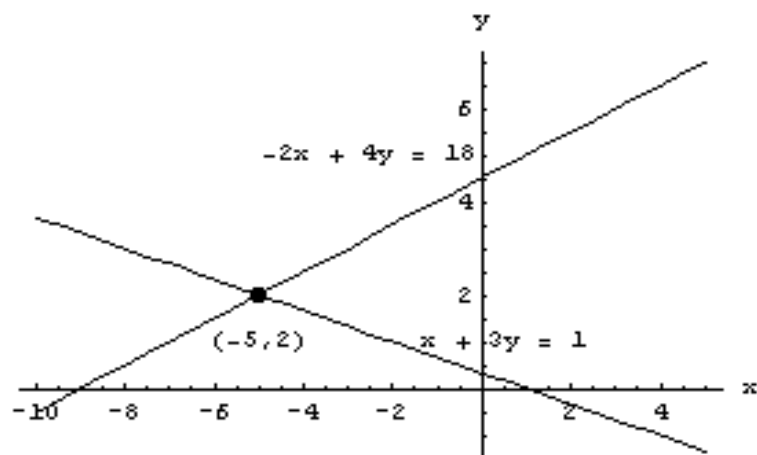
Example

The system $-2x + 4y = 18$

$$x + 3y = 1$$

has as its unique solution $x = -5$ and $y = 2$ or, more succinctly, the ordered pair $(-5, 2)$.

Geometrically, of course, the graph of each equation is a line and the solution of the system, the common solution of each, is the point where the lines intersect as the accompanying graph shows.



Problem

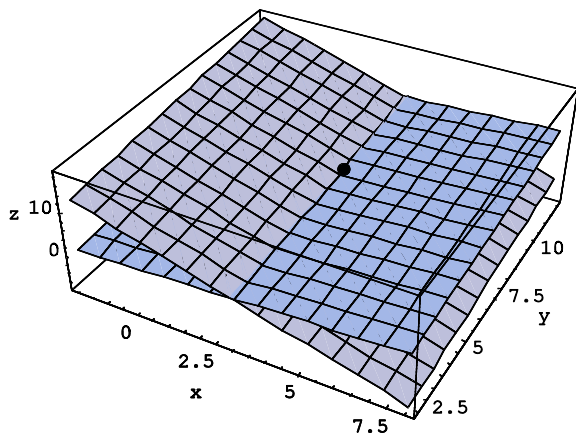
What other possibilities (geometrically) are there for the solution of a linear system with two variables and two equations? What if there are three equations? Or just one?

Example

The following system with three variables $10x - 3y + 5z = 36$

$$7x + 11y - 20z = -11$$

has the point $(3, 8, 6)$ as one solution. Each equation has a plane as its graph as we see below.



Problem

What other solutions does this system have? What other possibilities (geometrically) are there for the solution of a linear system with three variables and two equations? What if there are three equations?

It is convenient to write the general system of linear equations using double subscripts for the coefficients of the variables.

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

The first subscript of each constant indicates the equation in which it is found, while the second tells which variable it multiplies. All of the information of the system is contained in these constants so we have the following definitions.

Definitions

1. A **matrix** is a rectangular array of numbers. The **size** of a matrix with m rows and n columns is $m \times n$ (read m by n .)
2. The **coefficient matrix** of the system above is the matrix, A , whose i, j^{th} entry (i.e., entry in the i^{th} row and j^{th} column) is a_{ij} . As a shorthand we shall write $A = (a_{ij})$.
3. The **augmented matrix** of the system consists of its coefficient matrix along with an additional column, its $(n+1)^{\text{st}}$, whose entries are the constants on the right-hand side of the

equations, so the last column is the vector $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

Example

The coefficient and augmented matrices of the system in the previous example with two variables are

$$\begin{pmatrix} -2 & 4 \\ 1 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} -2 & 4 & 18 \\ 1 & 3 & 1 \end{pmatrix} \text{ respectively.}$$

Note that, for the general system that was given above, the first column of its coefficient

matrix, $\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$, and, in general, its j^{th} column, $\mathbf{c}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$, are vectors. For each i , the i^{th} row

of the coefficient matrix, $\mathbf{r}_i = (a_{i1} \ a_{i2} \ \dots \ a_{in})$, is also a vector and we can write the matrix in several

ways $A = (a_{ij}) = (\mathbf{c}_1 \ \mathbf{c}_2 \ \dots \ \mathbf{c}_n) = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix}$. Using the vector operations introduced in the last chapter, we

can form the linear combination of the columns of A and observe that x_1, x_2, \dots, x_n is a solution of the system if and only if $x_1 \mathbf{c}_1 + x_2 \mathbf{c}_2 + \dots + x_n \mathbf{c}_n = \mathbf{b}$, i.e. \mathbf{b} is a linear combination of the columns of A .

Example

Returning to the 2×2 example (2 equations with 2 unknowns) that we introduced earlier,

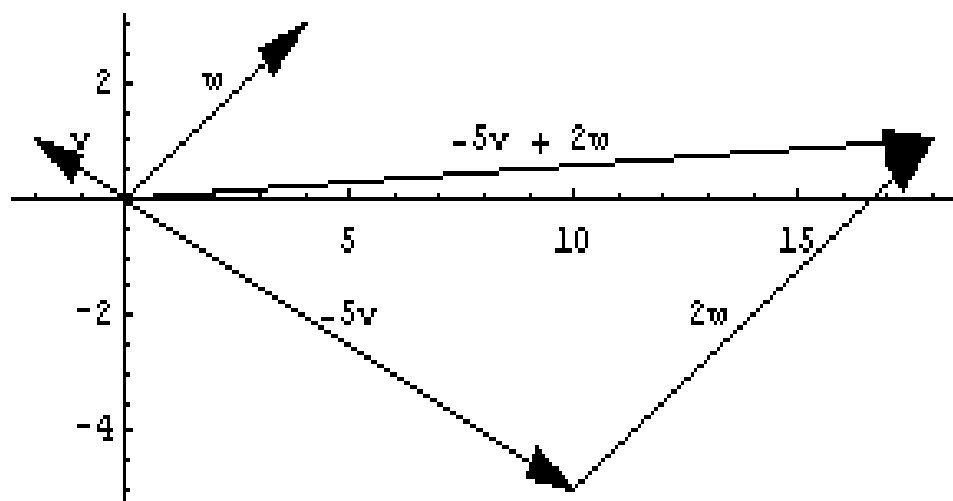
$$-2x + 4y = 18$$

$$x + 3y = 1,$$

we have as its corresponding equation with column vectors

$$x \mathbf{v} + y \mathbf{w} = \mathbf{b} \text{ where } \mathbf{v} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \text{ and } \mathbf{b} = \begin{pmatrix} 18 \\ 1 \end{pmatrix}.$$

Then the solution $(-5, 2)$ can be pictured in the vector diagram that follows.



Thus we have two distinct pictures for systems, both of which have their value and to which we'll return. They are, as Strang refers to them in his text, the row picture (the intersecting lines shown earlier) and the column picture (the linear combination of vectors in the graph above) of the system. The latter picture was a result of viewing the system as a vector equation with the right hand side of the system, \mathbf{b} , written as a linear combination of the column vectors of the coefficient matrix. With the dot product and the row expression of A , we can write the system

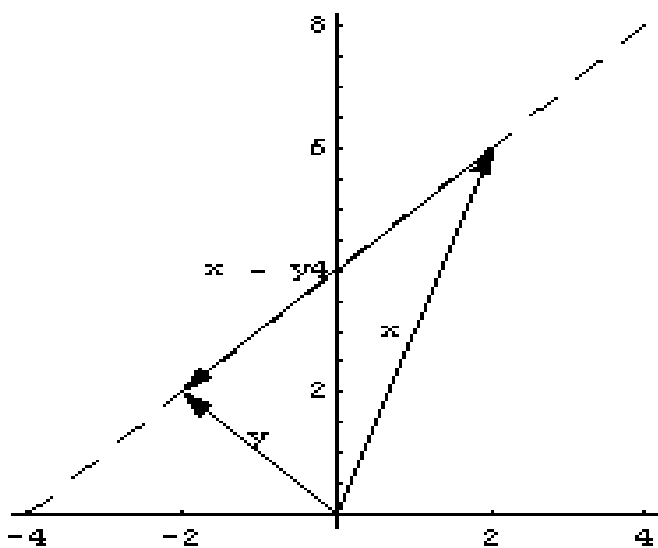
$$\begin{pmatrix} r_1 \cdot \mathbf{x} \\ r_2 \cdot \mathbf{x} \\ \vdots \\ r_m \cdot \mathbf{x} \end{pmatrix} = \mathbf{b} \text{ where } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Let's exploit this view point in considering the possibilities for the system's solutions.

We have seen from the row pictures with intersecting (and parallel) lines and planes that a linear system can have no solution (in which case it is called *inconsistent*) or, if it is *consistent*, can have a single, unique solution or infinitely many solutions. We now show that these are, in fact, the only possibilities. Indeed, suppose that the general system has as distinct solutions the vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$. Let \mathbf{z} be the linear combination of \mathbf{x} and \mathbf{y} given by

$$\mathbf{z} = t \mathbf{x} + (1 - t) \mathbf{y}$$

where t is an arbitrary real number. Consider for any $i = 1, 2, \dots, m$, the i^{th} equation of the system. Since \mathbf{x} and \mathbf{y} are solutions of the system (and, hence, each equation) $\mathbf{r}_i \cdot \mathbf{x} = b_i = \mathbf{r}_i \cdot \mathbf{y}$. Then $\mathbf{r}_i \cdot \mathbf{z} = \mathbf{r}_i \cdot (t\mathbf{x} + (1 - t)\mathbf{y}) = t(\mathbf{r}_i \cdot \mathbf{x}) + (1 - t)(\mathbf{r}_i \cdot \mathbf{y}) = t b_i + (1 - t) b_i = b_i$. Thus \mathbf{z} is a solution of the i^{th} equation for each i and so is a solution of the system. (Note: the second step in the string of equalities uses the result of the problem below.) Consequently any system with two distinct solutions has infinitely many solutions. (Writing $\mathbf{z} = t \mathbf{x} + (1 - t) \mathbf{y} = \mathbf{y} + t(\mathbf{x} - \mathbf{y})$ it is not hard to recognize that the terminal points of \mathbf{z} drawn in standard position (with the origin as its initial point) trace the line through the heads of the vectors \mathbf{x} and \mathbf{y} (in standard position) as the parameter t is varied throughout the real numbers. This is the parametric form of the equation of this line.)



Problem

Show that *the dot product is linear*: for any vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} in \mathbf{R}^n and any scalar c

1. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
2. $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$