Solving Linear Systems

Linear Algebra, edition four

Jim Hefferon

http://joshua.smcvt.edu/linearalgebra

Gauss's Method

Linear systems

1.1 *Definition* A *linear combination* of $x_1, \ldots, x_n$ has the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n$$

where the numbers $a_1, \ldots, a_n \in \mathbb{R}$ are the combination's *coefficients*.

Example This is a linear combination of x, y, and z.

$$(1/4)x + y - z$$

1.1 *Definition* A *linear equation* in the variables $x_1, \ldots, x_n$ has the form $a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = d$ where $d \in \mathbb{R}$ is the *constant*.

An n-tuple $(s_1, s_2, \ldots, s_n) \in \mathbb{R}^n$ is a *solution* of, or *satisfies*, that equation if substituting the numbers $s_1, \ldots, s_n$ for the variables gives a true statement: $a_1s_1 + a_2s_2 + \cdots + a_ns_n = d$. A *system of linear equations*

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= d_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= d_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= d_m \end{aligned}$$

has the solution $(s_1, s_2, \ldots, s_n)$ if that n-tuple is a solution of all of the equations.

Example There are three linear equations in this linear system.

$$\begin{aligned} (1/4)x + y - z &= 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - z &= 3 \end{aligned}$$

Solving a linear system

Example To find the solution of this system

$$\begin{aligned} (1/4)x + y - z &= 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - z &= 3 \end{aligned}$$

we transform it to one whose solution is easy. Start by clearing the fraction.

$$\xrightarrow{4\rho_1} \quad \begin{aligned} x + 4y - 4z &= 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - z &= 3 \end{aligned}$$

Next use the first row to act on the rows below, eliminating their x terms.

$$\xrightarrow[-2\rho_1+\rho_3]{-\rho_1+\rho_2} \quad \begin{aligned} x + 4y - 4z &= 0 \\ +6z &= 12 \\ -11y + 7z &= 3 \end{aligned}$$

Then swap to bring a y term to the second row.

$$\xrightarrow{\rho_2 \leftrightarrow \rho_3} \quad \begin{aligned} x + \quad 4y - 4z &= 0 \\ -11y + 7z &= 3 \\ 6z &= 12 \end{aligned}$$

Now solve the bottom row: $z = 2$. With that, the shape of the transformed system lets us solve for y by substituting into the second row: $-11y + 7(2) = 3$ shows $y = 1$. The shape also lets us solve for x by substituting into the first row: $x + 4(1) - 4(2) = 0$, so that $x = 4$.

1.10 *Definition* In each row of a system, the first variable with a nonzero coefficient is the row's *leading variable*. A system is in *echelon form* if each leading variable is to the right of the leading variable in the row above it, except for the leading variable in the first row, and any rows with all-zero coefficients are at the bottom.

Example

$$\begin{array}{rcl}
2x - 3y - z + 2w &=& -2 \\
x \phantom{{}-3y} + 3z + 1w &=& 6 \\
2x - 3y - z + 3w &=& -3 \\
y + z - 2w &=& 4
\end{array}
\xrightarrow[-\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2}
\begin{array}{rcl}
2x - 3y - z + 2w &=& -2 \\
(3/2)y + (7/2)z &=& 7 \\
w &=& -1 \\
y + z - 2w &=& 4
\end{array}$$

$$\xrightarrow{(-2/3)\rho_2+\rho_4}
\begin{array}{rcl}
2x - 3y - z + 2w &=& -2 \\
(3/2)y + (7/2)z &=& 7 \\
w &=& -1 \\
-(4/3)z - 2w &=& -2/3
\end{array}$$

$$\xrightarrow{\rho_3\leftrightarrow\rho_4}
\begin{array}{rcl}
2x - 3y - z + 2w &=& -2 \\
(3/2)y + (7/2)z &=& 7 \\
-(4/3)z - 2w &=& -2/3 \\
w &=& -1
\end{array}$$

The fourth equation says $w = -1$. Substituting back into the third equation gives $z = 2$. Then back substitution into the second and first rows gives $y = 0$ and $x = 1$. The unique solution is $(1, 0, 2, -1)$.

Gauss's Method

1.5 *Theorem* If a linear system is changed to another by one of these operations

1) an equation is swapped with another
2) an equation has both sides multiplied by a nonzero constant
3) an equation is replaced by the sum of itself and a multiple of another

then the two systems have the same set of solutions.

The proof is in the book.

1.6 *Definition* The three operations from Theorem 1.5 are the *elementary reduction operations*, or *row operations*, or *Gaussian operations*. They are *swapping*, *multiplying by a scalar* (or *rescaling*), and *row combination*.

Systems without a unique solution

Example This system has no solution.

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + z &= 8 \\ 2x + 3y + 2z &= 13 \end{aligned}$$

On the left the sum of the first two rows equals the third row, while on the right that is not so. So there is no triple of reals that makes all three equations true.

Gauss' Method makes the inconsistency clear.

$$\xrightarrow[-2\rho_1+\rho_3]{-\rho_1+\rho_2} \begin{aligned} x + y + z &= 6 \\ y &= 2 \\ y &= 1 \end{aligned} \xrightarrow{-\rho_2+\rho_3} \begin{aligned} x + y + z &= 6 \\ y &= 2 \\ 0 &= -1 \end{aligned}$$

Example This system has infinitely many solutions.

$$\begin{array}{rcl} -x - y + 3z = 3 & & -x - y + 3z = 3 \\ x \phantom{{}- y} + z = 3 & \xrightarrow[3\rho_1+\rho_3]{\rho_1+\rho_2} & -y + 4z = 6 \\ 3x - y + 7z = 15 & & -4y + 16z = 24 \\ & & -x - y + 3z = 3 \\ & \xrightarrow{-4\rho_2+\rho_3} & -y + 4z = 6 \\ & & 0 = 0 \end{array}$$

Taking $z = 0$ gives $(3, -6, 0)$ while taking $z = 1$ gives $(2, -2, 1)$.

Example It is not the '$0 = 0$' that counts. This also has infinitely many solutions.

$$\begin{array}{rcl} x - y + z = 4 & \xrightarrow{-\rho_1+\rho_2} & x - y + z = 4 \\ x + y - 2z = -1 & & 2y - 3z = -5 \end{array}$$

Taking $z = 0$ gives the solution $(3/2, -5/2, 0)$. Taking $z = -1$ gives $(1, -4, -1)$.

Describing the solution set

Parametrizing

We've seen that this system has infinitely many solutions.

$$\begin{array}{rcl} -x-y+3z= & 3 \\ x \qquad + \ \ z= & 3 \\ 3x-y+7z= & 15 \end{array} \xrightarrow[3\rho_1+\rho_3]{-\rho_1+\rho_2} \xrightarrow{-4\rho_2+\rho_3} \begin{array}{rcl} -x- \ \ y+3z=3 \\ -y+4z=6 \\ 0=0 \end{array}$$

We want to describe the solution set.

Use the second row to express y in terms of z as $y=-6+4z$. Now substitute into the first row $-x-(-6+4z)+3z=3$ to express x also in terms of z with $x=3-z$.

This description of the solution set is convenient. Here we pick some z's to show a few of the infinitely many solutions.

z	0	1	2	$-1/2$
solution (x, y, z)	$(3,-6,0)$	$(2,-2,1)$	$(1,2,2)$	$(3.5,-8,-0.5)$

Also, we can tell that $(1,6,3)$ is not a solution without plugging it into the equations because it does not satisfy $x=3-z$.

2.2 *Definition* In an echelon form linear system the variables that are not leading are *free*. A variable that we use to describe a family of solutions is a *parameter*.

We shall routinely parametrize linear systems with the free variables.

Example This system is already in echelon form.

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

The leading variables are x and y so we will parametrize the solution set with z and w. The second row gives $y = -6 + z + 4w$. Substituting into the first row gives $2x + (-6 + z + 4w) + z - w = 5$, so $x = (11/2) - z - (3/2)w$.

Example This is also already in echelon form.

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

We parametrize with y and w. The second row gives $z = 1/4 + (1/2)w$. Substituting back into the first row leaves $x = -(7/8) + (1/2)y + (1/4)w$.

$$\begin{aligned} x &= -(7/8) + (1/2)y + (1/4)w \\ y &= y \\ z &= 1/4 + (1/2)w \\ w &= w \end{aligned}$$

Example

$$\begin{array}{rcl} x - y + 2z + 3w &=& 14 \\ 2x - 2y - z + 2w &=& 6 \\ -3z + 2w &=& 0 \end{array} \xrightarrow{-2\rho_1+\rho_2} \begin{array}{rcl} x - y + 2z + 3w &=& 14 \\ -5z - 4w &=& -22 \\ -3z + 2w &=& 0 \end{array}$$

$$\xrightarrow{-(3/5)\rho_2+\rho_3} \begin{array}{rcl} x - y + 2z + 3w &=& 14 \\ -5z - 4w &=& -22 \\ (22/5)w &=& 66/5 \end{array}$$

The leading variables are x, z, and w. We will parametrize with the free variable y.

The bottom row gives $w = 3$ and substituting that into the next row up gives $z = 2$. The top equation is $x - y + 2 \cdot 2 + 3 \cdot 3 = 14$ so we have $x = 1 + y$.

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

Matrices and vectors

2.6 *Definition* An $m\times n$ *matrix* is a rectangular array of numbers with m *rows* and n *columns*. Each number in the matrix is an *entry*.

Example This is a 2×3 matrix

$$B = \begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \end{pmatrix}$$

because it has 2 rows and 3 columns. The entry in row 2 and column 1 is $b_{2,1} = 4$.

2.8 *Definition* A *column vector*, often just called a *vector*, is a matrix with a single column. A matrix with a single row is a *row vector*. The entries of a vector are sometimes called *components*. A column or row vector whose components are all zeros is a *zero vector*.

We denote vectors with an over-arrow (many authors use boldface).

Example This column vector has three components.

$$\vec{v} = \begin{pmatrix} -1 \\ -0.5 \\ 0 \end{pmatrix}$$

Example This row vector has three components

$$\vec{w} = \begin{pmatrix} -1 & -0.5 & 0 \end{pmatrix}$$

Example This is the two-component zero vector.

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vector operations

2.10 *Definition* The *vector sum* of $\vec{u}$ and $\vec{v}$ is the vector of the sums.

$$\vec{u}+\vec{v}=\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}+\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\begin{pmatrix}u_1+v_1\\ \vdots\\ u_n+v_n\end{pmatrix}$$

2.11 *Definition* The *scalar multiplication* of the real number r and the vector $\vec{v}$ is the vector of the multiples.

$$r\cdot\vec{v}=r\cdot\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}=\begin{pmatrix}rv_1\\ \vdots\\ rv_n\end{pmatrix}$$

Example

$$3\begin{pmatrix}1\\ 2\end{pmatrix}-2\begin{pmatrix}0\\ 1\end{pmatrix}=\begin{pmatrix}3\\ 4\end{pmatrix}$$

Matrix notation for linear systems

Example We can reduce the clerical load in solving this system

$$\begin{aligned} -3x \qquad\quad + 2z &= -1 \\ x - 2y + 2z &= -5/3 \\ -x - 4y + 6z &= -13/3 \end{aligned}$$

by writing it as an *augmented matrix*.

$$\left(\begin{array}{ccc|c} -3 & 0 & 2 & -1 \\ 1 & -2 & 2 & -5/3 \\ -1 & -4 & 6 & -13/3 \end{array}\right) \xrightarrow[-(1/3)\rho_1+\rho_3]{(1/3)\rho_1+\rho_2} \left(\begin{array}{ccc|c} -3 & 0 & 2 & -1 \\ 0 & -2 & 8/3 & -2 \\ 0 & -4 & 16/3 & -4 \end{array}\right)$$

$$\xrightarrow{-2\rho_2+\rho_3} \left(\begin{array}{ccc|c} -3 & 0 & 2 & -1 \\ 0 & -2 & 8/3 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

The two nonzero rows give $-3x + 2z = -1$ and $-2y + (8/3)z = -2$.

Parametrizing $-3x + 2z = -1$ and $-2y + (8/3)z = -2$ gives this.

$$\begin{aligned} x &= (1/3) + (2/3)z \\ y &= 1 + (4/3)z \\ z &= z \end{aligned}$$

We can write the solution set in vector notation.

$$\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2/3 \\ 4/3 \\ 1 \end{pmatrix} z \mid z \in \mathbb{R} \}$$

This description helps us understand the set of solutions. For instance, each value of z gives a different solution.

z	0	1	2	$-1/2$
solution $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$	$\begin{pmatrix} 1/3 \\ 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 7/3 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 5/3 \\ 11/3 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1/3 \\ -1/2 \end{pmatrix}$

Example Reducing this system

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

using the augmented matrix notation

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 2 & -1 & -2 & 1 & 5 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & -5 & 0 & 1 & 1 \end{array}\right)$$

leads to this vector description of the solution set.

$$\{ \begin{pmatrix} 12/5 \\ -1/5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \mid z, w \in \mathbb{R} \}$$

$$\text{General} = \text{Particular} + \text{Homogeneous}$$

Form of solution sets

Example This system

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

has solutions of this form.

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 12/5 \\ -1/5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \qquad z, w \in \mathbb{R}$$

Taking $z = w = 0$ shows that the first vector is a particular solution of the system.

3.2 *Definition* A linear equation is *homogeneous* if it has a constant of zero, so that it can be written as $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$.

Example From the prior system

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

we get this associated system of homogeneous equations.

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

The same Gauss's Method steps reduce it to echelon form.

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 0 \\ 2 & -1 & -2 & 1 & 0 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 0 \\ 0 & -5 & 0 & 1 & 0 \end{array}\right)$$

The vector description of the solution set is like the earlier one but the zero vector is a particular solution.

$$\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \mid z, w \in \mathbb{R}\}$$

3.6 *Lemma* For any homogeneous linear system there exist vectors $\vec{\beta}_1, \ldots,$ $\vec{\beta}_k$ such that the solution set of the system is

$$\{c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

where k is the number of free variables in an echelon form version of the system.

Example The book has the proof. For the main idea consider this system of homogeneous equations.

$$\begin{aligned} x + y + z + w &= 0 \\ y - z + w &= 0 \end{aligned}$$

Using the bottom equation, express the leading variable y in terms of the free variables $y = z - w$. Next substitute $x + (z - w) + z + w = 0$ and solve for the leading variable $x = -2z$. Finish by describing the solution in vector notation.

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = z\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + w\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \qquad z, w \in \mathbb{R}$$

and recognize the two vectors as the lemma's $\vec{\beta}_1$ and $\vec{\beta}_2$.

3.7 *Lemma* For a linear system and for any particular solution $\vec{p}$, the solution set equals $\{\vec{p}+\vec{h} \mid \vec{h}$ satisfies the associated homogeneous system$\}$.

3.7 *Proof* For set inclusion the first way, that if a vector solves the system then it is in the set described above, assume that $\vec{s}$ solves the system. Then $\vec{s}-\vec{p}$ solves the associated homogeneous system since for each equation index i,

$$\begin{aligned} a_{i,1}(s_1-p_1)+\cdots+a_{i,n}(s_n-p_n) \\ =(a_{i,1}s_1+\cdots+a_{i,n}s_n)-(a_{i,1}p_1+\cdots+a_{i,n}p_n)=d_i-d_i=0 \end{aligned}$$

where p_j and s_j are the j-th components of $\vec{p}$ and $\vec{s}$. Express $\vec{s}$ in the required $\vec{p}+\vec{h}$ form by writing $\vec{s}-\vec{p}$ as $\vec{h}$.

For set inclusion the other way, take a vector of the form $\vec{p}+\vec{h}$, where $\vec{p}$ solves the system and $\vec{h}$ solves the associated homogeneous system and note that $\vec{p}+\vec{h}$ solves the given system since for any equation index i,

$$\begin{aligned} a_{i,1}(p_1+h_1)+\cdots+a_{i,n}(p_n+h_n) \\ &= (a_{i,1}p_1+\cdots+a_{i,n}p_n)+(a_{i,1}h_1+\cdots+a_{i,n}h_n) = d_i+0 = d_i \end{aligned}$$

where as earlier p_j and h_j are the j-th components of $\vec{p}$ and $\vec{h}$. QED

3.1 *Theorem* Any linear system's solution set has the form

$$\{\vec{p} + c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

where $\vec{p}$ is any particular solution and where the number of vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ equals the number of free variables that the system has after a Gaussian reduction.

Proof This restates the prior two lemmas. QED

3.10 *Corollary* Solution sets of linear systems are either empty, have one element, or have infinitely many elements.

Summary: Kinds of Solution Sets

		number of solutions of the homogeneous system	
		one	*infinitely many*
particular solution exists?	*yes*	unique solution	infinitely many solutions
	no	no solutions	no solutions

An important special case is when there are the same number of equations as unknowns.

3.11 *Definition* A square matrix is *nonsingular* if it is the matrix of coefficients of a homogeneous system with a unique solution. It is *singular* otherwise, that is, if it is the matrix of coefficients of a homogeneous system with infinitely many solutions.