Introducing Linear Algebra

Systems of Real Linear Equations

A. Havens

Department of Mathematics
University of Massachusetts, Amherst

January 22, 2018

Outline

A Geometric Problem in Two Dimensions

- Suppose we have two lines in the plane, and wish to find out if they intersect, and if so, we wish to locate the point of intersection.
- E.g. we might wish to find the intersection point, if it exists, for the pair of lines whose equations in "standard form" are given as

$$\begin{aligned} 2x + 4y &= 6, \\ x - y &= 0\,. \end{aligned}$$

- We want a general algebraic procedure to solve problems such as this, generalized to many variables.

A Geometric Problem in Two Dimensions

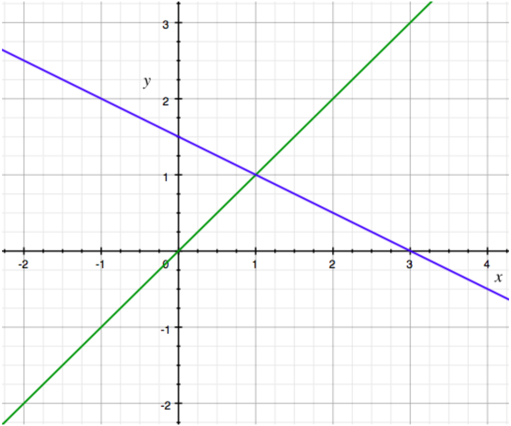

Figure: The two lines plotted in the Cartesian coordinate plane.

A Geometric Problem in Three Dimensions

- In three dimensions, one can show that a plane has an equation of the form

$$ax + by + cz = d\,,$$

where $a, b, c, d \in \mathbb{R}$ are constant coefficients, and x, y, and z are real coordinate variables on $\mathbb{R}^3$.

- In analogy to the preceding intersection problem for lines, we can consider an intersection problem for planes:

$$\begin{cases} x + y + z = 6 \\ x - 2y + 3z = 6 \\ 4x - 5y + 6z = 12 \end{cases},$$

A Geometric Problem in Three Dimensions

- Here's a picture of these planes in 3D space:

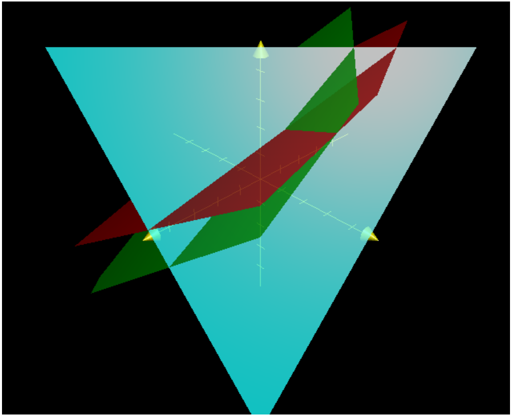

Figure: Note: two planes that intersect meet in a line, and the three lines of intersection meet in a point which is common to all three

Definition of a system of real linear equations

- We are now ready to define a general real linear system of equations. Let $m, n \geq 2$ be positive integers. If we are given mn real numbers a_{ij}, $1 \leq i \leq m$ and $1 \leq j \leq n$, and m constants b_i, then we can form m equations in n unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \qquad\qquad \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \end{cases}.$$

- This is called an $m \times n$ *real linear system of equations.*
- We will describe methods for producing general solutions, and study the existence and uniqueness of solutions.

Applications

There are too many applications to provide a comprehensive list.

- Within mathematics and statistics, linear systems occur in many contexts, such as solving differential equations, finding equations of certain curves and surfaces from sample points or finding curves/surfaces to fit data, or studying geometric transformations of mathematical spaces, such as rotations and reflections.
- Many of the above applications occur when modeling real physical systems, biological systems, chemical reactions, electrical circuits...
- Search engines such as Google use linear algebra to create page rankings.
- Recursive neural nets in machine learning require solving linear systems to obtain weights.

Memories of High School Algebra

- Writing out equations many times while performing manipulations to solve them can be clunky and tedious, so we will introduce a book-keeping device that turns out to have deeper, more practical uses than mere notational convenience.
- For subsequent discussion, we often consider an "abstract" $m \times n$ system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \end{cases} .$$

Defining real matrices

- A real $m \times n$ matrix A is an array of numbers with m rows and n columns:

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

- If the entries of the matrix arise from a linear system such as the one above, we say it is the *coefficient matrix* for that linear system. To refer to the entry of a matrix A in the ith row and jth column we write a_{ij}. A convenient shorthand for the whole matrix is $\mathrm{A} = (a_{ij})$.
- We can also encode the constants as a matrix, called a *column vector*. More on this in a moment.

- E.g. For the system

$$\begin{cases} 2x + 4y & = 6 \\ x - y & = 0 \end{cases}.$$

the coefficient matrix is

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

- What is the coefficient matrix for the system

$$\begin{cases} x + y + z & = 6 \\ x - 2y + 3z & = 6 \\ 4x - 5y + 6z & = 12 \end{cases} ?$$

- Write your name, preferred email, and the answer to this on a piece of paper to hand in later.

Constants and column vectors

- We can collect constants $b_1, \ldots, b_m$ into an $m \times 1$ matrix called a *column vector*:

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

- What are the constant column vectors associated to the above systems with 2 and 3 variables? For the 3 variable system, write it down below your answer for the coefficient matrix.

Augmented Matrix of a System

- To describe a whole linear system, we need only the coefficient matrix and the column vector of constants. This is encoded together in what is called an *augmented matrix*:

$$\left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right].$$

- Given a matrix A and a column vector $\mathbf{b}$, the augmented matrix obtained from appending $\mathbf{b}$ to A is often denoted $\left[\mathrm{A} \mid \mathbf{b}\right]$.

- For the two variable system

$$\begin{cases} 2x + 4y & = 6 \\ x - y & = 0 \end{cases}$$

the augmented matrix is

$$\left[\begin{array}{cc|c} 2 & 4 & 6 \\ 1 & -1 & 0 \end{array}\right].$$

- Write down the systems associated to the following augmented matrices:

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 0 \\ -2 & 4 & 6 & -3 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Solution by substitution

- We can solve the system

$$\begin{cases} 2x + 4y &= 6 \\ x - y &= 0 \end{cases}$$

by first solving one equation for a variable, such as y, and then substituting into the other equation.
- The equation $x - y = 0$ tells us $y = x$, and so substituting into the first equation, we have $2x + 4x = 6x = 6 \implies x = 1$.
- *Back-substituting*, we get $y = x = 1$, so the point $(1, 1)$ is the solution.

The Method of Elimination

- With more variables, the substitution approach becomes much more tedious.
- Another method is to use *elimination*: one can add a multiple of one equation to another to try to reduce the number of variables.
- After eliminating sufficiently many variables, one obtains a single equation in a single variable. Solving and then back-substituting, one can hope to obtain the solution.
- What if we can't eliminate all of the variables?
- We will see that some systems have no solution at all, or have infinitely many solutions, expressed in terms of *free variables*, arising from those which one fails to eliminate.

Elimination Example

In the system

$$\begin{cases} x + y + z &= 6 \\ x - 2y + 3z &= 6 \\ 4x - 5y + 6z &= 12 \end{cases}$$

if we subtract the first equation from the second equation, and replace the second equation with the result, we obtain

$$\begin{cases} x + y + z &= 6 \\ 0 - 3y + 2z &= 0 \\ 4x - 5y + 6z &= 12 \end{cases}$$

Elimination Example

- Subtracting 4 times the first equation from the third equation and replacing the third equation with the result yields

$$\begin{cases} x + \ y + \ z &= 6 \\ 0 - \ 3y + 2z &= 0 \\ 0 \ - 9y + 2z &= -12 \end{cases}$$

- The last two equations now only involve the unknowns y and z, so we can solve for these more easily.
- What should we do to reduce to one variable by elimination with these two equations?

Elimination Example with matrices

- To make the process less cumbersome, we can use augmented matrices instead.
- The system can be rewritten

$$\begin{cases} x + \ \ y + \ \ z &= 6 \\ x - \ 2y + 3z &= 6 \\ 4x - 5y + 6z &= 12 \end{cases} \longleftrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & -2 & 3 & 6 \\ 4 & -5 & 6 & 12 \end{array}\right].$$

- The steps performed above correspond to *row operations*.

Elimination Example with matrices

- We can describe the row operations already performed symbolically:

$$\text{first}\ \ R_2 - R_1 \mapsto R_2\,,\ \ \text{then}\ \ R_3 - 4R_1 \mapsto R_3\,.$$

- To find a unique solution, we'd need to manipulate the rows, as we would equations, until our matrix has the form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & p \\ 0 & 1 & 0 & q \\ 0 & 0 & 1 & r \end{array}\right]\,,$$

  which corresponds to a solution $(x, y, z) = (p, q, r)$ for some $p, q, r \in \mathbb{R}$.

- A system in this form is said to be in *reduced row echelon form*, or RREF.

Row Operations

There are three kinds of elementary operations on rows used in reducing. Before we complete the solution of the above system, let's list these operations. Consider a general real linear system described by some augmented matrix $\left[\mathrm{A} \mid \mathbf{b}\right]$.

1. We may swap two rows, just as we may write the equations in any order we please. We notate a swap of the ith and jth rows of an augmented matrix by $R_i \leftrightarrow R_j$.
2. We may replace a row R_i with the row obtained by scaling the original row by a *nonzero* real number. We notate this by $sR_i \mapsto R_i$.
3. We may replace a row R_i by the difference of that row and a multiple of another row. We notate this by $R_i - sR_j \mapsto R_i$.

Finishing the example (sort-of)

- One should check (exercise!) that the following sequence of operations (including the initial two described above) reduces the system to RREF:

1.) $R_2 - R_1 \mapsto R_2$,
2.) $R_3 - 4R_1 \mapsto R_3$,
3.) $-\frac{1}{3}R_2 \mapsto R_2$,
4.) $R_3 + 9R_2 \mapsto R_3$,
5.) $-\frac{1}{4}R_3 \mapsto R_3$,
6.) $R_2 + \frac{2}{3}R_3 \mapsto R_2$,
7.) $R_1 - R_3 \mapsto R_1$,
8.) and $R_1 - R_2 \mapsto R_1$.

- After performing in sequence the moves , the matrix reduces to

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

- Thus the solution to our system is $(1, 2, 3)$, which is the point where these planes intersect.
- Note that the process used by these row operations gradually transforms all entries below the *main diagonal* into zeros, before back-substituting.
- Thus, we focus on one column at a time, using one entry to transform every entry below it into a zero.
- This process is called "pivoting down", and the nonzero entry used to accomplish this is called a *pivot*. Usually, the rows are scaled to make the pivots equal to 1.
- It may happen that there are some zeros above the pivot. Sometimes you can swap rows to collect such zeroes below the pivot, so long as you don't undo previous work.

- Once all the pivoting down was complete, pivots were used to zero out the entries above them. This is called *pivoting up*.
- You should convince yourself that pivoting up is equivalent algebraically to back substituting.
- Next time we will explicitly describe how to use the row operations to solve a real linear system given by an augmented matrix.
- You can guess that the method involves locating pivots, pivoting-down, then pivoting-up/back-substituting.
- Even with all integer entries, fractions will often appear in the calculation, whether or not the solution can be expressed as a collection of integers. However, there is a way to modify the algorithm to avoid fractions until the very end, provided the matrix has only integer entries.

Any Two Lines. . .

- Let us apply row operations to attempt to solve the abstract system
$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases} \longleftrightarrow \left[\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right].$$
- We assume temporarily that $a \neq 0$. We will discuss this assumption in more depth later.
- Since our goal is to make the coefficient matrix have ones along the diagonal from left top to right bottom, and zeros elsewhere, we work to first zero out the bottom left entry.
- This can be done, for example, by taking a times the second row and subtracting c times the first row, and replacing the second row with the result.

Any Two Lines. . .

- Thus, applying $aR_2 - cR_1 \mapsto R_2$:

$$\left[\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right] \longmapsto \left[\begin{array}{cc|c} a & b & e \\ 0 & ad-bc & af-ce \end{array}\right].$$

- We see that if $ad - bc = 0$, then either there is no solution, or we must have $af - ce = 0$.
- Let's plug on assuming that $ad - bc \neq 0$. We may eliminate the upper right position held by b in the coefficient matrix by $(ad - bc)R_1 - bR_2 \mapsto R_1$, yielding

Any Two Lines. . .

- We may eliminate the upper right position held by b in the coefficient matrix by $(ad-bc)R_1 - bR_2 \mapsto R_1$, yielding

$$\left[\begin{array}{cc|c} a(ad-bc) & 0 & (ad-bc)e - b(af-ce) \\ 0 & ad-bc & af-ce \end{array}\right]$$

$$= \left[\begin{array}{cc|c} a(ad-bc) & 0 & ade-abf \\ 0 & ad-bc & af-ce \end{array}\right].$$

Any Two Lines...

- Since we assumed a and $ad-bc$ nonzero, we may apply the final row operations $\frac{1}{a(ad-bc)}R_1 \mapsto R_1$ and $\frac{1}{ad-bc}R_2 \mapsto R_2$ to obtain
$$\left[\begin{array}{cc|c} 1 & 0 & (de-bf)/(ad-bc) \\ 0 & 1 & (af-ce)/(ad-bc) \end{array}\right],$$
so we obtain the solution as
$$x = \frac{de-bf}{ad-bc}, \quad y = \frac{af-ce}{ad-bc}.$$
- About that assumption, $a \neq 0 \ldots$

Any Two Lines. . .

- Note that if $a = 0$ but $bc \neq 0$, the solutions are still well defined.
- One can obtain the corresponding expressions with $a = 0$ substituted in by instead performing elimination on

$$\left[\begin{array}{cc|c} 0 & b & e \\ c & d & f \end{array}\right],$$

where the first step might be a simple row swap.

Any Two Lines. . .

- However, if $ad - bc = 0$, there is no hope for the unique solution expressions we obtained, though there may still be solutions, or there may be none at all.
- How do we characterize this failure geometrically?
- A solution is unique precisely when the two lines $ax + by = e$ and $cx + dy = f$ have distinct slopes, and thus intersect in a unique point. One can show that $ad - bc$ measures whether the slopes are distinct!
- If $ad - bc = 0$, there could be no solutions at all (two distinct parallel lines) or infinitely many solutions!

Existence and Uniqueness of Solutions for Two-Dimensional Systems

Proposition

For a given two variable linear system described by the equations

$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases}$$

the quantity $ad - bc = 0$ *if and only if the lines described by the equations have the same slope.*

Corollary

There exists a unique solution to the system above if and only if $ad - bc$ *is nonzero.*

Consistency

- If no solution exists, the system is said to be *inconsistent*. Otherwise, it is said to be a consistent system.
- Can we easily characterize when a two dimensional system has infinitely many solutions, just from the terms of the augmented matrix?
- We'll come back to this, and prove the proposition and its corollary in time.

What About Planes?

- Will three nonparallel planes always intersect in a unique point?
- For two nonparallel planes, the geometric intersection is a line. We don't get unique solutions in this case, but we can still use row reduction to describe the line.
- Try to draw a picture of an inconsistent system of planes.

The Big Questions (for now)

The general questions we want to ask about solutions to a linear system are:

- Existence: does *some* solution exist, i.e. is there some point satisfying all the equations? Equivalently, is the system consistent?
- Uniqueness: If a solution exists is it the only one?

Homework for Week 1

- Go to http://people.math.umass.edu/~havens/math235-4.html, review the policy and expectations for this section and the overall course.
- Read the course overview on the section website.
- Use the course ID to log into MyMathLab.
- Please read sections 1.1-1.3 of the textbook for Friday.