Contents

Gaussian Elimination

Objectives:

1. Given a multivariable problem, write it as a linear system of equations (when this is possible).
2. Express a linear system of equations as a single equation involving matrices and vectors.
3. Write matrix equations in the shorthand augmented matrix notation.
4. Understand that different augmented matrices can correspond to the same solution(s).
5. Know which row operations leave the solutions of an augmented matrix unchanged.
6. Understand the reduced row echelon form of an augmented matrix and how to extract solutions from it.
7. Gaussian elimination: systematically apply row operations to an augmented matrix to obtain reduced row echelon form.

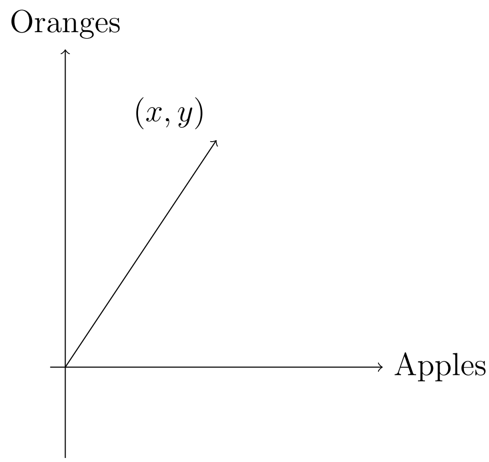
1 Linear Systems

1.1 A Simple Linear System

Reading Guide

Example Suppose I have a bunch of apples and oranges. Let x be the number of apples and y be the number of oranges. As everyone knows, apples and oranges don't mix, so if I want to keep track of the number of apples and oranges I have, I should put them in a list. We'll call this list a *vector,* and write it like this: (x, y). The order here matters! I should remember to always write the number of apples first and then the number of oranges–otherwise if I see the vector $(1, 2)$, I won't know whether I have two apples or two oranges.

The vector (x, y) in the example is just a list of two numbers, so if we want to, we can represent it as a point in the plane with the corresponding coordinates, like so:



In the plane, we can imagine each point as some combination of apples and oranges (or parts thereof, for the points that don't have integer coordinates). So each point corresponds to some vector. The collection of all such vectors—all the points in our apple-orange plane—is an example of a *vector space*.

Example There are 27 pieces of fruit in a barrel, and twice as many oranges as apples. How many apples and oranges are in the barrel?

How to solve this conundrum? We can re-write the question mathematically as follows:

$$\begin{aligned} x + y &= 27 \\ y &= 2x \end{aligned}$$

This is an example of a *Linear System*. It's a collection of equations in which variables are multiplied by constants and summed, and no variables are multiplied together: There are no powers of x or y greater than one, no fractional or negative powers of x or y, and no places where x and y are multiplied together.

Reading homework: problem 1.1

Notice that we can solve the system by manipulating the equations involved. First, notice that the second equation is the same as $-2x + y = 0$. Then if you subtract the second equation from the first, you get on the left side $x + y - (-2x + y) = 3x$, and on the left side you get $27 - 0 = 27$. Then

$3x = 27$, so we learn that $x = 9$. Using the second equation, we then see that $y = 18$. Then there are 9 apples and 18 oranges.

Let's do it again, by working with the list of equations as an object in itself. First we rewrite the equations tidily:

$$\begin{aligned} x + y &= 27 \\ 2x - y &= 0 \end{aligned}$$

We can express this set of equations with a matrix as follows:

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 27 \\ 0 \end{pmatrix}$$

The square list of numbers is an example of a *matrix*. We can multiply the matrix by the vector to get back the linear system using the following rule for multiplying matrices by vectors:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \tag{1}$$

 Reading homework: problem 1.2

The matrix is an example of a *Linear Transformation*, because it takes one vector and turns it into another in a "linear" way. Of course, we can have much larger matrices if our system has more variables:

`A` 3×3 `matrix example`

Our next task is to solve linear systems. We'll learn a general method called Gaussian Elimination.

2 Gaussian Elimination

2.1 Notation for Linear Systems

We just—thanks to the oranges and apples—studied the linear system

$$\begin{array}{rcrcr} x & + & y & = & 27 \\ 2x & - & y & = & 0 \end{array}$$

and found that

$$\begin{array}{rcr} x & = & 9 \\ y & = & 18\,. \end{array}$$

We learned to write the linear system using a matrix and two vectors:

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 27 \\ 0 \end{pmatrix}$$

Likewise, we can write the solution as:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 9 \\ 18 \end{pmatrix}$$

The matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is called the *Identity Matrix*. You should check that if v is any vector, then

$$Iv = v\,.$$

A useful shorthand for a linear system is an *Augmented Matrix*, which looks like this for the linear system we've been dealing with:

$$\left(\begin{array}{cc|c} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right)$$

We don't bother writing the vector $\begin{pmatrix} x \\ y \end{pmatrix}$, since it will show up in any linear system we deal with. The solution to the linear system looks like this:

$$\left(\begin{array}{cc|c} 1 & 0 & 9 \\ 0 & 1 & 18 \end{array}\right)$$

Augmented Matrix Notation

Here's another example of an augmented matrix, for a linear system with three equations and four unknowns:

$$\left(\begin{array}{cccc|c} 1 & 3 & 2 & 0 & 9 \\ 6 & 2 & 0 & -2 & 0 \\ -1 & 0 & 1 & 1 & 3 \end{array}\right)$$

And finally, here's the general case. The number of equations in the linear system is the number of rows r in the augmented matrix, and the number of columns k in the matrix left of the vertical line is the number of unknowns.

$$\left(\begin{array}{cccc|c} a^1_1 & a^1_2 & \cdots & a^1_k & b^1 \\ a^2_1 & a^2_2 & \cdots & a^2_k & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a^r_1 & a^r_2 & \cdots & a^r_k & b^r \end{array}\right)$$

Reading homework: problem 2.1

Here's the idea: Gaussian Elimination is a set of rules for taking a general augmented matrix and turning it into a very simple augmented matrix consisting of the identity matrix on the left and a bunch of numbers (the solution) on the right.

Equivalence Relations for Linear Systems

Equivalence Example

It often happens that two mathematical objects will appear to be different but in fact are exactly the same. The best-known example of this are fractions. For example, the fractions $\frac{1}{2}$ and $\frac{6}{12}$ describe the same number. We could certainly call the two fractions *equivalent*.

In our running example, we've noticed that the two augmented matrices

$$\left(\begin{array}{cc|c} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right), \qquad \left(\begin{array}{cc|c} 1 & 0 & 9 \\ 0 & 1 & 18 \end{array}\right)$$

both contain the same information: $x = 9, y = 18$.

Two augmented matrices corresponding to linear systems *that actually have solutions* are said to be (row) *equivalent* if they have the *same* solutions. To denote this, we write:

$$\left(\begin{array}{cc|c} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 9 \\ 0 & 1 & 18 \end{array}\right)$$

The symbol $\sim$ is read "is equivalent to".

A small excursion into the philosophy of mathematical notation: Suppose I have a large pile of equivalent fractions, such as $\frac{2}{4}$, $\frac{27}{54}$, $\frac{100}{200}$, and so on. Most people will agree that their favorite way to write the number represented by all these different factors is $\frac{1}{2}$, in which the numerator and denominator are relatively prime. We usually call this a *reduced fraction*. This is an example of a *canonical form*, which is an extremely impressive way of saying "favorite way of writing it down". There's a theorem telling us that every rational number can be specified by a unique fraction whose numerator and denominator are relatively prime. To say that again, but slower, *every* rational number *has* a reduced fraction, and furthermore, that reduced fraction is *unique*.

`A 3 × 3 example`

2.2 Reduced Row Echelon Form

Since there are many different augmented matrices that have the same set of solutions, we should find a canonical form for writing our augmented matrices. This canonical form is called *Reduced Row Echelon Form*, or RREF for short. RREF looks like this in general:

$$\left(\begin{array}{cccccccc|c} 1 & * & 0 & * & 0 & \cdots & & 0 & b^1 \\ 0 & & 1 & * & 0 & \cdots & & 0 & b^2 \\ 0 & & 0 & & 1 & \cdots & & 0 & b^3 \\ \vdots & & \vdots & & \vdots & & & 0 & \vdots \\ & & & & & & & 1 & b^k \\ 0 & & 0 & & 0 & \cdots & & 0 & 0 \\ \vdots & & \vdots & & \vdots & & & \vdots & \vdots \\ 0 & & 0 & & 0 & \cdots & & 0 & 0 \end{array}\right)$$

The first non-zero entry in each row is called the *pivot*. The asterisks denote arbitrary content which could be several columns long. The following properties describe the RREF.

1. In RREF, the pivot of any row is always 1.
2. The pivot of any given row is always to the right of the pivot of the row above it.
3. The pivot is the only non-zero entry in its column.

Example $\left(\begin{array}{ccc|c} 1 & 0 & 7 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$

Here is a NON-Example, which breaks all three of the rules:

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

The RREF is a very useful way to write linear systems: it makes it very easy to write down the solutions to the system.

Example

$$\left(\begin{array}{cccc|c} 1 & 0 & 7 & 0 & 4 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

When we write this augmented matrix as a system of linear equations, we get the following:

$$\begin{aligned} x \quad + 7z \quad &= 4 \\ y + 3z \quad &= 1 \\ w &= 2 \end{aligned}$$

Solving from the bottom variables up, we see that $w = 2$ immediately. z is not a pivot, so it is still undetermined. Set $z = \lambda$. Then $y = 1 - 3\lambda$ and $x = 4 - 7\lambda$. More concisely:

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -7 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

So we can read off the solution *set* directly from the RREF. (Notice that we use the word "set" because there is not just one solution, but one for every choice of λ.)

Reading homework: problem 2.2

You need to become very adept at reading off solutions of linear systems from the RREF of their augmented matrix. The general method is to work from the bottom up and set any non-pivot variables to unknowns. Here is another example.

Example

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Here we were not told the names of the variables, so lets just call them x_1, x_2, x_3, x_4, x_5. (There are always as many of these as there are columns in the matrix before the vertical line; the number of rows, on the other hand is the number of linear equations.)

To begin with we immediately notice that there are no pivots in the second and fourth columns so x_2 and x_4 are undetermined and we set them to

$$x_2 = \lambda_1 , \qquad x_4 = \lambda_2 .$$

(Note that you get to be creative here, we could have used λ and μ or any other names we like for a pair of unknowns.)

Working from the bottom up we see that the last row just says $0 = 0$, a well known fact! *Note that a row of zeros save for a non-zero entry after the vertical line would be mathematically inconsistent and indicates that the system has NO solutions at all.*

Next we see from the second last row that $x_5 = 3$. The second row says $x_3 = 2 - 2x_4 = 2 - 2\lambda_2$. The top row then gives $x_1 = 1 - x_2 - x_4 = 1 - \lambda_1 - \lambda_2$. Again we can write this solution as a vector

$$\begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}.$$

Observe, that since no variables were given at the beginning, we do not really need to state them in our solution. As a challenge, look carefully at this solution and make sure you can see how every part of it comes from the original augmented matrix without every having to reintroduce variables and equations.

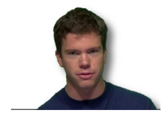
Perhaps unsurprisingly in light of the previous discussions of RREF, we have a theorem:

Theorem 2.1. *Every augmented matrix is row-equivalent to a* unique *augmented matrix in reduced row echelon form.*

Later you will explore why this is true.

3 Elementary Row Operations

Reading Guide



Our goal is to begin with an arbitrary matrix and apply operations that respect row equivalence until we have a matrix in Reduced Row Echelon Form (RREF). The three elementary row operations are:

- (Row Swap) Exchange any two rows.
- (Scalar Multiplication) Multiply any row by a non-zero constant.
- (Row Sum) Add a multiple of one row to another row.

Example

Why do these preserve the linear system in question? Swapping rows is just changing the order of the equations begin considered, which certainly should not alter the solutions. Scalar multiplication is just multiplying the equation by the same number on both sides, which does not change the solution(s) of the equation. Likewise, if two equations share a common solution, adding one to the other preserves the solution. Therefore we can define augmented matrices to be row equivalent if the are related by a sequence of elementary row operations. This definition can also be applied to augmented matrices corresponding to linear systems with no solutions at all!

There is a very simple process for row-reducing a matrix, working column by column. This process is called *Gauss–Jordan elimination* or simply Gaussian elimination.

1. If all entries in a given column are zero, then the associated variable is undetermined; make a note of the undetermined variable(s) and then ignore all such columns.

2. Swap rows so that the first entry in the first column is non-zero.

3. Multiply the first row by λ so that this pivot entry is 1.

4. Add multiples of the first row to each other row so that the first entry of every other row is zero.

5. Before moving on to step 6, add multiples of the first row any rows above that you have ignored to ensure there are zeros in the column above the current pivot entry.

6. Now ignore the first row and first column and repeat steps 2-5 until the matrix is in RREF.

Reading homework: problem 3.1

Reading Guide

Example

$$\begin{array}{rrrl} & & 3x_3 &= 9 \\ x_1 & +5x_2 & -2x_3 &= 2 \\ \frac{1}{3}x_1 & +2x_2 & &= 3 \end{array}$$

First we write the system as an augmented matrix:

$$\left(\begin{array}{ccc|c} 0 & 0 & 3 & 9 \\ 1 & 5 & -2 & 2 \\ \frac{1}{3} & 2 & 0 & 3 \end{array}\right) \quad \underset{\sim}{R_1 \leftrightarrow R_3} \quad \left(\begin{array}{ccc|c} \frac{1}{3} & 2 & 0 & 3 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

$$\underset{\sim}{3R_1} \quad \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 1 & 5 & -2 & 2 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

$$\underset{\sim}{R_2 = R_2 - R_1} \quad \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

$$\underset{\sim}{-R_2} \quad \left(\begin{array}{ccc|c} 1 & 6 & 0 & 9 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

$$\underset{\sim}{R_1 = R_1 - 6R_2} \quad \left(\begin{array}{ccc|c} 1 & 0 & -12 & -33 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 3 & 9 \end{array}\right)$$

$$\underset{\sim}{\frac{1}{3}R_3} \quad \left(\begin{array}{ccc|c} 1 & 0 & -12 & -33 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\underset{\sim}{R_1 = R_1 + 12R_3} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\underset{\sim}{R_2 = R_2 - 2R_3} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Now we're in RREF and can see that the solution to the system is given by $x_1 = 3$, $x_2 = 1$, and $x_3 = 3$; it happens to be a unique solution. Notice that we kept track of the steps we were taking; this is important for checking your work!

Example

$$
\left(\begin{array}{cccc|c}
1 & 0 & -1 & 2 & -1 \\
1 & 1 & 1 & -1 & 2 \\
0 & -1 & -2 & 3 & -3 \\
5 & 2 & -1 & 4 & 1
\end{array}\right)
$$

$$
\overset{R_2-R_1;R_4-5R_1}{\sim}
\left(\begin{array}{cccc|c}
1 & 0 & -1 & 2 & -1 \\
0 & 1 & 2 & -3 & 3 \\
0 & -1 & -2 & 3 & -3 \\
0 & 2 & 4 & -6 & 6
\end{array}\right)
$$

$$
\overset{R_3+R_2;R_4-2R_3}{\sim}
\left(\begin{array}{cccc|c}
1 & 0 & -1 & 2 & -1 \\
0 & 1 & 2 & -3 & 3 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{array}\right)
$$

Here the variables x_3 and x_4 are undetermined; the solution is not unique. Set $x_3 = \lambda$ and $x_4 = \mu$ where λ and μ are *arbitrary* real numbers. Then we can write x_1 and x_2 in terms of λ and μ as follows:

$$
\begin{aligned}
x_1 &= \lambda - 2\mu - 1 \\
x_2 &= -2\lambda + 3\mu + 3
\end{aligned}
$$

We can write the solution set with vectors like so:

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 3 \\ 0 \\ 1 \end{pmatrix}
$$

This is (almost) our preferred form for writing the set of solutions for a linear system with many solutions.

`Worked examples of Gaussian elimination`

Uniqueness of Gauss-Jordan Elimination

Theorem 3.1. *Gauss-Jordan Elimination produces a unique augmented matrix in RREF.*

Proof. Suppose Alice and Bob compute the RREF for a linear system but get different results, A and B. Working from the left, discard all columns except for the pivots and the first column in which A and B differ. By Review Problem 1b, removing columns does not affect row equivalence. Call the new, smaller, matrices $\hat{A}$ and $\hat{B}$. The new matrices should look this:

$$\hat{A} = \left(\begin{array}{c|c} I_N & a \\ 0 & 0 \end{array}\right) \text{ and } \hat{B} = \left(\begin{array}{c|c} I_N & b \\ 0 & 0 \end{array}\right),$$

where I_N is an $N \times N$ identity matrix and a and b are vectors.

Now if $\hat{A}$ and $\hat{B}$ have the same solution, then we must have $a = b$. But this is a contradiction! Then $A = B$. $\square$

`Explanation of the proof`

4 Solution Sets for Systems of Linear Equations

For a system of equations with r equations and k unknowns, one can have a number of different outcomes. For example, consider the case of r equations in three variables. Each of these equations is the equation of a plane in three-dimensional space. To find solutions to the system of equations, we look for the common intersection of the planes (if an intersection exists). Here we have five different possibilities:

1. **No solutions.** Some of the equations are contradictory, so no solutions exist.

2. **Unique Solution.** The planes have a unique point of intersection.

3. **Line.** The planes intersect in a common line; any point on that line then gives a solution to the system of equations.

4. **Plane.** Perhaps you only had one equation to begin with, or else all of the equations coincide geometrically. In this case, you have a plane of solutions, with two free parameters.

`Planes`

5. **All of $\mathbb{R}^3$.** If you start with no information, then any point in $\mathbb{R}^3$ is a solution. There are three free parameters.

In general, for systems of equations with k unknowns, there are $k + 2$ possible outcomes, corresponding to the number of free parameters in the solutions set, plus the possibility of no solutions. These types of "solution sets" are hard to visualize, but luckily "hyperplanes" behave like planes in $\mathbb{R}^3$ in many ways.

`Pictures and Explanation`

Reading homework: problem 4.1

4.1 Non-Leading Variables

Variables that are not a pivot in the reduced row echelon form of a linear system are *free*. We set them equal to arbitrary parameters $\mu_1, \mu_2, \ldots$

Example $\left(\begin{array}{cccc|c} 1 & 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$

Here, x_1 and x_2 are the pivot variables and x_3 and x_4 are non-leading variables, and thus free. The solutions are then of the form $x_3 = \mu_1$, $x_4 = \mu_2$, $x_2 = 1 + \mu_1 - \mu_2$, $x_1 = 1 - \mu_1 + \mu_2$.

The preferred way to write a solution set is with set notation. Let S be the set of solutions to the system. Then:

$$S = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

`Example`

We have already seen how to write a linear system of two equations in two unknowns as a matrix multiplying a vector. We can apply exactly the same idea for the above system of three equations in four unknowns by calling

$$M = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \text{ and } V = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Then if we take for the product of the matrix M with the vector X of unknowns

$$MX = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + x_3 - x_4 \\ x_2 - x_3 + x_4 \\ 0 \end{pmatrix}$$

our system becomes simply

$$MX = V.$$

Stare carefully at our answer for the product MX above. First you should notice that each of the three rows corresponds to the left hand side of one of the equations in the system. Also observe that each entry was obtained by matching the entries in the corresponding row of M with the column entries of X. For example, using the second row of M we obtained the second entry of MX

$$\begin{matrix} 0 & 1 & -1 & 1 \end{matrix} \quad \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} \quad \longmapsto \quad x_2 - x_3 + x_4 \,.$$

Later we will study matrix multiplication in detail, but you can already try to discover the main rules for yourself by working through Review Question 3 on multiplying matrices by vectors.

Given two vectors we can *add* them term-by-term:

$$\begin{pmatrix} a^1 \\ a^2 \\ a^3 \\ \vdots \\ a^r \end{pmatrix} + \begin{pmatrix} b^1 \\ b^2 \\ b^3 \\ \vdots \\ b^r \end{pmatrix} = \begin{pmatrix} a^1 + b^1 \\ a^2 + b^2 \\ a^3 + b^3 \\ \vdots \\ a^r + b^r \end{pmatrix}$$

We can also multiply a vector by a scalar, like so:

$$\lambda \begin{pmatrix} a^1 \\ a^2 \\ a^3 \\ \vdots \\ a^r \end{pmatrix} = \begin{pmatrix} \lambda a^1 \\ \lambda a^2 \\ \lambda a^3 \\ \vdots \\ \lambda a^r \end{pmatrix}$$

Then yet another way to write the solution set for the example is:

$$X = X_0 + \mu_1 Y_1 + \mu_2 Y_2$$

where

$$X_0 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, Y_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, Y_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

Definition Let X and Y be vectors and α and β be scalars. A function f is *linear* if

$$f(\alpha X + \beta Y) = \alpha f(X) + \beta f(Y)$$

This is called the *linearity property* for matrix multiplication.

The notion of linearity is a core concept in this course. Make sure you understand what it means and how to use it in computations!

Example Consider our example system above with

$$M = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \text{ and } Y = \begin{pmatrix} y_1 \\ y_2 \\ y^3 \\ y^4 \end{pmatrix},$$

and take for the function of vectors

$$f(X) = MX\,.$$

Now let us check the linearity property for f. The property needs to hold for *any* scalars α and β, so for simplicity let us concentrate first on the case $\alpha = \beta = 1$. This means that we need to compare the following two calculations:

1. First add $X + Y$, then compute $f(X + Y)$.
2. First compute $f(X)$ and $f(Y)$, then compute the sum $f(X) + f(Y)$.

The second computation is slightly easier:

$$f(X) = MX = \begin{pmatrix} x_1 + x_3 - x_4 \\ x_2 - x_3 + x_4 \\ 0 \end{pmatrix} \text{ and } f(Y) = MY = \begin{pmatrix} y_1 + y_3 - y_4 \\ y_2 - y_3 + y_4 \\ 0 \end{pmatrix},$$

(using our result above). Adding these gives

$$f(X) + f(Y) = \begin{pmatrix} x_1 + x_3 - x_4 + y_1 + y_3 - y_4 \\ x_2 - x_3 + x_4 + y_2 - y_3 + y_4 \\ 0 \end{pmatrix}.$$

Next we perform the first computation beginning with:

$$X + Y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ x_4 + y_4 \end{pmatrix},$$

from which we calculate

$$f(X+Y) = \begin{pmatrix} x_1 + y_2 + x_3 + y_3 - (x_4 + y_4) \\ x_2 + y_2 - (x_3 + y_3) + x_4 + y_4 \\ 0 \end{pmatrix}.$$

Distributing the minus signs and remembering that the order of adding numbers like $x_1, x_2, \ldots$ does not matter, we see that the two computations give exactly the same answer.

Of course, you should complain that we took a special choice of α and β. Actually, to take care of this we only need to check that $f(\alpha X) = \alpha f(X)$. It is your job to explain this in Review Question 1

Later we will show that matrix multiplication is always linear. Then we will know that:

$$M(\alpha X + \beta Y) = \alpha MX + \beta MY$$

Then the two equations $MX = V$ and $X = X_0 + \mu_1 Y_1 + \mu_2 Y_2$ together say that:

$$MX_0 + \mu_1 MY_1 + \mu_2 MY_2 = V$$

for *any* $\mu_1, \mu_2 \in \mathbb{R}$. Choosing $\mu_1 = \mu_2 = 0$, we obtain

$$MX_0 = V\,.$$

Here, X_0 is an example of what is called a *particular solution* to the system.

Given the particular solution to the system, we can then deduce that $\mu_1 MY_1 + \mu_2 MY_2 = 0$. Setting $\mu_1 = 1, \mu_2 = 0$, and recalling the particular solution $MX_0 = V$, we obtain

$$MY_1 = 0\,.$$

Likewise, setting $\mu_1 = 0, \mu_1 = 0$, we obtain

$$MY_2 = 0\,.$$

Here Y_1 and Y_2 are examples of what are called *homogeneous* solutions to the system. They *do not* solve the original equation $MX = V$, but instead its associated homogeneous system of equations $MY = 0$.

Example Consider the linear system with the augmented matrix we've been working with.

$$\begin{array}{rrrrl} x & & +z & -w & = 1 \\ & y & -z & +w & = 1 \end{array}$$

Recall that the system has the following solution set:

$$S = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Then $MX_0 = V$ says that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ solves the original system of equations, which is certainly true, but this is not the only solution.

$MY_1 = 0$ says that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ solves the homogeneous system.

$MY_2 = 0$ says that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$ solves the homogeneous system.

Notice how adding any multiple of a homogeneous solution to the particular solution yields another particular solution.

Definition Let M a matrix and V a vector. Given the linear system $MX = V$, we call X_0 a *particular solution* if $MX_0 = V$. We call Y a *homogeneous solution* if $MY = 0$. The linear system

$$MX = 0$$

is called the (associated) *homogeneous system*.

If X_0 is a particular solution, then the general solution to the system is[1]:

[1] The notation $S = \{X_0 + Y : MY = 0\}$ is read, "S is the set of all $X_0 + Y$ such that $MY = 0$," and means exactly that. Sometimes a pipe $|$ is used instead of a colon.

$$S = \{X_0 + Y : MY = 0\}$$

In other words, the general solution = particular + homogeneous.

Reading homework: problem 4.2

Wikipedia

- Systems of Linear Equations
- Row Echelon Form
- Row Echelon Form
- Elementary Matrix Operations

Review Problems

Linear Systems

1. Let M be a matrix and u and v vectors:
   $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, v = \begin{pmatrix} x \\ y \end{pmatrix}, u = \begin{pmatrix} w \\ z \end{pmatrix}.$

   (a) *Propose* a definition for $u + v$.

   (b) *Check* that your definition obeys $Mv + Mu = M(u + v)$.

2. *Matrix Multiplication:* Let M and N be matrices

   $$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } N = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

   and v a vector

   $$v = \begin{pmatrix} x \\ y \end{pmatrix}.$$

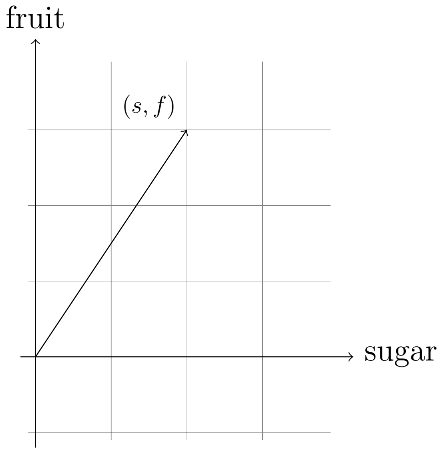
   Compute the vector Nv using the rule given above. Now multiply this vector by the matrix M, *i.e.*, compute the vector $M(Nv)$.

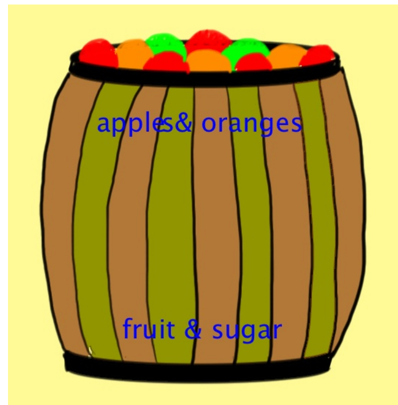
   Next recall that multiplication of ordinary numbers is associative, namely the order of brackets does not matter: $(xy)z = x(yz)$. Let us try to demand the same property for matrices and vectors, that is

   $$M(Nv) = (MN)v.$$

   We need to be careful reading this equation because Nv is a vector and so is $M(Nv)$. Therefore the right hand side, $(MN)v$ should also be a vector. This means that MN must be a matrix; in fact it is the matrix obtained by multiplying the matrices M and N. Use your result for $M(Nv)$ to find the matrix MN.

3. Pablo is a nutritionist who knows that oranges always have twice as much sugar as apples. When considering the sugar intake of schoolchildren eating a barrel of fruit, he represents the barrel like so:





Find a linear transformation relating Pablo's representation to the one in the lecture. Write your answer as a matrix.

Hint: Let λ represent the amount of sugar in each apple.

Hint

4. There are methods for solving linear systems other than Gauss' method. One often taught in high school is to solve one of the equations for a variable, then substitute the resulting expression into other equations. That step is repeated until there is an equation with only one variable. From that, the first number in the solution is derived, and then back-substitution can be done. This method takes longer than Gauss' method, since it involves more arithmetic operations, and is also more likely to lead to errors. To illustrate how it can lead to wrong conclusions, we will use the system

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

(a) Solve the first equation for x and substitute that expression into the second equation. Find the resulting y.

(b) Again solve the first equation for x, but this time substitute that expression into the third equation. Find this y.

What extra step must a user of this method take to avoid erroneously concluding a system has a solution?

Gaussian Elimination

1. State whether the following augmented matrices are in RREF and compute their solution sets.

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 3 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 & 2 & 0 \end{array}\right),$$

$$\left(\begin{array}{cccccc|c} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

$$\left(\begin{array}{ccccccc|c} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right).$$

2. Show that this pair of augmented matrices are row equivalent, assuming $ad - bc \neq 0$:

$$\left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \frac{de-bf}{ad-bc} \\ 0 & 1 & \frac{af-ce}{ad-bc} \end{array}\right)$$

3. Consider the augmented matrix: $\left(\begin{array}{cc|c} 2 & -1 & 3 \\ -6 & 3 & 1 \end{array}\right)$

   Give a *geometric* reason why the associated system of equations has no solution. (Hint, plot the three vectors given by the columns of this augmented matrix in the plane.) Given a general augmented matrix

$$\left(\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}\right),$$

   can you find a condition on the numbers a, b, c and d that create the geometric condition you found?

4. List as many operations on augmented matrices that *preserve* row equivalence as you can. Explain your answers. Give examples of operations that break row equivalence.

5. Row equivalence of matrices is an example of an *equivalence relation*. Recall that a relation $\sim$ on a set of objects U is an equivalence relation if the following three properties are satisfied:

   - Reflexive: For any $x \in U$, we have $x \sim x$.
   - Symmetric: For any $x, y \in U$, if $x \sim y$ then $y \sim x$.
   - Transitive: For any x, y and $z \in U$, if $x \sim y$ and $y \sim z$ then $x \sim z$.

   (For a fuller discussion of equivalence relations, see Homework 0, Problem 4)

   Show that row equivalence of augmented matrices is an equivalence relation.

`Hints for Questions 4 and 5`

Elementary Row Operations

1. (Row Equivalence)

   (a) Solve the following linear system using Gauss-Jordan elimination:

   $$\begin{aligned} 2x_1 + 5x_2 - 8x_3 + 2x_4 + 2x_5 &= 0 \\ 6x_1 + 2x_2 - 10x_3 + 6x_4 + 8x_5 &= 6 \\ 3x_1 + 6x_2 + 2x_3 + 3x_4 + 5x_5 &= 6 \\ 3x_1 + 1x_2 - 5x_3 + 3x_4 + 4x_5 &= 3 \\ 6x_1 + 7x_2 - 3x_3 + 6x_4 + 9x_5 &= 9 \end{aligned}$$

   Be sure to set your work out carefully with equivalence signs $\sim$ between each step, labeled by the row operations you performed.

   (b) Check that the following two matrices are row-equivalent:

   $$\left(\begin{array}{ccc|c} 1 & 4 & 7 & 10 \\ 2 & 9 & 6 & 0 \end{array}\right) \text{ and } \left(\begin{array}{ccc|c} 0 & -1 & 8 & 20 \\ 4 & 18 & 12 & 0 \end{array}\right)$$

   Now remove the third column from each matrix, and show that the resulting two matrices (shown below) are row-equivalent:

   $$\left(\begin{array}{cc|c} 1 & 4 & 10 \\ 2 & 9 & 0 \end{array}\right) \text{ and } \left(\begin{array}{cc|c} 0 & -1 & 20 \\ 4 & 18 & 0 \end{array}\right)$$

   Now remove the fourth column from each of the original two matrices, and show that the resulting two matrices, viewed as augmented matrices (shown below) are row-equivalent:

   $$\left(\begin{array}{cc|c} 1 & 4 & 7 \\ 2 & 9 & 6 \end{array}\right) \text{ and } \left(\begin{array}{cc|c} 0 & -1 & 8 \\ 4 & 18 & 12 \end{array}\right)$$

   Explain why row-equivalence is never affected by removing columns.

   (c) Check that the matrix $\left(\begin{array}{cc|c} 1 & 4 & 10 \\ 3 & 13 & 9 \\ 4 & 17 & 20 \end{array}\right)$ has no solutions. If you remove one of the rows of this matrix, does the new matrix have any solutions? In general, can row equivalence be affected by removing rows? Explain why or why not.

2. (Gaussian Elimination) Another method for solving linear systems is to use row operations to bring the augmented matrix to row echelon form. In row echelon form, the pivots are not necessarily set to one, and we only require that all entries left of the pivots are zero, not necessarily entries above a pivot. Provide a counterexample to show that row echelon form is not unique.

   Once a system is in row echelon form, it can be solved by "back substitution." Write the following row echelon matrix as a system of equations, then solve the system using back-substitution.

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 3 \end{array}\right)$$

3. Explain why the linear system has no solutions:

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 6 \end{array}\right)$$

   For which values of k does the system below have a solution?

$$\begin{array}{rcrcrcr} x & - & 3y & & & = & 6 \\ x & & & + & 3z & = & -3 \\ 2x & + & ky & + & (3-k)z & = & 1 \end{array}$$

Hint for question 3

Solution Sets for Systems of Linear Equations

1. Let $f(X) = MX$ where

$$M = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

   Suppose that α is any number. Compute the following four quantities:

$$\alpha X\,,\ f(X)\,,\ \alpha f(X) \text{ and } f(\alpha X)\,.$$

   Check your work by verifying that

$$\alpha f(X) = f(\alpha X)\,.$$

   Now explain why the result checked in the Lecture, namely

$$f(X+Y) = f(X) + f(Y)\,,$$

   and your result $f(\alpha X) = \alpha f(X)$ together imply

$$f(\alpha X + \beta Y) = \alpha f(X) + \beta f(Y)\,.$$

2. Write down examples of augmented matrices corresponding to each of the five types of solution sets for systems of equations with three unknowns.

3. Let

$$M = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_k^1 \\ a_1^2 & a_2^2 & \cdots & a_k^2 \\ \vdots & \vdots & & \vdots \\ a_1^r & a_2^r & \cdots & a_k^r \end{pmatrix}, \qquad X = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \end{pmatrix}$$

Propose a rule for MX so that $MX = 0$ is equivalent to the linear system:

$$\begin{array}{cccc} a_1^1 x^1 & +a_2^1 x^2 & \cdots +a_k^1 x^k & = 0 \\ a_1^2 x^1 & +a_2^2 x^2 & \cdots +a_k^2 x^k & = 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_1^r x^1 & +a_2^r x^2 & \cdots +a_k^r x^k & = 0 \end{array}$$

Show that your rule for multiplying a matrix by a vector obeys the linearity property.

Note that in this problem, x^2 does not denote the square of x. Instead x^1, x^2, x^3, etc... denote different variables. Although confusing at first, this notation was invented by Albert Einstein who noticed that quantities like $a_1^2 x^1 + a_2^2 x^2 \cdots + a_k^2 x^k$ could be written in summation notation as $\sum_{j=1}^k a_j^2 x^j$. Here j is called a summation index. Einstein observed that you could even drop the summation sign $\sum$ and simply write $a_j^2 x^j$.

Problem 3 hint

4. Use the rule you developed in the problem 3 to compute the following products

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 14 \\ 14 \\ 21 \\ 35 \\ 62 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 42 & 97 & 2 & -23 & 46 \\ 0 & 1 & 3 & 1 & 0 & 33 \\ 11 & \pi & 1 & 0 & 46 & 29 \\ -98 & 12 & 0 & 33 & 99 & 98 \\ \log 2 & 0 & \sqrt{2} & 0 & e & 23 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Now that you are good at multiplying a matrix with a column vector, try your hand at a product of two matrices

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Hint, to do this problem view the matrix on the right as three column vectors next to one another.

5. The *standard basis vector* e_i is a column vector with a one in the ith row, and zeroes everywhere else. Using the rule for multiplying a matrix times a vector in problem 3, find a simple rule for multiplying Me_i, where M is the general matrix defined there.

5 Scripts

5.1 What is Linear Algebra: 3×3 Matrix Example

Your friend places a jar on a table and tells you that there is 65 cents in this jar with 7 coins consisting of quarters, nickels, and dimes, and that there are twice as many dimes as quarters. Your friend wants to know how many nickels, dimes, and quarters are in the jar.

We can translate this into a system of the following linear equations:

$$\begin{aligned} 5n + 10d + 25q &= 65 \\ n + d + q &= 7 \\ d &= 2q \end{aligned}$$

Now we can rewrite the last equation in the form of $-d + 2q = 0$, and thus express this problem as the matrix equation

$$\begin{pmatrix} 5 & 10 & 25 \\ 1 & 1 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} n \\ d \\ q \end{pmatrix} = \begin{pmatrix} 65 \\ 7 \\ 0 \end{pmatrix}.$$

or as an augmented matrix (see also this script on the notation).

$$\left(\begin{array}{ccc|c} 5 & 10 & 25 & 65 \\ 1 & 1 & 1 & 7 \\ 0 & -1 & 2 & 0 \end{array}\right)$$

Now to solve it, using our original set of equations and by substitution, we have

$$\begin{aligned} 5n + 20q + 25q = 5n + 45q &= 65 \\ n + 2q + q = n + 3q &= 7 \end{aligned}$$

and by subtracting 5 times the bottom equation from the top, we get

$$45q - 15q = 30q = 65 - 35 = 30$$

and hence $q = 1$. Clearly $d = 2$, and hence $n = 7-2-1 = 4$. Therefore there are four nickels, two dimes, and one quarter.

5.2 What is Linear Algebra: Hint

Looking at the problem statement we find some important information, first that oranges always have twice as much sugar as apples, and second that the information about the barrel is recorded as (s, f), where $s =$ units of sugar in the barrel and $f =$ number of pieces of fruit in the barrel.

We are asked to find a linear transformation relating this new representation to the one in the lecture, where in the lecture $x =$ the number of apples and $y =$ the number of oranges. This means we must create a system of equations relating the variable x and y to the variables s and f in matrix form. Your answer should be the matrix that transforms one set of variables into the other.

Hint: Let λ represent the amount of sugar in each apple.

1. To find the first equation find a way to relate f to the variables x and y.
2. To find the second equation, use the hint to figure out how much sugar is in x apples, and y oranges in terms of λ. Then write an equation for s using x, y and λ.

5.3 Gaussian Elimination: Augmented Matrix Notation

Why is the augmented matrix

$$\left(\begin{array}{cc|c} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right),$$

equivalent to the system of equations

$$\begin{aligned} x + y &= 27 \\ 2x - y &= 0\,? \end{aligned}$$

Well the augmented matrix is just a new notation for the matrix equation

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 27 \\ 0 \end{pmatrix}$$

and if you review your matrix multiplication remember that

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 2x - y \end{pmatrix}$$

This means that

$$\begin{pmatrix} x + y \\ 2x - y \end{pmatrix} = \begin{pmatrix} 27 \\ 0 \end{pmatrix}$$

Which is our original equation.

5.4 Gaussian Elimination: Equivalence of Augmented Matrices

Lets think about what it means for the two augmented matrices

$$\left(\begin{array}{cc|c} 1 & 1 & 27 \\ 2 & -1 & 0 \end{array}\right),$$

and

$$\left(\begin{array}{cc|c} 1 & 0 & 9 \\ 0 & 1 & 18 \end{array}\right),$$

to be equivalent?

They are certainly not equal, because they don't match in each component, but since these augmented matrices represent a system, we might want to introduce a new kind of equivalence relation.

Well we could look at the system of linear equations this represents

$$\begin{aligned} x + y &= 27 \\ 2x - y &= 0\,? \end{aligned}$$

and notice that the solution is $x = 9$ and $y = 18$. The other augmented matrix represents the system

$$\begin{aligned} x + 0 \cdot y &= 9 \\ 0 \cdot x + y &= 18\,? \end{aligned}$$

This which clearly has the same solution. The first and second system are related in the sense that their solutions are the same. Notice that it is really nice to have the augmented matrix in the second form, because the matrix multiplication can be done in your head.

5.5 Gaussian Elimination: Hints for Review Questions 4 and 5

The hint for Review Question 4 is simple--just read the lecture on Elementary Row Operations.

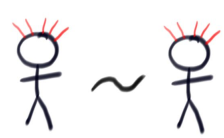
Question 5 looks harder than it actually is:

> Row equivalence of matrices is an example of an *equivalence relation*. Recall that a relation $\sim$ on a set of objects U is an equivalence relation if the following three properties are satisfied:
>
> - Reflexive: For any $x \in U$, we have $x \sim x$.
> - Symmetric: For any $x, y \in U$, if $x \sim y$ then $y \sim x$.
> - Transitive: For any x, y and $z \in U$, if $x \sim y$ and $y \sim z$ then $x \sim z$.
>
> (For a more complete discussion of equivalence relations, see Webwork Homework 0, Problem 4)
>
> Show that row equivalence of augmented matrices is an equivalence relation.

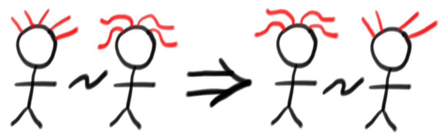
Firstly remember that an equivalence relation is just a more general version of ''equals''. Here we defined row equivalence for augmented matrices whose linear systems have solutions by the property that their solutions are the same.

So this question is really about the word *same*. Lets do a silly example: Lets replace the set of augmented matrices by the set of people who have hair. We will call two people equivalent if they have the same hair color. There are three properties to check:

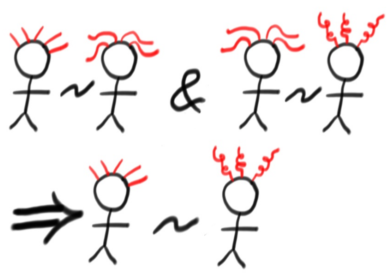
- Reflexive: This just requires that you have the same hair color as yourself so obviously holds.



- Symmetric: If the first person, Bob (say) has the same hair color as a second person Betty(say), then Bob has the same hair color as Betty, so this holds too.



- Transitive: If Bob has the same hair color as Betty (say) and Betty has the same color as Brenda (say), then it follows that Bob and Brenda have the same hair color, so the transitive property holds too and we are done.



5.6 Gaussian Elimination: 3×3 Example

We'll start with the matrix from the What is Linear Algebra: 3×3 Matrix Example which was

$$\left(\begin{array}{ccc|c} 5 & 10 & 25 & 65 \\ 1 & 1 & 1 & 7 \\ 0 & -1 & 2 & 0 \end{array}\right),$$

and recall the solution to the problem was $n = 4$, $d = 2$, and $q = 1$. So as a matrix equation we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} n \\ d \\ q \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

or as an augmented matrix

$$\left(\begin{array}{ccc|c} 1 & & & 4 \\ & 1 & & 2 \\ & & 1 & 1 \end{array}\right)$$

Note that often in diagonal matrices people will either omit the zeros or write in a single large zero. Now the first matrix is equivalent to the second matrix and is written as

$$\left(\begin{array}{ccc|c} 5 & 10 & 25 & 65 \\ 1 & 1 & 1 & 7 \\ 0 & -1 & 2 & 0 \end{array}\right), \sim \left(\begin{array}{ccc|c} 1 & & & 4 \\ & 1 & & 2 \\ & & 1 & 1 \end{array}\right)$$

since they have the same solutions.

5.7 Elementary Row Operations: Example

We have three basic rules

1. Row Swap
2. Scalar Multiplication
3. Row Sum

Lets look at an example. The system

$$3x + y = 7$$
$$x + 2y = 4$$

is something we learned to solve in high school algebra. Now we can write it in augmented matrix for this way

$$\left(\begin{array}{cc|c} 3 & 1 & 7 \\ 1 & 2 & 4 \end{array}\right).$$

We can see what these operations allow us to do:

1. Row swap allows us to switch the order of rows. In this example there are only two equations, so I will switch them. This will work with a larger system as well, but you have to decide which equations to switch. So we get

   $$x + 2y = 4$$
   $$3x + y = 7$$

   The augmented matrix looks like

   $$\left(\begin{array}{cc|c} 1 & 2 & 4 \\ 3 & 1 & 7 \end{array}\right).$$

   Notice that this won't change the solution of the system, but the augmented matrix will look different. This is where we can say that the original augmented matrix is equivalent to the one with the rows swapped. This will work with a larger system as well, but you have to decide which equations, or rows to switch. Make sure that you don't forget to switch the entries in the right-most column.

2. Scalar multiplication allows us to multiply both sides of an equation by a non-zero constant. So if we are starting with

$$\begin{aligned} x + 2y &= 4 \\ 3x + y &= 7 \end{aligned}$$

Then we can multiply the first equation by -3 which is a non-zero scalar. This operation will give us

$$\begin{aligned} -3x + -6y &= -12 \\ 3x + y &= 7 \end{aligned}$$

which has a corresponding augmented matrix

$$\left(\begin{array}{cc|c} -3 & -6 & -12 \\ 3 & 1 & 7 \end{array} \right).$$

Notice that we have multiplied the entire first row by -3, and this changes the augmented matrix, but not the solution of the system. We are not allowed to multiply by zero because it would be like replacing one of the equations with $0 = 0$, effectively destroying the information contained in the equation.

3. Row summing allows us to add one equation to another. In our example we could start with

$$\begin{aligned} -3x + -6y &= -12 \\ 3x + y &= 7 \end{aligned}$$

and replace the first equation with the sum of both equations. So we get

$$\begin{aligned} -3x + 3x + -6y + y &= -12 + 7 \\ 3x + y &= 7, \end{aligned}$$

which after some simplification is translates to

$$\left(\begin{array}{cc|c} 0 & -5 & -5 \\ 3 & 1 & 7 \end{array} \right).$$

When using this row operation make sure that you end up with as many equations as you started with. Here we replaced the first equation with a sum, but the second equation remained untouched.

In the example, notice that the x-terms in the first equation disappeared, which makes it much easier to solve for y. Think about what the next steps for solving this system would be using the language of elementary row operations.

5.8 Elementary Row Operations: Worked Examples

Let us consider that we are given two systems of equations that give rise to the following two (augmented) matrices:

$$\left(\begin{array}{cccc|c} 2 & 5 & 2 & 0 & 2 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 4 & 1 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{cc|c} 5 & 2 & 9 \\ 0 & 5 & 10 \\ 0 & 3 & 6 \end{array}\right)$$

and we want to find the solution to those systems. We will do so by doing Gaussian elimination.

For the first matrix we have

$$\begin{aligned} \left(\begin{array}{cccc|c} 2 & 5 & 2 & 0 & 2 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 4 & 1 & 0 & 1 \end{array}\right) \overset{R_1 \leftrightarrow R_2}{\sim} & \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 1 \\ 2 & 5 & 2 & 0 & 2 \\ 1 & 4 & 1 & 0 & 1 \end{array}\right) \\ \overset{R_2 - 2R_1; R_3 - R_1}{\sim} & \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \end{array}\right) \\ \overset{\frac{1}{3}R_2}{\sim} & \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \end{array}\right) \\ \overset{R_1 - R_2; R_3 - 3R_2}{\sim} & \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{aligned}$$

1. We begin by interchanging the first two rows in order to get a 1 in the upper-left hand corner and avoiding dealing with fractions.

2. Next we subtract row 1 from row 3 and twice from row 2 to get zeros in the left-most column.

3. Then we scale row 2 to have a 1 in the eventual pivot.

4. Finally we subtract row 2 from row 1 and three times from row 2 to get it into Row-Reduced Echelon Form.

Therefore we can write $x = 1 - \lambda$, $y = 0$, $z = \lambda$ and $w = \mu$, or in vector form

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now for the second system we have

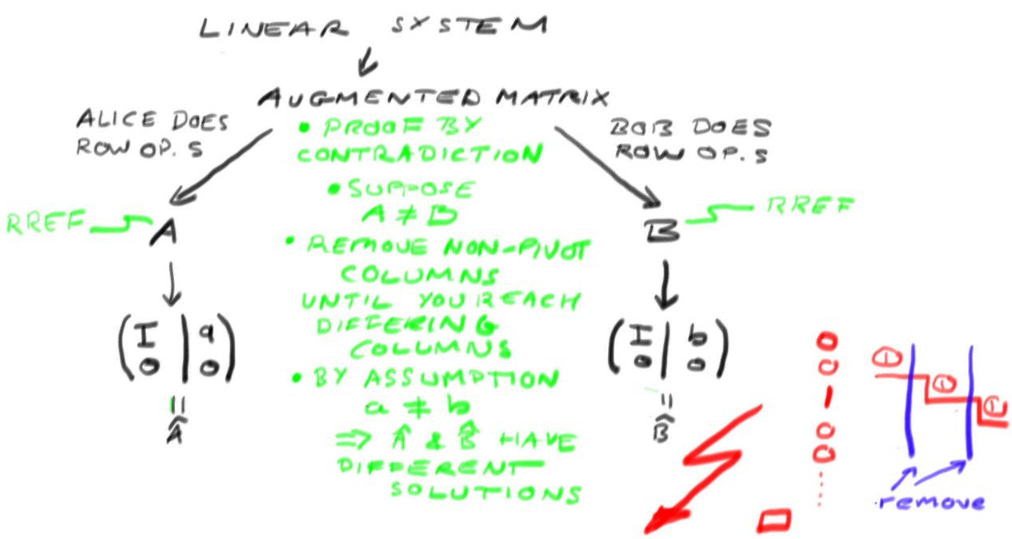
$$\begin{aligned} \left(\begin{array}{cc|c} 5 & 2 & 9 \\ 0 & 5 & 10 \\ 0 & 3 & 6 \end{array}\right) &\overset{\frac{1}{5}R_2}{\sim} \left(\begin{array}{cc|c} 5 & 2 & 9 \\ 0 & 1 & 2 \\ 0 & 3 & 6 \end{array}\right) \\ &\overset{R_3-3R_2}{\sim} \left(\begin{array}{cc|c} 5 & 2 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right) \\ &\overset{R_1-2R_2}{\sim} \left(\begin{array}{cc|c} 5 & 0 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right) \\ &\overset{\frac{1}{5}R_1}{\sim} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right) \end{aligned}$$

We scale the second and third rows appropriately in order to avoid fractions, then subtract the corresponding rows as before. Finally scale the first row and hence we have $x = 1$ and $y = 2$ as a unique solution.

5.9 Elementary Row Operations: Explanation of Proof for Theorem 3.1

The first thing to realize is that there are choices in the Gaussian elimination recipe, so maybe that could lead to two different RREF's and in turn two different solution sets for the same linear system. But that would be weird, in fact this Theorem says that this can never happen!

Because this proof comes at the end of the section it is often glossed over, but it is a very important result. Here's a sketch of what happens in the video:



In words: we start with a linear system and convert it to an augmented matrix. Then, because we are studying a uniqueness statement, we try a proof by contradiction. That is the method where to show that a statement is true, you try to demonstrate that the opposite of the statement leads to a contradiction. Here, the opposite statement to the theorem would be to find two different RREFs for the same system.

Suppose, therefore, that Alice and Bob do find different RREF augmented matrices called A and B. Then remove all the non-pivot columns from A and B until you hit the first column that differs. Record that in the last column and call the results $\widehat{A}$ and $\widehat{B}$. Removing columns does change the solution sets, but it does not ruin row equivalence, so $\widehat{A}$ and $\widehat{B}$ have the same solution sets.

Now, because we left only the pivot columns (plus the first column that differs) we have

$$\hat{A} = \left(\begin{array}{c|c} I_N & a \\ 0 & 0 \end{array}\right) \text{ and } \hat{B} = \left(\begin{array}{c|c} I_N & b \\ 0 & 0 \end{array}\right),$$

where I_N is an identity matrix and a and b are column vectors. Importantly, by assumption,

$$a \neq b.$$

So if we try to wrote down the solution sets for $\widehat{A}$ and $\widehat{B}$ they would be different. But at all stages, we only performed operations that kept Alice's solution set the same as Bob's. This is a contradiction so the proof is complete.

5.10 Elementary Row Operations: Hint for Review Question 3

The first part for Review Question 3 is simple--just write out the associated linear system and you will find the equation $0 = 6$ which is inconsistent. Therefore we learn that we must avoid a row of zeros preceding a non-vanishing entry after the vertical bar.

Turning to the system of equations, we first write out the augmented matrix and then perform two row operations

$$\left(\begin{array}{ccc|c} 1 & -3 & 0 & 6 \\ 1 & 0 & 3 & -3 \\ 2 & k & 3-k & 1 \end{array}\right)$$

$$\stackrel{R_2-R_1;R_3-2R_1}{\sim} \left(\begin{array}{ccc|c} 1 & -3 & 0 & 6 \\ 0 & 3 & 3 & -9 \\ 0 & k+6 & 3-k & -11 \end{array}\right).$$

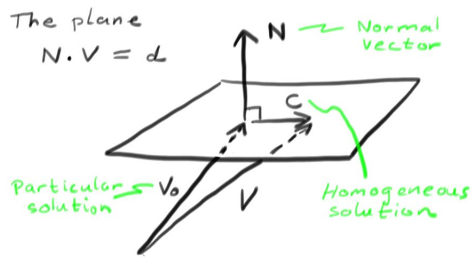
Next we would like to subtract some amount of R_2 from R_3 to achieve a zero in the third entry of the second column. But if

$$k + 6 = 3 - k \Rightarrow k = -\frac{3}{2},$$

this would produce zeros in the third row before the vertical line. You should also check that this does not make the whole third line zero. You now have enough information to write a complete solution.

5.11 Solution Sets for Systems of Linear Equations: Planes

Here we want to describe the mathematics of planes in space. The video is summarised by the following picture:



A plane is often called $\mathbb{R}^2$ because it is spanned by two coordinates, and space is called $\mathbb{R}^3$ and has three coordinates, usually called (x, y, z). The equation for a plane is

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

Lets simplify this by calling $V = (x, y, z)$ the vector of unknowns and $N = (a, b, c)$. Using the dot product in $\mathbb{R}^3$ we have

$$N \cdot V = d\,.$$

Remember that when vectors are perpendicular their dot products vanish. *I.e.* $U \cdot V = 0 \Leftrightarrow U \perp V$. This means that if a vector V_0 solves our equation $N \cdot V = d$, then so too does $V_0 + C$ whenever C is perpendicular to N. This is because

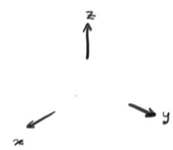
$$N \cdot (V_0 + C) = N \cdot V_0 + N \cdot C = d + 0 = d\,.$$

But C is ANY vector perpendicular to N, so all the possibilities for C span a plane whose normal vector is N. Hence we have shown that solutions to the equation $ax + by + cz = 0$ are a plane with normal vector $N = (a, b, c)$.

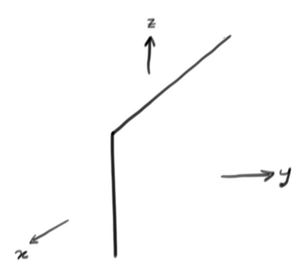
5.12 Solution Sets for Systems of Linear Equations: Pictures and Explanation

This video considers solutions sets for linear systems with three unknowns. These are often called (x, y, z) and label points in $\mathbb{R}^3$. Lets work case by case:

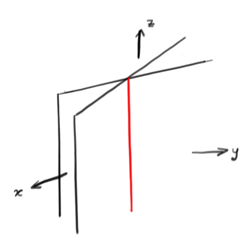
- If you have no equations at all, then any (x, y, z) is a solution, so the solution set is all of $\mathbb{R}^3$. The picture looks a little silly:



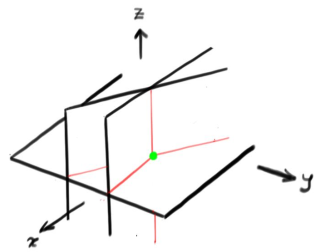
- For a single equation, the solution is a plane. This is explained in this video or the accompanying script. The picture looks like this:



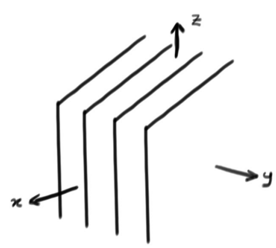
- For two equations, we must look at two planes. These usually intersect along a line, so the solution set will also (usually) be a line:



- For three equations, most often their intersection will be a single point so the solution will then be unique:



- Of course stuff can go wrong. Two different looking equations could determine the same plane, or worse equations could be inconsistent. If the equations are inconsistent, there will be no solutions at all. For example, if you had four equations determining four parallel planes the solution set would be empty. This looks like this:



5.13 Solution Sets for Systems of Linear Equations: Example

Here is an augmented matrix, let's think about what the solution set looks like

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 2 \\ 0 & 1 & 0 & 1 \end{array}\right)$$

This looks like the system

$$\begin{aligned} x_1 + 3x_3 &= 2 \\ x_2 &= 1 \end{aligned}$$

Notice that when the system is written this way the copy of the 2×2 identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ makes it easy to write a solution in terms of the variables x_1 and x_2. We will call x_1 and x_2 the *pivot* variables. The third column $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ does not look like part of an identity matrix, and there is no 3×3 identity in the augmented matrix. Notice there are more variables than equations and that this means we will have to write the solutions for the system in terms of the variable x_3. We'll call x_3 the *free* variable.

Let $x_3 = \mu$. Then we can rewrite the first equation in our system

$$\begin{aligned} x_1 + 3x_3 &= 2 \\ x_1 + 3\mu &= 2 \\ x_1 &= 2 - 3\mu. \end{aligned}$$

Then since the second equation doesn't depend on μ we can keep the equation

$$x_2 = 1,$$

and for a third equation we can write

$$x_3 = \mu$$

so that we get system

$$\begin{aligned}
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 2-3\mu \\ 1 \\ \mu \end{pmatrix} \\
&= \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -3\mu \\ 0 \\ \mu \end{pmatrix} \\
&= \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}$$

So for any value of μ will give a solution of the system, and any system can be written in this form for some value of μ. Since there are multiple solutions, we can also express them as a set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \,\middle|\, \mu \in \mathbb{R} \right\}.$$

5.14 Solution Sets for Systems of Linear Equations: Hint

For the first part of this problem, the key is to consider the vector as a $n \times 1$ matrix. For the second part, all you need to show is that

$$M(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot (MX) + \beta \cdot (MY)$$

where $\alpha, \beta \in \mathbb{R}$ (or whatever field we are using) and

$$Y = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^k \end{pmatrix}.$$

Note that this will be somewhat tedious, and many people use summation notation or Einstein's summation convention with the added notation of M_j denoting the j-th row of the matrix. For example, for any j we have

$$(MX)_j = \sum_{i=1}^{k} a_i^j x^i = a_i^j x^i.$$

You can see a concrete example after the definition of the linearity property.

Webwork Links

Reading homework: problem 1.1

Reading homework: problem 1.1

Reading homework: problem 2.1

Reading homework: problem 2.2

Reading homework: problem 3.1

Reading homework: problem 4.1

Reading homework: problem 4.2

Background homework set

Linear Systems

Solution Sets