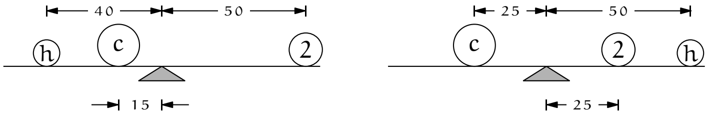
Chapter One

Linear Systems

I Solving Linear Systems

Systems of linear equations are common in science and mathematics. These two examples from high school science [Onan] give a sense of how they arise.

The first example is from Statics. Suppose that we have three objects, we know that one has a mass of 2 kg, and we want to find the two unknown masses. Suppose further that experimentation with a meter stick produces these two balances.



For the masses to balance we must have that the sum of moments on the left equals the sum of moments on the right, where the moment of an object is its mass times its distance from the balance point. That gives a system of two linear equations.

$$\begin{aligned} 40h + 15c &= 100 \\ 25c &= 50 + 50h \end{aligned}$$

The second example is from Chemistry. We can mix, under controlled conditions, toluene C_7H_8 and nitric acid HNO_3 to produce trinitrotoluene $C_7H_5O_6N_3$ along with the byproduct water (conditions have to be very well controlled — trinitrotoluene is better known as TNT). In what proportion should we mix them? The number of atoms of each element present before the reaction

$$x\, C_7H_8 \;+\; y\, HNO_3 \quad\longrightarrow\quad z\, C_7H_5O_6N_3 \;+\; w\, H_2O$$

must equal the number present afterward. Applying that in turn to the elements C, H, N, and O gives this system.

$$\begin{aligned} 7x &= 7z \\ 8x + 1y &= 5z + 2w \\ 1y &= 3z \\ 3y &= 6z + 1w \end{aligned}$$

Both examples come down to solving a system of equations. In each system, the equations involve only the first power of each variable. This chapter shows how to solve any such system.

I.1 Gauss's Method

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*. 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*

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

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

1.2 Example The combination $3x_1 + 2x_2$ of x_1 and x_2 is linear. The combination $3x_1^2 + 2\sin(x_2)$ is not linear, nor is $3x_1^2 + 2x_2$.

1.3 Example The ordered pair $(-1, 5)$ is a solution of this system.

$$\begin{aligned} 3x_1 + 2x_2 &= 7 \\ -x_1 + x_2 &= 6 \end{aligned}$$

In contrast, $(5, -1)$ is not a solution.

Finding the set of all solutions is *solving* the system. We don't need guesswork or good luck; there is an algorithm that always works. This algorithm is *Gauss's Method* (or *Gaussian elimination* or *linear elimination*).

1.4 Example To solve this system

$$\begin{aligned} 3x_3 &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ \tfrac{1}{3}x_1 + 2x_2 \qquad &= 3 \end{aligned}$$

we transform it, step by step, until it is in a form that we can easily solve.

The first transformation rewrites the system by interchanging the first and third row.

$$\xrightarrow{\text{swap row 1 with row 3}} \begin{aligned} \tfrac{1}{3}x_1 + 2x_2 \qquad &= 3 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ 3x_3 &= 9 \end{aligned}$$

The second transformation rescales the first row by a factor of 3.

$$\xrightarrow{\text{multiply row 1 by 3}} \begin{aligned} x_1 + 6x_2 \qquad &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2 \\ 3x_3 &= 9 \end{aligned}$$

The third transformation is the only nontrivial one in this example. We mentally multiply both sides of the first row by -1, mentally add that to the second row, and write the result in as the new second row.

$$\xrightarrow{\text{add } -1 \text{ times row 1 to row 2}} \begin{aligned} x_1 + 6x_2 \qquad &= 9 \\ -x_2 - 2x_3 &= -7 \\ 3x_3 &= 9 \end{aligned}$$

The point of these steps is that we've brought the system to a form where we can easily find the value of each variable. The bottom equation shows that $x_3 = 3$. Substituting 3 for x_3 in the middle equation shows that $x_2 = 1$. Substituting those two into the top equation gives that $x_1 = 3$. Thus the system has a unique solution; the solution set is $\{(3, 1, 3)\}$.

Most of this subsection and the next one consists of examples of solving linear systems by Gauss's Method, which we will use throughout the book. It is fast and easy. But before we do those examples we will first show that it is also safe: Gauss's Method never loses solutions (any solution to the system before you apply the method is also a solution after), nor does it ever picks up extraneous solutions (any tuple that is not a solution before is also not a solution after).

1.5 Theorem (Gauss's Method) 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.

Each of the three Gauss's Method operations has a restriction. Multiplying a row by 0 is not allowed because obviously that can change the solution set. Similarly, adding a multiple of a row to itself is not allowed because adding -1 times the row to itself has the effect of multiplying the row by 0. We disallow swapping a row with itself to make some results in the fourth chapter easier, and also because it's pointless.

Proof We will cover the equation swap operation here. The other two cases are Exercise 31.

Consider a linear system.

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &=& d_1 \\ &\vdots& \\ a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n &=& d_i \\ &\vdots& \\ a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n &=& d_j \\ &\vdots& \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &=& d_m \end{array}$$

The tuple $(s_1, \ldots, s_n)$ satisfies this system if and only if substituting the values for the variables, the s's for the x's, gives a conjunction of true statements: $a_{1,1}s_1 + a_{1,2}s_2 + \cdots + a_{1,n}s_n = d_1$ and … $a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n = d_i$ and … $a_{j,1}s_1 + a_{j,2}s_2 + \cdots + a_{j,n}s_n = d_j$ and … $a_{m,1}s_1 + a_{m,2}s_2 + \cdots + a_{m,n}s_n = d_m$.

In a list of statements joined with 'and' we can rearrange the order of the statements. Thus this requirement is met if and only if $a_{1,1}s_1 + a_{1,2}s_2 + \cdots + a_{1,n}s_n = d_1$ and … $a_{j,1}s_1 + a_{j,2}s_2 + \cdots + a_{j,n}s_n = d_j$ and … $a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n = d_i$ and … $a_{m,1}s_1 + a_{m,2}s_2 + \cdots + a_{m,n}s_n = d_m$. This is exactly the requirement that $(s_1, \ldots, s_n)$ solves the system after the row swap. QED

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*.

When writing out the calculations, we will abbreviate 'row i' by 'ρ_i'. For instance, we will denote a row combination operation by $k\rho_i + \rho_j$, with the row that changes written second. To save writing we will often combine addition steps when they use the same ρ_i as in the next example.

1.7 Example Gauss's Method systematically applies the row operations to solve a system. Here is a typical case.

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

We begin by using the first row to eliminate the $2x$ in the second row and the x in the third. To get rid of the $2x$ we mentally multiply the entire first row by -2, add that to the second row, and write the result in as the new second row. To eliminate the x in the third row we multiply the first row by -1, add that to the third row, and write the result in as the new third row.

$$\xrightarrow[-\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{aligned} x + y \qquad &= 0 \\ -3y + 3z &= 3 \\ -3y - z &= 3 \end{aligned}$$

We finish by transforming the second system into a third, where the bottom equation involves only one unknown. We do that by using the second row to eliminate the y term from the third row.

$$\xrightarrow{-\rho_2+\rho_3} \begin{aligned} x + y \qquad &= 0 \\ -3y + 3z &= 3 \\ -4z &= 0 \end{aligned}$$

Now finding the system's solution is easy. The third row gives $z = 0$. Substitute that back into the second row to get $y = -1$. Then substitute back into the first row to get $x = 1$.

1.8 Example For the Physics problem from the start of this chapter, Gauss's Method gives this.

$$\begin{aligned} 40h + 15c &= 100 \\ -50h + 25c &= 50 \end{aligned} \xrightarrow{5/4\rho_1+\rho_2} \begin{aligned} 40h + 15c &= 100 \\ (175/4)c &= 175 \end{aligned}$$

So $c = 4$, and back-substitution gives that $h = 1$. (We will solve the Chemistry problem later.)

1.9 Example The reduction

$$\begin{array}{rcl} x+\ y+\ z=9 \\ 2x+4y-3z=1 \\ 3x+6y-5z=0 \end{array} \xrightarrow[-3\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{array}{r} x+\ y+\ z=\ 9 \\ 2y-5z=-17 \\ 3y-8z=-27 \end{array}$$

$$\xrightarrow{-(3/2)\rho_2+\rho_3} \begin{array}{r} x+\ y+\ z=\ 9 \\ 2y-\ 5z=\ -17 \\ -(1/2)z=-(3/2) \end{array}$$

shows that $z=3$, $y=-1$, and $x=7$.

As illustrated above, the point of Gauss's Method is to use the elementary reduction operations to set up back-substitution.

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 all-zero rows are at the bottom.

1.11 Example The prior three examples only used the operation of row combination. This linear system requires the swap operation to get it into echelon form because after the first combination

$$\begin{array}{r} x-\ y\qquad\qquad=0 \\ 2x-2y+\ z+2w=4 \\ y\qquad+\ w=0 \\ 2z+\ w=5 \end{array} \xrightarrow{-2\rho_1+\rho_2} \begin{array}{r} x-y\qquad\qquad=0 \\ z+2w=4 \\ y\qquad+\ w=0 \\ 2z+\ w=5 \end{array}$$

the second equation has no leading y. To get one, we put in place a lower-down row that has a leading y.

$$\xrightarrow{\rho_2\leftrightarrow\rho_3} \begin{array}{r} x-y\qquad\qquad=0 \\ y\qquad+\ w=0 \\ z+2w=4 \\ 2z+\ w=5 \end{array}$$

(Had there been more than one suitable row below the second then we could have swapped in any one.) With that, Gauss's Method proceeds as before.

$$\xrightarrow{-2\rho_3+\rho_4} \begin{array}{r} x-y\qquad\qquad=\ 0 \\ y\qquad+\ w=\ 0 \\ z+\ 2w=\ 4 \\ -3w=-3 \end{array}$$

Back-substitution gives $w=1$, $z=2$, $y=-1$, and $x=-1$.

Strictly speaking, to solve linear systems we don't need the row rescaling operation. We have introduced it here because it is convenient and because we will use it later in this chapter as part of a variation of Gauss's Method, the Gauss-Jordan Method.

All of the systems seen so far have the same number of equations as unknowns. All of them have a solution and for all of them there is only one solution. We finish this subsection by seeing other things that can happen.

1.12 Example This system has more equations than variables.

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

Gauss's Method helps us understand this system also, since this

$$\xrightarrow[-2\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ -4y &= -4 \end{aligned}$$

shows that one of the equations is redundant. Echelon form

$$\xrightarrow{-(4/5)\rho_2+\rho_3} \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ 0 &= 0 \end{aligned}$$

gives that $y = 1$ and $x = -2$. The '$0 = 0$' reflects the redundancy.

Gauss's Method is also useful on systems with more variables than equations. The next subsection has many examples.

Another way that linear systems can differ from the examples shown above is that some linear systems do not have a unique solution. This can happen in two ways. The first is that a system can fail to have any solution at all.

1.13 Example Contrast the system in the last example with this one.

$$\begin{aligned} x + 3y &= 1 \\ 2x + y &= -3 \\ 2x + 2y &= 0 \end{aligned} \quad \xrightarrow[-2\rho_1+\rho_3]{-2\rho_1+\rho_2} \quad \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ -4y &= -2 \end{aligned}$$

Here the system is inconsistent: no pair of numbers (s_1, s_2) satisfies all three equations simultaneously. Echelon form makes the inconsistency obvious.

$$\xrightarrow{-(4/5)\rho_2+\rho_3} \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ 0 &= 2 \end{aligned}$$

The solution set is empty.

1.14 Example The prior system has more equations than unknowns but that is not what causes the inconsistency — Example 1.12 has more equations than unknowns and yet is consistent. Nor is having more equations than unknowns necessary for inconsistency, as we see with this inconsistent system that has the same number of equations as unknowns.

$$\begin{array}{rcl} x + 2y = 8 \\ 2x + 4y = 8 \end{array} \xrightarrow{-2\rho_1+\rho_2} \begin{array}{rcl} x + 2y = 8 \\ 0 = -8 \end{array}$$

The other way that a linear system can fail to have a unique solution, besides having no solutions, is to have many solutions.

1.15 Example In this system

$$\begin{aligned} x + y &= 4 \\ 2x + 2y &= 8 \end{aligned}$$

any pair of numbers satisfying the first equation also satisfies the second. The solution set $\{(x, y) \mid x + y = 4\}$ is infinite; some example member pairs are $(0, 4)$, $(-1, 5)$, and $(2.5, 1.5)$.

The result of applying Gauss's Method here contrasts with the prior example because we do not get a contradictory equation.

$$\xrightarrow{-2\rho_1+\rho_2} \begin{aligned} x + y &= 4 \\ 0 &= 0 \end{aligned}$$

Don't be fooled by that example. A $0 = 0$ equation does not signal that a system has many solutions.

1.16 Example The absence of a $0 = 0$ equation does not keep a system from having many different solutions. This system is in echelon form, has no $0 = 0$, but has infinitely many solutions.

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

Some solutions are: $(0, 1, -1)$, $(0, 1/2, -1/2)$, $(0, 0, 0)$, and $(0, -\pi, \pi)$. Any triple whose first component is 0 and whose second component is the negative of the third is a solution.

Nor does the presence of a $0 = 0$ mean that the system must have many solutions. Example 1.12 shows that. So does this system, which does not have any solutions at all despite that in echelon form it has a $0 = 0$ row.

$$\begin{array}{rcl} \begin{aligned} 2x \qquad - 2z &= 6 \\ y + \ z &= 1 \\ 2x + \ y - \ z &= 7 \\ 3y + 3z &= 0 \end{aligned} & \xrightarrow{-\rho_1+\rho_3} & \begin{aligned} 2x \qquad - 2z &= 6 \\ y + \ z &= 1 \\ y + \ z &= 1 \\ 3y + 3z &= 0 \end{aligned} \\ & \xrightarrow[-3\rho_2+\rho_4]{-\rho_2+\rho_3} & \begin{aligned} 2x \qquad - 2z &= 6 \\ y + \ z &= 1 \\ 0 &= 0 \\ 0 &= -3 \end{aligned} \end{array}$$

In summary, Gauss's Method uses the row operations to set a system up for back substitution. If any step shows a contradictory equation then we can stop with the conclusion that the system has no solutions. If we reach echelon form without a contradictory equation, and each variable is a leading variable in its row, then the system has a unique solution and we find it by back substitution. Finally, if we reach echelon form without a contradictory equation, and there is not a unique solution — that is, at least one variable is not a leading variable — then the system has many solutions.

The next subsection explores the third case. We will see that such a system must have infinitely many solutions and we will describe the solution set.

Note *Here, and in the rest of the book, you must justify all of your exercise answers. For instance, if a question asks whether a system has a solution then you must justify a yes response by producing the solution and must justify a no response by showing that no solution exists.*

Exercises

✓ **1.17** Use Gauss's Method to find the unique solution for each system.

(a) $\begin{aligned} 2x + 3y &= 13 \\ x - \ y &= -1 \end{aligned}$ (b) $\begin{aligned} x \qquad - z &= 0 \\ 3x + y \qquad &= 1 \\ -x + y + z &= 4 \end{aligned}$

✓ **1.18** Use Gauss's Method to solve each system or conclude 'many solutions' or 'no solutions'.

(a) $\begin{aligned} 2x + 2y &= 5 \\ x - 4y &= 0 \end{aligned}$ (b) $\begin{aligned} -x + y &= 1 \\ x + y &= 2 \end{aligned}$ (c) $\begin{aligned} x - 3y + \ z &= 1 \\ x + \ y + 2z &= 14 \end{aligned}$ (d) $\begin{aligned} -x - \ y &= 1 \\ -3x - 3y &= 2 \end{aligned}$

(e) $\begin{aligned} 4y + z &= 20 \\ 2x - 2y + z &= 0 \\ x \qquad + z &= 5 \\ x + \ y - z &= 10 \end{aligned}$ (f) $\begin{aligned} 2x \qquad + \ z + w &= 5 \\ y \qquad - w &= -1 \\ 3x \qquad - \ z - w &= 0 \\ 4x + y + 2z + w &= 9 \end{aligned}$

✓ **1.19** We can solve linear systems by methods other than Gauss's. One often taught in high school is to solve one of the equations for a variable, then substitute the resulting expression into other equations. Then we repeat that step until there is an equation with only one variable. From that we get the first number in the solution and then we get the rest with back-substitution. This method takes longer than Gauss's 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

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

from Example 1.13.

(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?

✓ 1.20 For which values of k are there no solutions, many solutions, or a unique solution to this system?

$$\begin{aligned} x-y&=1\\ 3x-3y&=k \end{aligned}$$

✓ 1.21 This system is not linear in that it says $\sin\alpha$ instead of α

$$\begin{aligned} 2\sin\alpha-\cos\beta+3\tan\gamma&=3\\ 4\sin\alpha+2\cos\beta-2\tan\gamma&=10\\ 6\sin\alpha-3\cos\beta+\tan\gamma&=9 \end{aligned}$$

and yet we can apply Gauss's Method. Do so. Does the system have a solution?

✓ 1.22 [Anton] What conditions must the constants, the b's, satisfy so that each of these systems has a solution? *Hint.* Apply Gauss's Method and see what happens to the right side.

(a) $$\begin{aligned} x-3y&=b_1\\ 3x+y&=b_2\\ x+7y&=b_3\\ 2x+4y&=b_4 \end{aligned}$$

(b) $$\begin{aligned} x_1+2x_2+3x_3&=b_1\\ 2x_1+5x_2+3x_3&=b_2\\ x_1+8x_3&=b_3 \end{aligned}$$

1.23 True or false: a system with more unknowns than equations has at least one solution. (As always, to say 'true' you must prove it, while to say 'false' you must produce a counterexample.)

1.24 Must any Chemistry problem like the one that starts this subsection — a balance the reaction problem — have infinitely many solutions?

✓ 1.25 Find the coefficients a, b, and c so that the graph of $f(x)=ax^2+bx+c$ passes through the points $(1,2)$, $(-1,6)$, and $(2,3)$.

1.26 After Theorem 1.5 we note that multiplying a row by 0 is not allowed because that could change a solution set. Give an example of a system with solution set S_0 where after multiplying a row by 0 the new system has a solution set S_1 and S_0 is a proper subset of S_1. Give an example where $S_0=S_1$.

1.27 Gauss's Method works by combining the equations in a system to make new equations.

(a) Can we derive the equation $3x - 2y = 5$ by a sequence of Gaussian reduction steps from the equations in this system?

$$\begin{aligned} x + y &= 1 \\ 4x - y &= 6 \end{aligned}$$

(b) Can we derive the equation $5x - 3y = 2$ with a sequence of Gaussian reduction steps from the equations in this system?

$$\begin{aligned} 2x + 2y &= 5 \\ 3x + y &= 4 \end{aligned}$$

(c) Can we derive $6x - 9y + 5z = -2$ by a sequence of Gaussian reduction steps from the equations in the system?

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

1.28 Prove that, where $a, b, \ldots, e$ are real numbers and $a \neq 0$, if

$$ax + by = c$$

has the same solution set as

$$ax + dy = e$$

then they are the same equation. What if $a = 0$?

✓ 1.29 Show that if $ad - bc \neq 0$ then

$$\begin{aligned} ax + by &= j \\ cx + dy &= k \end{aligned}$$

has a unique solution.

✓ 1.30 In the system

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

each of the equations describes a line in the xy-plane. By geometrical reasoning, show that there are three possibilities: there is a unique solution, there is no solution, and there are infinitely many solutions.

1.31 Finish the proof of Theorem 1.5.

1.32 Is there a two-unknowns linear system whose solution set is all of $\mathbb{R}^2$?

✓ 1.33 Are any of the operations used in Gauss's Method redundant? That is, can we make any of the operations from a combination of the others?

1.34 Prove that each operation of Gauss's Method is reversible. That is, show that if two systems are related by a row operation $S_1 \to S_2$ then there is a row operation to go back $S_2 \to S_1$.

? 1.35 [Anton] A box holding pennies, nickels and dimes contains thirteen coins with a total value of 83 cents. How many coins of each type are in the box? (These are US coins; a penny is 1 cent, a nickel is 5 cents, and a dime is 10 cents.)

? 1.36 [Con. Prob. 1955] Four positive integers are given. Select any three of the integers, find their arithmetic average, and add this result to the fourth integer. Thus the numbers 29, 23, 21, and 17 are obtained. One of the original integers is:

(a) 19 **(b)** 21 **(c)** 23 **(d)** 29 **(e)** 17

? ✓ **1.37** [Am. Math. Mon., Jan. 1935] Laugh at this: AHAHA + TEHE = TEHAW. It resulted from substituting a code letter for each digit of a simple example in addition, and it is required to identify the letters and prove the solution unique.

? **1.38** [Wohascum no. 2] The Wohascum County Board of Commissioners, which has 20 members, recently had to elect a President. There were three candidates (A, B, and C); on each ballot the three candidates were to be listed in order of preference, with no abstentions. It was found that 11 members, a majority, preferred A over B (thus the other 9 preferred B over A). Similarly, it was found that 12 members preferred C over A. Given these results, it was suggested that B should withdraw, to enable a runoff election between A and C. However, B protested, and it was then found that 14 members preferred B over C! The Board has not yet recovered from the resulting confusion. Given that every possible order of A, B, C appeared on at least one ballot, how many members voted for B as their first choice?

? **1.39** [Am. Math. Mon., Jan. 1963] "This system of n linear equations with n unknowns," said the Great Mathematician, "has a curious property."

"Good heavens!" said the Poor Nut, "What is it?"

"Note," said the Great Mathematician, "that the constants are in arithmetic progression."

"It's all so clear when you explain it!" said the Poor Nut. "Do you mean like $6x + 9y = 12$ and $15x + 18y = 21$?"

"Quite so," said the Great Mathematician, pulling out his bassoon. "Indeed, the system has a unique solution. Can you find it?"

"Good heavens!" cried the Poor Nut, "I am baffled."

Are you?

I.2 Describing the Solution Set

A linear system with a unique solution has a solution set with one element. A linear system with no solution has a solution set that is empty. In these cases the solution set is easy to describe. Solution sets are a challenge to describe only when they contain many elements.

2.1 Example This system has many solutions because in echelon form

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

$$\xrightarrow{-\rho_2+\rho_3} \begin{array}{rcl} 2x \phantom{{}-y} + z &=& 3 \\ -y - (3/2)z &=& -1/2 \\ 0 &=& 0 \end{array}$$

not all of the variables are leading variables. Theorem 1.5 shows that an (x, y, z) satisfies the first system if and only if it satisfies the third. So we can describe the solution set $\{(x, y, z) \mid 2x + z = 3 \text{ and } x - y - z = 1 \text{ and } 3x - y = 4\}$ in this way.

$$\{(x, y, z) \mid 2x + z = 3 \text{ and } -y - 3z/2 = -1/2\} \qquad (*)$$

This description is better because it has two equations instead of three but it is not optimal because it still has some hard to understand interactions among the variables.

To improve it, use the variable that does not lead any equation, z, to describe the variables that do lead, x and y. The second equation gives $y = (1/2) - (3/2)z$ and the first equation gives $x = (3/2) - (1/2)z$. Thus we can describe the solution set in this way.

$$\{(x, y, z) = ((3/2) - (1/2)z, (1/2) - (3/2)z, z) \mid z \in \mathbb{R}\} \qquad (**)$$

Compared with $(*)$, the advantage of $(**)$ is that z can be any real number. This makes the job of deciding which tuples are in the solution set much easier. For instance, taking $z = 2$ shows that $(1/2, -5/2, 2)$ is a solution.

2.2 Definition In an echelon form linear system the variables that are not leading are *free*.

2.3 Example Reduction of a linear system can end with more than one variable free. Gauss's Method on this system

$$\begin{array}{rcrcrcrcr} x & + & y & + & z & - & w & = & 1 \\ & & y & - & z & + & w & = & -1 \\ 3x & & & + & 6z & - & 6w & = & 6 \\ & & -y & + & z & - & w & = & 1 \end{array} \xrightarrow{-3\rho_1+\rho_3} \begin{array}{rcrcrcrcr} x & + & y & + & z & - & w & = & 1 \\ & & y & - & z & + & w & = & -1 \\ & & -3y & + & 3z & - & 3w & = & 3 \\ & & -y & + & z & - & w & = & 1 \end{array}$$

$$\xrightarrow[\rho_2+\rho_4]{3\rho_2+\rho_3} \begin{array}{rcrcrcrcr} x & + & y & + & z & - & w & = & 1 \\ & & y & - & z & + & w & = & -1 \\ & & & & & & 0 & = & 0 \\ & & & & & & 0 & = & 0 \end{array}$$

leaves x and y leading and both z and w free. To get the description that we prefer, we work from the bottom. We first express the leading variable y in terms of z and w, as $y = -1 + z - w$. Moving up to the top equation, substituting for y gives $x + (-1 + z - w) + z - w = 1$ and solving for x leaves $x = 2 - 2z + 2w$. The solution set

$$\{(2 - 2z + 2w, -1 + z - w, z, w) \mid z, w \in \mathbb{R}\} \qquad (**)$$

has the leading variables in terms of the variables that are free.

2.4 Example The list of leading variables may skip over some columns. After this reduction

$$\begin{array}{rcl} 2x-2y &=& 0 \\ z+3w &=& 2 \\ 3x-3y &=& 0 \\ x-y+2z+6w &=& 4 \end{array} \quad \xrightarrow[-(1/2)\rho_1+\rho_4]{-(3/2)\rho_1+\rho_3} \quad \begin{array}{rcl} 2x-2y &=& 0 \\ z+3w &=& 2 \\ 0 &=& 0 \\ 2z+6w &=& 4 \end{array}$$

$$\xrightarrow{-2\rho_2+\rho_4} \quad \begin{array}{rcl} 2x-2y &=& 0 \\ z+3w &=& 2 \\ 0 &=& 0 \\ 0 &=& 0 \end{array}$$

x and z are the leading variables, not x and y. The free variables are y and w and so we can describe the solution set as $\{(y, y, 2-3w, w) \mid y, w \in \mathbb{R}\}$. For instance, $(1, 1, 2, 0)$ satisfies the system — take $y = 1$ and $w = 0$. The four-tuple $(1, 0, 5, 4)$ is not a solution since its first coordinate does not equal its second.

A variable that we use to describe a family of solutions is a *parameter*. We say that the solution set in the prior example is *parametrized* with y and w.

(The terms 'parameter' and 'free variable' do not mean the same thing. In the prior example y and w are free because in the echelon form system they do not lead while they are parameters because of how we used them to describe the set of solutions. Had we instead rewritten the second equation as $w = 2/3 - (1/3)z$ then the free variables would still be y and w but the parameters would be y and z.)

In the rest of this book we will solve linear systems by bringing them to echelon form and then parametrizing with the free variables.

2.5 Example This is another system with infinitely many solutions.

$$\begin{array}{rcl} x+2y &=& 1 \\ 2x +z &=& 2 \\ 3x+2y+z-w &=& 4 \end{array} \quad \xrightarrow[-3\rho_1+\rho_3]{-2\rho_1+\rho_2} \quad \begin{array}{rcl} x+2y &=& 1 \\ -4y+z &=& 0 \\ -4y+z-w &=& 1 \end{array}$$

$$\xrightarrow{-\rho_2+\rho_3} \quad \begin{array}{rcl} x+2y &=& 1 \\ -4y+z &=& 0 \\ -w &=& 1 \end{array}$$

The leading variables are x, y, and w. The variable z is free. Notice that, although there are infinitely many solutions, the value of w doesn't vary but is constant at -1. To parametrize, write w in terms of z with $w = -1 + 0z$. Then $y = (1/4)z$. Substitute for y in the first equation to get $x = 1 - (1/2)z$. The solution set is $\{(1 - (1/2)z, (1/4)z, z, -1) \mid z \in \mathbb{R}\}$.

Parametrizing solution sets shows that systems with free variables have infinitely many solutions. For instance, above z takes on all of infinitely many real number values, each associated with a different solution.

We finish this subsection by developing a streamlined notation for linear systems and their solution sets.

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*.

We usually denote a matrix with an upper case roman letters. For instance,

$$A = \begin{pmatrix} 1 & 2.2 & 5 \\ 3 & 4 & -7 \end{pmatrix}$$

has 2 rows and 3 columns and so is a 2×3 matrix. Read that aloud as "two-by-three"; the number of rows is always given first. (The matrix has parentheses on either side so that when two matrices are adjacent we can tell where one ends and the other begins.) We name matrix entries with the corresponding lower-case letter so that $a_{2,1} = 3$ is the entry in the second row and first column of the above array. Note that the order of the subscripts matters: $a_{1,2} \neq a_{2,1}$ since $a_{1,2} = 2.2$.

We write $\mathcal{M}_{n \times m}$ for the set of all $n \times m$ matrices.

We use matrices to do Gauss's Method in essentially the same way that we did it for systems of equations: where a row's *leading entry*. is its first nonzero entry (if it has one), we perform row operations to arrive at *matrix echelon form*,where the leading entry in lower rows are to the right of those in the rows above. We switch to this notation because it lightens the clerical load of Gauss's Method — the copying of variables and the writing of $+$'s and $=$'s.

2.7 Example We can abbreviate this linear system

$$\begin{aligned} x + 2y \phantom{{}-2z} &= 4 \\ y - z &= 0 \\ x \phantom{{}+2y} + 2z &= 4 \end{aligned}$$

with this matrix.

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

The vertical bar reminds a reader of the difference between the coefficients on the system's left hand side and the constants on the right. With a bar, this is an *augmented* matrix.

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 2 & 4 \end{array}\right) \xrightarrow{-\rho_1+\rho_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & -2 & 2 & 0 \end{array}\right) \xrightarrow{2\rho_2+\rho_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

The second row stands for $y - z = 0$ and the first row stands for $x + 2y = 4$ so the solution set is $\{(4 - 2z, z, z) \mid z \in \mathbb{R}\}$.

Matrix notation also clarifies the descriptions of solution sets. Example 2.3's $\{(2 - 2z + 2w, -1 + z - w, z, w) \mid z, w \in \mathbb{R}\}$ is hard to read. We will rewrite it to group all of the constants together, all of the coefficients of z together, and all of the coefficients of w together. We write them vertically, in one-column matrices.

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

For instance, the top line says that $x = 2 - 2z + 2w$ and the second line says that $y = -1 + z - w$. (Our next section gives a geometric interpretation that will help picture the solution sets.)

2.8 Definition A *vector* (or *column vector*) is a matrix with a single column. A matrix with a single row is a *row vector*. The entries of a vector are its *components*. A column or row vector whose components are all zeros is a *zero vector*.

Vectors are an exception to the convention of representing matrices with capital roman letters. We use lower-case roman or greek letters overlined with an arrow: $\vec{a}$, $\vec{b}$, ... or $\vec{\alpha}$, $\vec{\beta}$, ... (boldface is also common: $\mathbf{a}$ or $\boldsymbol{\alpha}$). For instance, this is a column vector with a third component of 7.

$$\vec{v} = \begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix}$$

A zero vector is denoted $\vec{0}$. There are many different zero vectors — the one-tall zero vector, the two-tall zero vector, etc. — but nonetheless we will often say "the" zero vector, expecting that the size will be clear from the context.

2.9 Definition The linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = d$ with unknowns $x_1, \ldots, x_n$ is *satisfied* by

$$\vec{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

if $a_1s_1 + a_2s_2 + \cdots + a_ns_n = d$. A vector satisfies a linear system if it satisfies each equation in the system.

The style of description of solution sets that we use involves adding the vectors, and also multiplying them by real numbers. Before we give the examples showing the style we first need to define these operations.

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

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

Note that for the addition to be defined the vectors must have the same number of entries. This entry-by-entry addition works for any pair of matrices, not just vectors, provided that they have the same number of rows and columns.

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

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

As with the addition operation, the entry-by-entry scalar multiplication operation extends beyond vectors to apply to any matrix.

We write scalar multiplication either as $r\cdot\vec{v}$ or $\vec{v}\cdot r$, or even without the '$\cdot$' symbol: $r\vec{v}$. (Do not refer to scalar multiplication as 'scalar product' because we will use that name for a different operation.)

2.12 Example

$$\begin{pmatrix}2\\3\\1\end{pmatrix}+\begin{pmatrix}3\\-1\\4\end{pmatrix}=\begin{pmatrix}2+3\\3-1\\1+4\end{pmatrix}=\begin{pmatrix}5\\2\\5\end{pmatrix}\qquad 7\cdot\begin{pmatrix}1\\4\\-1\\-3\end{pmatrix}=\begin{pmatrix}7\\28\\-7\\-21\end{pmatrix}$$

Observe that the definitions of addition and scalar multiplication agree where they overlap; for instance, $\vec{v}+\vec{v}=2\vec{v}$.

With these definitions, we are set to use matrix and vector notation to both solve systems and express the solution.

2.13 Example This system

$$\begin{array}{rcl}2x+y-w&=&4\\ y+w+u&=&4\\ x-z+2w&=&0\end{array}$$

reduces in this way.

$$\left(\begin{array}{ccccc|c} 2 & 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 & 0 & 0 \end{array}\right) \xrightarrow{-(1/2)\rho_1+\rho_3} \left(\begin{array}{ccccc|c} 2 & 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & -1/2 & -1 & 5/2 & 0 & -2 \end{array}\right)$$

$$\xrightarrow{(1/2)\rho_2+\rho_3} \left(\begin{array}{ccccc|c} 2 & 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 1 & 4 \\ 0 & 0 & -1 & 3 & 1/2 & 0 \end{array}\right)$$

The solution set is $\{(w+(1/2)u, 4-w-u, 3w+(1/2)u, w, u) \mid w, u \in \mathbb{R}\}$. We write that in vector form.

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

Note how well vector notation sets off the coefficients of each parameter. For instance, the third row of the vector form shows plainly that if u is fixed then z increases three times as fast as w. Another thing shown plainly is that setting both w and u to zero gives that

$$\begin{pmatrix} x \\ y \\ z \\ w \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

is a particular solution of the linear system.

2.14 Example In the same way, this system

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

reduces

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

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

to a one-parameter solution set.

$$\{\begin{pmatrix}1\\0\\0\end{pmatrix}+\begin{pmatrix}-1/3\\2/3\\1\end{pmatrix}z \mid z\in\mathbb{R}\}$$

As in the prior example, the vector not associated with the parameter

$$\begin{pmatrix}1\\0\\0\end{pmatrix}$$

is a particular solution of the system.

Before the exercises, we will consider what we have accomplished and what we have yet to do.

So far we have done the mechanics of Gauss's Method. We have not stopped to consider any of the interesting questions that arise, except for proving Theorem 1.5 — which justifies the method by showing that it gives the right answers.

For example, can we always describe solution sets as above, with a particular solution vector added to an unrestricted linear combination of some other vectors? We've noted that the solution sets we described in this way have infinitely many solutions so an answer to this question would tell us about the size of solution sets.

Many questions arise from our observation that we can do Gauss's Method in more than one way (for instance, when swapping rows we may have a choice of more than one row). Theorem 1.5 says that we must get the same solution set no matter how we proceed but if we do Gauss's Method in two ways must we get the same number of free variables in each echelon form system? Must those be the same variables, that is, is it impossible to solve a problem one way to get y and w free and solve it another way to get y and z free?

In the rest of this chapter we will answer these questions. The answer to each is 'yes'. In the next subsection we do the first one: we will prove that we can always describe solution sets in that way. Then, in this chapter's second section, we will use that understanding to describe the geometry of solution sets. In this chapter's final section, we will settle the questions about the parameters.

When we are done, we will not only have a solid grounding in the practice of Gauss's Method but we will also have a solid grounding in the theory. We will know exactly what can and cannot happen in a reduction.

Exercises

✓ **2.15** Find the indicated entry of the matrix, if it is defined.

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

(a) $a_{2,1}$ (b) $a_{1,2}$ (c) $a_{2,2}$ (d) $a_{3,1}$

✓ 2.16 Give the size of each matrix.

(a) $\begin{pmatrix}1&0&4\\2&1&5\end{pmatrix}$ (b) $\begin{pmatrix}1&1\\-1&1\\3&-1\end{pmatrix}$ (c) $\begin{pmatrix}5&10\\10&5\end{pmatrix}$

✓ 2.17 Do the indicated vector operation, if it is defined.

(a) $\begin{pmatrix}2\\1\\1\end{pmatrix}+\begin{pmatrix}3\\0\\4\end{pmatrix}$ (b) $5\begin{pmatrix}4\\-1\end{pmatrix}$ (c) $\begin{pmatrix}1\\5\\1\end{pmatrix}-\begin{pmatrix}3\\1\\1\end{pmatrix}$ (d) $7\begin{pmatrix}2\\1\end{pmatrix}+9\begin{pmatrix}3\\5\end{pmatrix}$

(e) $\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}1\\2\\3\end{pmatrix}$ (f) $6\begin{pmatrix}3\\1\\1\end{pmatrix}-4\begin{pmatrix}2\\0\\3\end{pmatrix}+2\begin{pmatrix}1\\1\\5\end{pmatrix}$

✓ 2.18 Solve each system using matrix notation. Express the solution using vectors.

(a) $\begin{aligned}3x+6y&=18\\x+2y&=6\end{aligned}$ (b) $\begin{aligned}x+y&=1\\x-y&=-1\end{aligned}$ (c) $\begin{aligned}x_1+x_3&=4\\x_1-x_2+2x_3&=5\\4x_1-x_2+5x_3&=17\end{aligned}$

(d) $\begin{aligned}2a+b-c&=2\\2a+c&=3\\a-b&=0\end{aligned}$ (e) $\begin{aligned}x+2y-z&=3\\2x+y+w&=4\\x-y+z+w&=1\end{aligned}$ (f) $\begin{aligned}x+z+w&=4\\2x+y-w&=2\\3x+y+z&=7\end{aligned}$

✓ 2.19 Solve each system using matrix notation. Give each solution set in vector notation.

(a) $\begin{aligned}2x+y-z&=1\\4x-y&=3\end{aligned}$ (b) $\begin{aligned}x-z&=1\\y+2z-w&=3\\x+2y+3z-w&=7\end{aligned}$ (c) $\begin{aligned}x-y+z&=0\\y+w&=0\\3x-2y+3z+w&=0\\-y-w&=0\end{aligned}$

(d) $\begin{aligned}a+2b+3c+d-e&=1\\3a-b+c+d+e&=3\end{aligned}$

✓ 2.20 The vector is in the set. What value of the parameters produces that vector?

(a) $\begin{pmatrix}5\\-5\end{pmatrix}$, $\{\begin{pmatrix}1\\-1\end{pmatrix}k \mid k\in\mathbb{R}\}$

(b) $\begin{pmatrix}-1\\2\\1\end{pmatrix}$, $\{\begin{pmatrix}-2\\1\\0\end{pmatrix}i+\begin{pmatrix}3\\0\\1\end{pmatrix}j \mid i,j\in\mathbb{R}\}$

(c) $\begin{pmatrix}0\\-4\\2\end{pmatrix}$, $\{\begin{pmatrix}1\\1\\0\end{pmatrix}m+\begin{pmatrix}2\\0\\1\end{pmatrix}n \mid m,n\in\mathbb{R}\}$

2.21 Decide if the vector is in the set.

(a) $\begin{pmatrix}3\\-1\end{pmatrix}$, $\{\begin{pmatrix}-6\\2\end{pmatrix}k \mid k\in\mathbb{R}\}$

(b) $\begin{pmatrix}5\\4\end{pmatrix}$, $\{\begin{pmatrix}5\\-4\end{pmatrix}j \mid j\in\mathbb{R}\}$

(c) $\begin{pmatrix}2\\1\\-1\end{pmatrix}, \{\begin{pmatrix}0\\3\\-7\end{pmatrix}+\begin{pmatrix}1\\-1\\3\end{pmatrix}r \mid r\in\mathbb{R}\}$

(d) $\begin{pmatrix}1\\0\\1\end{pmatrix}, \{\begin{pmatrix}2\\0\\1\end{pmatrix}j+\begin{pmatrix}-3\\-1\\1\end{pmatrix}k \mid j,k\in\mathbb{R}\}$

2.22 [Cleary] A farmer with 1200 acres is considering planting three different crops, corn, soybeans, and oats. The farmer wants to use all 1200 acres. Seed corn costs \$20 per acre, while soybean and oat seed cost \$50 and \$12 per acre respectively. The farmer has \$40 000 available to buy seed and intends to spend it all.

(a) Use the information above to formulate two linear equations with three unknowns and solve it.

(b) Solutions to the system are choices that the farmer can make. Write down two reasonable solutions.

(c) Suppose that in the fall when the crops mature, the farmer can bring in revenue of \$100 per acre for corn, \$300 per acre for soybeans and \$80 per acre for oats. Which of your two solutions in the prior part would have resulted in a larger revenue?

2.23 Parametrize the solution set of this one-equation system.

$$x_1+x_2+\cdots+x_n=0$$

✓ 2.24 **(a)** Apply Gauss's Method to the left-hand side to solve

$$\begin{aligned} x+2y-w&=a\\ 2x+z&=b\\ x+y+2w&=c \end{aligned}$$

for x, y, z, and w, in terms of the constants a, b, and c.

(b) Use your answer from the prior part to solve this.

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

✓ 2.25 Why is the comma needed in the notation '$a_{i,j}$' for matrix entries?

✓ 2.26 Give the 4×4 matrix whose i,j-th entry is

(a) $i+j$; **(b)** -1 to the $i+j$ power.

2.27 For any matrix A, the *transpose* of A, written A^{T}, is the matrix whose columns are the rows of A. Find the transpose of each of these.

(a) $\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}$ **(b)** $\begin{pmatrix}2&-3\\1&1\end{pmatrix}$ **(c)** $\begin{pmatrix}5&10\\10&5\end{pmatrix}$ **(d)** $\begin{pmatrix}1\\1\\0\end{pmatrix}$

✓ 2.28 **(a)** Describe all functions $f(x)=ax^2+bx+c$ such that $f(1)=2$ and $f(-1)=6$.

(b) Describe all functions $f(x)=ax^2+bx+c$ such that $f(1)=2$.

2.29 Show that any set of five points from the plane $\mathbb{R}^2$ lie on a common conic section, that is, they all satisfy some equation of the form $ax^2+by^2+cxy+dx+ey+f=0$ where some of $a, \ldots, f$ are nonzero.

2.30 Make up a four equations/four unknowns system having

(a) a one-parameter solution set;

(b) a two-parameter solution set;
(c) a three-parameter solution set.

? 2.31 [Shepelev] This puzzle is from a Russian web-site http://www.arbuz.uz/ and there are many solutions to it, but mine uses linear algebra and is very naive. There's a planet inhabited by arbuzoids (watermeloners, to translate from Russian). Those creatures are found in three colors: red, green and blue. There are 13 red arbuzoids, 15 blue ones, and 17 green. When two differently colored arbuzoids meet, they both change to the third color.

The question is, can it ever happen that all of them assume the same color?

? 2.32 [USSR Olympiad no. 174]
(a) Solve the system of equations.

$$\begin{aligned} ax + y &= a^2 \\ x + ay &= 1 \end{aligned}$$

For what values of a does the system fail to have solutions, and for what values of a are there infinitely many solutions?
(b) Answer the above question for the system.

$$\begin{aligned} ax + y &= a^3 \\ x + ay &= 1 \end{aligned}$$

? 2.33 [Math. Mag., Sept. 1952] In air a gold-surfaced sphere weighs 7588 grams. It is known that it may contain one or more of the metals aluminum, copper, silver, or lead. When weighed successively under standard conditions in water, benzene, alcohol, and glycerin its respective weights are 6588, 6688, 6778, and 6328 grams. How much, if any, of the forenamed metals does it contain if the specific gravities of the designated substances are taken to be as follows?

Aluminum	2.7	Alcohol	0.81
Copper	8.9	Benzene	0.90
Gold	19.3	Glycerin	1.26
Lead	11.3	Water	1.00
Silver	10.8		

I.3 General = Particular + Homogeneous

In the prior subsection the descriptions of solution sets all fit a pattern. They have a vector that is a particular solution of the system added to an unrestricted combination of some other vectors. The solution set from Example 2.13

illustrates.

$$\{\underbrace{\begin{pmatrix}0\\4\\0\\0\\0\end{pmatrix}}_{\substack{\text{particular}\\\text{solution}}}\underbrace{+w\begin{pmatrix}1\\-1\\3\\1\\0\end{pmatrix}+u\begin{pmatrix}1/2\\-1\\1/2\\0\\1\end{pmatrix}}_{\substack{\text{unrestricted}\\\text{combination}}}\mid w,u\in\mathbb{R}\}$$

The combination is unrestricted in that w and u can be any real numbers — there is no condition like "such that $2w-u=0$" to restrict which pairs w,u we can use.

That example shows an infinite solution set fitting the pattern. The other two kinds of solution sets also fit. A one-element solution set fits because it has a particular solution, and the unrestricted combination part is trivial. (That is, instead of being a combination of two vectors or of one vector, it is a combination of no vectors. By convention the sum of an empty set of vectors is the zero vector.) An empty solution set fits the pattern because there is no particular solution and thus there are no sums of that form at all.

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 $\vec{\beta}_1$, ..., $\vec{\beta}_k$ equals the number of free variables that the system has after a Gaussian reduction.

The solution description has two parts, the particular solution $\vec{p}$ and the unrestricted linear combination of the $\vec{\beta}$'s. We shall prove the theorem with two corresponding lemmas.

We will focus first on the unrestricted combination. For that we consider systems that have the vector of zeroes as a particular solution so that we can shorten $\vec{p}+c_1\vec{\beta}_1+\cdots+c_k\vec{\beta}_k$ to $c_1\vec{\beta}_1+\cdots+c_k\vec{\beta}_k$.

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

3.3 Example With any linear system like

$$\begin{aligned}3x+4y&=3\\2x-\ \ y&=1\end{aligned}$$

we associate a system of homogeneous equations by setting the right side to zeros.

$$\begin{aligned} 3x + 4y &= 0 \\ 2x - y &= 0 \end{aligned}$$

Compare the reduction of the original system

$$\begin{array}{rl} 3x + 4y = 3 \\ 2x - y = 1 \end{array} \xrightarrow{-(2/3)\rho_1+\rho_2} \begin{array}{rl} 3x + 4y &= 3 \\ -(11/3)y &= -1 \end{array}$$

with the reduction of the associated homogeneous system.

$$\begin{array}{rl} 3x + 4y = 0 \\ 2x - y = 0 \end{array} \xrightarrow{-(2/3)\rho_1+\rho_2} \begin{array}{rl} 3x + 4y &= 0 \\ -(11/3)y &= 0 \end{array}$$

Obviously the two reductions go in the same way. We can study how to reduce a linear systems by instead studying how to reduce the associated homogeneous system.

Studying the associated homogeneous system has a great advantage over studying the original system. Nonhomogeneous systems can be inconsistent. But a homogeneous system must be consistent since there is always at least one solution, the zero vector.

3.4 Example Some homogeneous systems have the zero vector as their only solution.

$$\begin{array}{rl} 3x + 2y + z &= 0 \\ 6x + 4y &= 0 \\ y + z &= 0 \end{array} \xrightarrow{-2\rho_1+\rho_2} \begin{array}{rl} 3x + 2y + z &= 0 \\ -2z &= 0 \\ y + z &= 0 \end{array} \xrightarrow{\rho_2\leftrightarrow\rho_3} \begin{array}{rl} 3x + 2y + z &= 0 \\ y + z &= 0 \\ -2z &= 0 \end{array}$$

3.5 Example Some homogeneous systems have many solutions. One is the Chemistry problem from the first page of the first subsection.

$$\begin{array}{rl} 7x - 7z &= 0 \\ 8x + y - 5z - 2w &= 0 \\ y - 3z &= 0 \\ 3y - 6z - w &= 0 \end{array} \xrightarrow{-(8/7)\rho_1+\rho_2} \begin{array}{rl} 7x - 7z &= 0 \\ y + 3z - 2w &= 0 \\ y - 3z &= 0 \\ 3y - 6z - w &= 0 \end{array}$$

$$\xrightarrow[-3\rho_2+\rho_4]{-\rho_2+\rho_3} \begin{array}{rl} 7x - 7z &= 0 \\ y + 3z - 2w &= 0 \\ -6z + 2w &= 0 \\ -15z + 5w &= 0 \end{array}$$

$$\xrightarrow{-(5/2)\rho_3+\rho_4} \begin{array}{rl} 7x - 7z &= 0 \\ y + 3z - 2w &= 0 \\ -6z + 2w &= 0 \\ 0 &= 0 \end{array}$$

The solution set

$$\{\begin{pmatrix}1/3\\1\\1/3\\1\end{pmatrix}w \mid w\in\mathbb{R}\}$$

has many vectors besides the zero vector (if we interpret w as a number of molecules then solutions make sense only when w is a nonnegative multiple of 3).

3.6 Lemma For any homogeneous linear system there exist vectors $\vec{\beta}_1$, ..., $\vec{\beta}_k$ 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.

We will make two points before the proof. The first is that the basic idea of the proof is straightforward. Consider this system of homogeneous equations in echelon form.

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

Start with the bottom equation. Express its leading variable in terms of the free variables with $u=-v$. For the next row up, substitute for the leading variable u of the row below $y+z+(-v)-v=0$ and solve for this row's leading variable $y=-z+2v$. Iterate: on the next row up, substitute expressions found in lower rows $x+(-z+2v)+2z+(-v)+v=0$ and solve for the leading variable $x=-z-2v$. To finish, write the solution in vector notation

$$\begin{pmatrix}x\\y\\z\\s\\t\end{pmatrix}=\begin{pmatrix}-1\\-1\\1\\0\\0\end{pmatrix}z+\begin{pmatrix}-2\\2\\0\\-1\\1\end{pmatrix}v \qquad \text{for } z,v\in\mathbb{R}$$

and recognize that the $\vec{\beta}_1$ and $\vec{\beta}_2$ of the lemma are the vectors associated with the free variables z and v.

The prior paragraph is an example, not a proof. But it does suggest the second point about the proof, its approach. The example moves row-by-row up the system, using the equations from lower rows to do the next row. This points to doing the proof by mathematical induction.*

* More information on mathematical induction is in the appendix.

Induction is an important and non-obvious proof technique that we shall use a number of times in this book. We will do proofs by induction in two steps, a base step and an inductive step. In the base step we verify that the statement is true for some first instance, here that for the bottom equation we can write the leading variable in terms of free variables. In the inductive step we must establish an implication, that if the statement is true for all prior cases then it follows for the present case also. Here we will establish that if for the bottom-most t rows we can express the leading variables in terms of the free variables, then for the $t+1$-th row from the bottom we can also express the leading variable in terms of those that are free.

Those two steps together prove the statement for all the rows because by the base step it is true for the bottom equation, and by the inductive step the fact that it is true for the bottom equation shows that it is true for the next one up. Then another application of the inductive step implies that it is true for the third equation up, etc.

Proof Apply Gauss's Method to get to echelon form. There may be some $0=0$ equations; we ignore these (if the system consists only of $0=0$ equations then the lemma is trivially true because there are no leading variables). But because the system is homogeneous there are no contradictory equations.

We will use induction to verify that each leading variable can be expressed in terms of free variables. That will finish the proof because we can use the free variables as parameters and the $\vec{\beta}$'s are the vectors of coefficients of those free variables.

For the base step consider the bottom-most equation

$$a_{m,\ell_m}x_{\ell_m} + a_{m,\ell_m+1}x_{\ell_m+1} + \cdots + a_{m,n}x_n = 0 \qquad (*)$$

where $a_{m,\ell_m} \neq 0$. (The 'ℓ' means "leading" so that x_{ℓ_m} is the leading variable in row m.) This is the bottom row so any variables after the leading one must be free. Move these to the right hand side and divide by a_{m,ℓ_m}

$$x_{\ell_m} = (-a_{m,\ell_m+1}/a_{m,\ell_m})x_{\ell_m+1} + \cdots + (-a_{m,n}/a_{m,\ell_m})x_n$$

to express the leading variable in terms of free variables. (There is a tricky technical point here: if in the bottom equation $(*)$ there are no variables to the right of x_{l_m} then $x_{\ell_m} = 0$. This satisfies the statement we are verifying because, as alluded to at the start of this subsection, it has x_{ℓ_m} written as a sum of a number of the free variables, namely as the sum of zero many, under the convention that a trivial sum totals to 0.)

For the inductive step assume that the statement holds for the bottom-most t rows, with $0 \leqslant t < m-1$. That is, assume that for the m-th equation, and the $(m-1)$-th equation, etc., up to and including the $(m-t)$-th equation, we can express the leading variable in terms of free ones. We must verify that this then also holds for the next equation up, the $(m-(t+1))$-th equation. For that, take each variable that leads in a lower equation $x_{\ell_m}, \ldots, x_{\ell_{m-t}}$ and substitute its expression in terms of free variables. We only need expressions for leading variables from lower equations because the system is in echelon form, so leading variables in higher equation do not appear in this equation. The result has a leading term of $a_{m-(t+1),\ell_{m-(t+1)}} x_{\ell_{m-(t+1)}}$ with $a_{m-(t+1),\ell_{m-(t+1)}} \neq 0$, and the rest of the left hand side is a linear combination of free variables. Move the free variables to the right side and divide by $a_{m-(t+1),\ell_{m-(t+1)}}$ to end with this equation's leading variable $x_{\ell_{m-(t+1)}}$ in terms of free variables.

We have done both the base step and the inductive step so by the principle of mathematical induction the proposition is true. QED

This shows, as discussed between the lemma and its proof, that we can parametrize solution sets using the free variables. We say that the set of vectors $\{c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$ is *generated by* or *spanned by* the set $\{\vec{\beta}_1, \ldots, \vec{\beta}_k\}$.

To finish the proof of Theorem 3.1 the next lemma considers the particular solution part of the solution set's description.

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

Proof We will show mutual set inclusion, that any solution to the system is in the above set and that anything in the set is a solution of the system.*

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,

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

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,

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

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

* More information on set equality is in the appendix.

The two lemmas together establish Theorem 3.1. Remember that theorem with the slogan, "General = Particular + Homogeneous".

3.8 Example This system illustrates Theorem 3.1.

$$\begin{array}{rcl} x + 2y - z &=& 1 \\ 2x + 4y \phantom{{}-3z} &=& 2 \\ y - 3z &=& 0 \end{array}$$

Gauss's Method

$$\xrightarrow{-2\rho_1+\rho_2} \begin{array}{rcl} x + 2y - z &=& 1 \\ 2z &=& 0 \\ y - 3z &=& 0 \end{array} \xrightarrow{\rho_2\leftrightarrow\rho_3} \begin{array}{rcl} x + 2y - z &=& 1 \\ y - 3z &=& 0 \\ 2z &=& 0 \end{array}$$

shows that the general solution is a singleton set.

$$\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\}$$

That single vector is obviously a particular solution. The associated homogeneous system reduces via the same row operations

$$\begin{array}{rcl} x + 2y - z &=& 0 \\ 2x + 4y \phantom{{}-3z} &=& 0 \\ y - 3z &=& 0 \end{array} \xrightarrow{-2\rho_1+\rho_2} \xrightarrow{\rho_2\leftrightarrow\rho_3} \begin{array}{rcl} x + 2y - z &=& 0 \\ y - 3z &=& 0 \\ 2z &=& 0 \end{array}$$

to also give a singleton set.

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

So, as discussed at the start of this subsection, in this single-solution case the general solution results from taking the particular solution and adding to it the unique solution of the associated homogeneous system.

3.9 Example The start of this subsection also discusses that the case where the general solution set is empty fits the General = Particular + Homogeneous pattern too. This system illustrates.

$$\begin{array}{rcl} x \phantom{{}-y} + z + w &=& -1 \\ 2x - y \phantom{{}+3z} + w &=& 3 \\ x + y + 3z + 2w &=& 1 \end{array} \xrightarrow[-\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{array}{rcl} x \phantom{{}-y} + z + w &=& -1 \\ -y - 2z - w &=& 5 \\ y + 2z + w &=& 2 \end{array}$$

It has no solutions because the final two equations conflict. But the associated homogeneous system does have a solution, as do all homogeneous systems.

$$\begin{array}{rcl} x \phantom{{}-y} + z + w &=& 0 \\ 2x - y \phantom{{}+3z} + w &=& 0 \\ x + y + 3z + 2w &=& 0 \end{array} \xrightarrow[-\rho_1+\rho_3]{-2\rho_1+\rho_2} \xrightarrow{\rho_2+\rho_3} \begin{array}{rcl} x \phantom{{}-y} + z + w &=& 0 \\ -y - 2z - w &=& 0 \\ 0 &=& 0 \end{array}$$

In fact, the solution set is infinite.

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

Nonetheless, because the original system has no particular solution, its general solution set is empty — there are no vectors of the form $\vec{p}+\vec{h}$ because there are no $\vec{p}$'s.

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

Proof We've seen examples of all three happening so we need only prove that there are no other possibilities.

First observe a homogeneous system with at least one non-$\vec{0}$ solution $\vec{v}$ has infinitely many solutions. This is because any scalar multiple of $\vec{v}$ also solves the homogeneous system and there are infinitely many vectors in the set of scalar multiples of $\vec{v}$: if $s,t\in\mathbb{R}$ are unequal then $s\vec{v}\neq t\vec{v}$, since $s\vec{v}-t\vec{v}=(s-t)\vec{v}$ is non-$\vec{0}$ as any non-0 component of $\vec{v}$, when rescaled by the non-0 factor $s-t$, will give a non-0 value.

Now apply Lemma 3.7 to conclude that a solution set

$$\{\vec{p}+\vec{h} \mid \vec{h}\text{ solves the associated homogeneous system}\}$$

is either empty (if there is no particular solution $\vec{p}$), or has one element (if there is a $\vec{p}$ and the homogeneous system has the unique solution $\vec{0}$), or is infinite (if there is a $\vec{p}$ and the homogeneous system has a non-$\vec{0}$ solution, and thus by the prior paragraph has infinitely many solutions). QED

This table summarizes the factors affecting the size of a general solution.

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

The dimension on the top of the table is the simpler one. When we perform Gauss's Method on a linear system, ignoring the constants on the right side and so paying attention only to the coefficients on the left-hand side, we either end with every variable leading some row or else we find some variable that does not lead a row, that is, we find some variable that is free. (We formalize "ignoring the constants on the right" by considering the associated homogeneous system.)

A notable special case is systems having the same number of equations as unknowns. Such a system will have a solution, and that solution will be unique, if and only if it reduces to an echelon form system where every variable leads its row (since there are the same number of variables as rows), which will happen if and only if the associated homogeneous system has a unique solution.

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.

3.12 Example The first of these matrices is nonsingular while the second is singular

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

because the first of these homogeneous systems has a unique solution while the second has infinitely many solutions.

$$\begin{aligned} x + 2y &= 0 \\ 3x + 4y &= 0 \end{aligned} \qquad \begin{aligned} x + 2y &= 0 \\ 3x + 6y &= 0 \end{aligned}$$

We have made the distinction in the definition because a system with the same number of equations as variables behaves in one of two ways, depending on whether its matrix of coefficients is nonsingular or singular. Where the matrix of coefficients is nonsingular the system has a unique solution for any constants on the right side: for instance, Gauss's Method shows that this system

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

has the unique solution $x = b - 2a$ and $y = (3a - b)/2$. On the other hand, where the matrix of coefficients is singular the system never has a unique solution — it has either no solutions or else has infinitely many, as with these.

$$\begin{aligned} x + 2y &= 1 \\ 3x + 6y &= 2 \end{aligned} \qquad \begin{aligned} x + 2y &= 1 \\ 3x + 6y &= 3 \end{aligned}$$

The definition uses the word 'singular' because it means "departing from general expectation." People often, naively, expect that systems with the same number of variables as equations will have a unique solution. Thus, we can think of the word as connoting “troublesome,” or at least “not ideal.” (That ‘singular’ applies to those systems that never have exactly one solution is ironic, but it is the standard term.)

3.13 Example The systems from Example 3.3, Example 3.4, and Example 3.8 each have an associated homogeneous system with a unique solution. Thus these matrices are nonsingular.

$$\begin{pmatrix} 3 & 4 \\ 2 & -1 \end{pmatrix} \qquad \begin{pmatrix} 3 & 2 & 1 \\ 6 & -4 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & 0 \\ 0 & 1 & -3 \end{pmatrix}$$

The Chemistry problem from Example 3.5 is a homogeneous system with more than one solution so its matrix is singular.

$$\begin{pmatrix} 7 & 0 & -7 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix}$$

The table above has two dimensions. We have considered the one on top: we can tell into which column a given linear system goes solely by considering the system’s left-hand side; the constants on the right-hand side play no role in this.

The table’s other dimension, determining whether a particular solution exists, is tougher. Consider these two systems with the same left side but different right sides.

$$\begin{aligned} 3x + 2y &= 5 \\ 3x + 2y &= 5 \end{aligned} \qquad \begin{aligned} 3x + 2y &= 5 \\ 3x + 2y &= 4 \end{aligned}$$

The first has a solution while the second does not, so here the constants on the right side decide if the system has a solution. We could conjecture that the left side of a linear system determines the number of solutions while the right side determines if solutions exist but that guess is not correct. Compare these two, with the same right sides but different left sides.

$$\begin{aligned} 3x + 2y &= 5 \\ 4x + 2y &= 4 \end{aligned} \qquad \begin{aligned} 3x + 2y &= 5 \\ 3x + 2y &= 4 \end{aligned}$$

The first has a solution but the second does not. Thus the constants on the right side of a system don’t alone determine whether a solution exists. Rather, that depends on some interaction between the left and right.

For some intuition about that interaction, consider this system with one of the coefficients left unspecified, as the variable c.

$$\begin{aligned} x + 2y + 3z &= 1 \\ x + y + z &= 1 \\ cx + 3y + 4z &= 0 \end{aligned}$$

If $c = 2$ then this system has no solution because the left-hand side has the third row as the sum of the first two, while the right-hand does not. If $c \neq 2$ then this system has a unique solution (try it with $c = 1$). For a system to have a solution, if one row of the matrix of coefficients on the left is a linear combination of other rows then on the right the constant from that row must be the same combination of constants from the same rows.

More intuition about the interaction comes from studying linear combinations. That will be our focus in the second chapter, after we finish the study of Gauss's Method itself in the rest of this chapter.

Exercises

✓ **3.14** Solve each system. Express the solution set using vectors. Identify the particular solution and the solution set of the homogeneous system. (These systems also appear in Exercise 18.)

(a) $\begin{aligned} 3x + 6y &= 18 \\ x + 2y &= 6 \end{aligned}$ (b) $\begin{aligned} x + y &= 1 \\ x - y &= -1 \end{aligned}$ (c) $\begin{aligned} x_1 + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 5 \\ 4x_1 - x_2 + 5x_3 &= 17 \end{aligned}$

(d) $\begin{aligned} 2a + b - c &= 2 \\ 2a + c &= 3 \\ a - b &= 0 \end{aligned}$ (e) $\begin{aligned} x + 2y - z &= 3 \\ 2x + y + w &= 4 \\ x - y + z + w &= 1 \end{aligned}$ (f) $\begin{aligned} x + z + w &= 4 \\ 2x + y - w &= 2 \\ 3x + y + z &= 7 \end{aligned}$

3.15 Solve each system, giving the solution set in vector notation. Identify the particular solution and the solution of the homogeneous system.

(a) $\begin{aligned} 2x + y - z &= 1 \\ 4x - y &= 3 \end{aligned}$ (b) $\begin{aligned} x - z &= 1 \\ y + 2z - w &= 3 \\ x + 2y + 3z - w &= 7 \end{aligned}$ (c) $\begin{aligned} x - y + z &= 0 \\ y + w &= 0 \\ 3x - 2y + 3z + w &= 0 \\ -y - w &= 0 \end{aligned}$

(d) $\begin{aligned} a + 2b + 3c + d - e &= 1 \\ 3a - b + c + d + e &= 3 \end{aligned}$

✓ **3.16** For the system

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

which of these can be used as the particular solution part of some general solution?

(a) $\begin{pmatrix} 0 \\ -3 \\ 5 \\ 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ (c) $\begin{pmatrix} -1 \\ -4 \\ 8 \\ -1 \end{pmatrix}$

✓ **3.17** Lemma 3.7 says that we can use any particular solution for $\vec{p}$. Find, if possible, a general solution to this system

$$\begin{aligned} x - \;\; y \quad\;\; + w &= 4 \\ 2x + 3y - z \quad\;\; &= 0 \\ y + z + w &= 4 \end{aligned}$$

that uses the given vector as its particular solution.

(a) $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 4 \end{pmatrix}$ (b) $\begin{pmatrix} -5 \\ 1 \\ -7 \\ 10 \end{pmatrix}$ (c) $\begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \end{pmatrix}$

3.18 One is nonsingular while the other is singular. Which is which?

(a) $\begin{pmatrix} 1 & 3 \\ 4 & -12 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 3 \\ 4 & 12 \end{pmatrix}$

✓ **3.19** Singular or nonsingular?

(a) $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 \\ -3 & -6 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 1 \end{pmatrix}$ (Careful!)

(d) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 3 \\ 3 & 4 & 7 \end{pmatrix}$ (e) $\begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 5 \\ -1 & 1 & 4 \end{pmatrix}$

✓ **3.20** Is the given vector in the set generated by the given set?

(a) $\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \{\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \end{pmatrix}\}$

(b) $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \{\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\}$

(c) $\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \{\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}\}$

(d) $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \{\begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ 2 \end{pmatrix}\}$

3.21 Prove that any linear system with a nonsingular matrix of coefficients has a solution, and that the solution is unique.

3.22 In the proof of Lemma 3.6, what happens if there are no non-$0 = 0$ equations?

✓ **3.23** Prove that if $\vec{s}$ and $\vec{t}$ satisfy a homogeneous system then so do these vectors.

(a) $\vec{s} + \vec{t}$ (b) $3\vec{s}$ (c) $k\vec{s} + m\vec{t}$ for $k, m \in \mathbb{R}$

What's wrong with this argument: "These three show that if a homogeneous system has one solution then it has many solutions — any multiple of a solution is another solution, and any sum of solutions is a solution also — so there are no homogeneous systems with exactly one solution."?

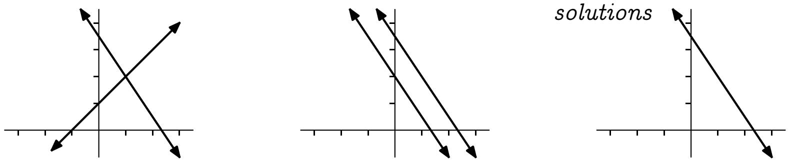
3.24 Prove that if a system with only rational coefficients and constants has a solution then it has at least one all-rational solution. Must it have infinitely many?

II Linear Geometry

If you have seen the elements of vectors then this section is an optional review. However, later work will refer to this material so if this is not a review then it is not optional.

In the first section we had to do a bit of work to show that there are only three types of solution sets — singleton, empty, and infinite. But this is easy to see geometrically in the case of systems with two equations and two unknowns. Draw each two-unknowns equation as a line in the plane, and then the two lines could have a unique intersection, be parallel, or be the same line.

Unique solution — *No solutions* — *Infinitely many solutions*



$$\begin{aligned} 3x+2y&=7\\ x-y&=-1 \end{aligned} \qquad \begin{aligned} 3x+2y&=7\\ 3x+2y&=4 \end{aligned} \qquad \begin{aligned} 3x+2y&=7\\ 6x+4y&=14 \end{aligned}$$

These pictures aren't a short way to prove the results from the prior section, because those results apply to linear systems of any size. But they do broaden our understanding of those results.

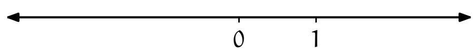
This section develops what we need to express our results geometrically. In particular, while the two-dimensional case is familiar, to extend to systems with more than two unknowns we shall need some higher-dimensional geometry.

II.1 Vectors in Space

"Higher-dimensional geometry" sounds exotic. It is exotic — interesting and eye-opening. But it isn't distant or unreachable.

We begin by defining one-dimensional space to be $\mathbb{R}$. To see that the definition is reasonable, picture a one-dimensional space

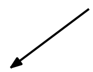
and pick a point to label 0 and another to label 1.



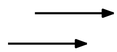
Now, with a scale and a direction, we have a correspondence with $\mathbb{R}$. For instance, to find the point matching +2.17, start at 0 and head in the direction of 1, and go 2.17 times as far.

The basic idea here, combining magnitude with direction, is the key to extending to higher dimensions.

An object comprised of a magnitude and a direction is a *vector* (we use the same word as in the prior section because we shall show below how to describe such an object with a column vector). We can draw a vector as having some length and pointing in some direction.

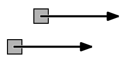


There is a subtlety involved in the definition of a vector as consisting of a magnitude and a direction — these

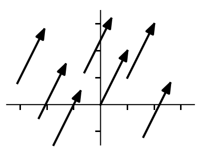


are equal, even though they start in different places They are equal because they have equal lengths and equal directions. Again: those vectors are not just alike, they are equal.

How can things that are in different places be equal? Think of a vector as representing a displacement (the word 'vector' is Latin for "carrier" or "traveler"). These two squares undergo displacements that are equal despite that they start in different places.



When we want to emphasize this property vectors have of not being anchored we refer to them as *free* vectors. Thus, these free vectors are equal, as each is a displacement of one over and two up.



More generally, vectors in the plane are the same if and only if they have the same change in first components and the same change in second components: the vector extending from (a_1, a_2) to (b_1, b_2) equals the vector from (c_1, c_2) to (d_1, d_2) if and only if $b_1 - a_1 = d_1 - c_1$ and $b_2 - a_2 = d_2 - c_2$.

Saying 'the vector that, were it to start at (a_1, a_2), would extend to (b_1, b_2)' would be unwieldy. We instead describe that vector as

$$\begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}$$

so that we represent the 'one over and two up' arrows shown above in this way.

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

We often draw the arrow as starting at the origin, and we then say it is in the *canonical position* (or *natural position* or *standard position*). When

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

is in canonical position then it extends from the origin to the endpoint (v_1, v_2).

We will typically say "the point

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}\text{"}$$

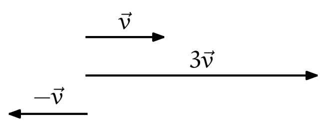
rather than "the endpoint of the canonical position of" that vector. Thus, we will call each of these $\mathbb{R}^2$.

$$\{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\} \qquad \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R}\}$$

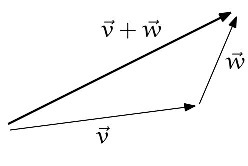
In the prior section we defined vectors and vector operations with an algebraic motivation;

$$r \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} rv_1 \\ rv_2 \end{pmatrix} \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}$$

we can now understand those operations geometrically. For instance, if $\vec{v}$ represents a displacement then $3\vec{v}$ represents a displacement in the same direction but three times as far and $-1\vec{v}$ represents a displacement of the same distance as $\vec{v}$ but in the opposite direction.

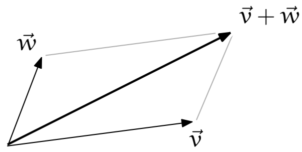


And, where $\vec{v}$ and $\vec{w}$ represent displacements, $\vec{v}+\vec{w}$ represents those displacements combined.



The long arrow is the combined displacement in this sense: imagine that you are walking on a ship's deck. Suppose that in one minute the ship's motion gives it a displacement relative to the sea of $\vec{v}$, and in the same minute your walking gives you a displacement relative to the ship's deck of $\vec{w}$. Then $\vec{v}+\vec{w}$ is your displacement relative to the sea.

Another way to understand the vector sum is with the *parallelogram rule*. Draw the parallelogram formed by the vectors $\vec{v}$ and $\vec{w}$. Then the sum $\vec{v}+\vec{w}$ extends along the diagonal to the far corner.



The above drawings show how vectors and vector operations behave in $\mathbb{R}^2$. We can extend to $\mathbb{R}^3$, or to even higher-dimensional spaces where we have no pictures, with the obvious generalization: the free vector that, if it starts at $(a_1, \ldots, a_n)$, ends at $(b_1, \ldots, b_n)$, is represented by this column.

$$\begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

Vectors are equal if they have the same representation. We aren't too careful about distinguishing between a point and the vector whose canonical representation ends at that point.

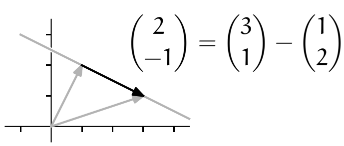
$$\mathbb{R}^n = \{ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mid v_1, \ldots, v_n \in \mathbb{R} \}$$

And, we do addition and scalar multiplication component-wise.

Having considered points, we next turn to lines. In $\mathbb{R}^2$, the line through $(1,2)$ and $(3,1)$ is comprised of (the endpoints of) the vectors in this set.

$$\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} 2 \\ -1 \end{pmatrix} \mid t \in \mathbb{R} \}$$

That description expresses this picture.

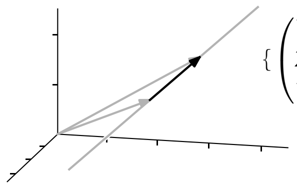


The vector that in the description is associated with the parameter t

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

is the one shown in the picture as having its whole body in the line — it is a *direction vector* for the line. Note that points on the line to the left of $x = 1$ are described using negative values of t.

In $\mathbb{R}^3$, the line through $(1, 2, 1)$ and $(2, 3, 2)$ is the set of (endpoints of) vectors of this form



$$\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \}$$

and lines in even higher-dimensional spaces work in the same way.

In $\mathbb{R}^3$, a line uses one parameter so that a particle on that line would be free to move back and forth in one dimension. A plane involves two parameters. For example, the plane through the points $(1, 0, 5)$, $(2, 1, -3)$, and $(-2, 4, 0.5)$ consists of (endpoints of) the vectors in this set.

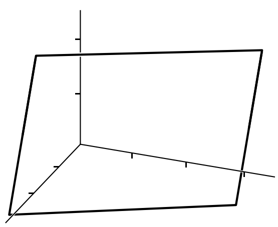
$$\{ \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ -8 \end{pmatrix} + s \begin{pmatrix} -3 \\ 4 \\ -4.5 \end{pmatrix} \mid t, s \in \mathbb{R} \}$$

The column vectors associated with the parameters come from these calculations.

$$\begin{pmatrix} 1 \\ 1 \\ -8 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} \qquad \begin{pmatrix} -3 \\ 4 \\ -4.5 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}$$

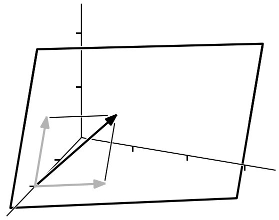
As with the line, note that we describe some points in this plane with negative t's or negative s's or both.

Calculus books often describe a plane by using a single linear equation.



$$P = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid 2x + y + z = 4 \}$$

To translate from this to the vector description, think of this as a one-equation linear system and parametrize: $x = 2 - y/2 - z/2$.



$$P = \{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} -1/2 \\ 0 \\ 1 \end{pmatrix} \mid y, z \in \mathbb{R} \}$$

Shown in grey are the vectors associated with y and z, offset from the origin by 2 units along the x-axis, so that their entire body lies in the plane. Thus the vector sum of the two, shown in black, has its entire body in the plane along with the rest of the parallelogram.

Generalizing, a set of the form $\{ \vec{p} + t_1\vec{v}_1 + t_2\vec{v}_2 + \cdots + t_k\vec{v}_k \mid t_1, \ldots, t_k \in \mathbb{R} \}$ where $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ and $k \leqslant n$ is a k-*dimensional linear surface* (or k-*flat*). For example, in $\mathbb{R}^4$

$$\{ \begin{pmatrix} 2 \\ \pi \\ 3 \\ -0.5 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mid t \in \mathbb{R} \}$$

is a line,

$$\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + s \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \mid t, s \in \mathbb{R} \}$$

is a plane, and

$$\{ \begin{pmatrix} 3 \\ 1 \\ -2 \\ 0.5 \end{pmatrix} + r \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \mid r, s, t \in \mathbb{R} \}$$

is a three-dimensional linear surface. Again, the intuition is that a line permits motion in one direction, a plane permits motion in combinations of two directions, etc. When the dimension of the linear surface is one less than the dimension of the space, that is, when in $\mathbb{R}^n$ we have an $(n-1)$-flat, the surface is called a *hyperplane*.

A description of a linear surface can be misleading about the dimension. For example, this

$$L = \{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} + s \begin{pmatrix} 2 \\ 2 \\ 0 \\ -2 \end{pmatrix} \mid t, s \in \mathbb{R} \}$$

is a *degenerate* plane because it is actually a line, since the vectors are multiples of each other and we can omit one.

$$L = \{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix} + r \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \mid r \in \mathbb{R} \}$$

We shall see in the Linear Independence section of Chapter Two what relationships among vectors causes the linear surface they generate to be degenerate.

We now can restate in geometric terms our conclusions from earlier. First, the solution set of a linear system with n unknowns is a linear surface in $\mathbb{R}^n$. Specifically, it is a k-dimensional linear surface, where k is the number of free variables in an echelon form version of the system. For instance, in the single equation case the solution set is an $n-1$-dimensional hyperplane in $\mathbb{R}^n$ (where $n > 0$). Second, the solution set of a homogeneous linear system is a linear surface passing through the origin. Finally, we can view the general solution set of any linear system as being the solution set of its associated homogeneous system offset from the origin by a vector, namely by any particular solution.

Exercises

✓ 1.1 Find the canonical name for each vector.
- **(a)** the vector from $(2, 1)$ to $(4, 2)$ in $\mathbb{R}^2$
- **(b)** the vector from $(3, 3)$ to $(2, 5)$ in $\mathbb{R}^2$
- **(c)** the vector from $(1, 0, 6)$ to $(5, 0, 3)$ in $\mathbb{R}^3$
- **(d)** the vector from $(6, 8, 8)$ to $(6, 8, 8)$ in $\mathbb{R}^3$

✓ 1.2 Decide if the two vectors are equal.
- **(a)** the vector from $(5, 3)$ to $(6, 2)$ and the vector from $(1, -2)$ to $(1, 1)$
- **(b)** the vector from $(2, 1, 1)$ to $(3, 0, 4)$ and the vector from $(5, 1, 4)$ to $(6, 0, 7)$

✓ 1.3 Does $(1, 0, 2, 1)$ lie on the line through $(-2, 1, 1, 0)$ and $(5, 10, -1, 4)$?

✓ 1.4 **(a)** Describe the plane through $(1, 1, 5, -1)$, $(2, 2, 2, 0)$, and $(3, 1, 0, 4)$.
- **(b)** Is the origin in that plane?

1.5 Describe the plane that contains this point and line.

$$\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \qquad \{ \begin{pmatrix} -1 \\ 0 \\ -4 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} t \mid t \in \mathbb{R} \}$$

✓ 1.6 Intersect these planes.

$$\{\begin{pmatrix}1\\1\\1\end{pmatrix}t+\begin{pmatrix}0\\1\\3\end{pmatrix}s \mid t,s\in\mathbb{R}\}\qquad \{\begin{pmatrix}1\\1\\0\end{pmatrix}+\begin{pmatrix}0\\3\\0\end{pmatrix}k+\begin{pmatrix}2\\0\\4\end{pmatrix}m \mid k,m\in\mathbb{R}\}$$

✓ 1.7 Intersect each pair, if possible.

(a) $\{\begin{pmatrix}1\\1\\2\end{pmatrix}+t\begin{pmatrix}0\\1\\1\end{pmatrix} \mid t\in\mathbb{R}\}$, $\{\begin{pmatrix}1\\3\\-2\end{pmatrix}+s\begin{pmatrix}0\\1\\2\end{pmatrix} \mid s\in\mathbb{R}\}$

(b) $\{\begin{pmatrix}2\\0\\1\end{pmatrix}+t\begin{pmatrix}1\\1\\-1\end{pmatrix} \mid t\in\mathbb{R}\}$, $\{s\begin{pmatrix}0\\1\\2\end{pmatrix}+w\begin{pmatrix}0\\4\\1\end{pmatrix} \mid s,w\in\mathbb{R}\}$

1.8 When a plane does not pass through the origin, performing operations on vectors whose bodies lie in it is more complicated than when the plane passes through the origin. Consider the picture in this subsection of the plane

$$\{\begin{pmatrix}2\\0\\0\end{pmatrix}+\begin{pmatrix}-0.5\\1\\0\end{pmatrix}y+\begin{pmatrix}-0.5\\0\\1\end{pmatrix}z \mid y,z\in\mathbb{R}\}$$

and the three vectors with endpoints $(2,0,0)$, $(1.5,1,0)$, and $(1.5,0,1)$.

(a) Redraw the picture, including the vector in the plane that is twice as long as the one with endpoint $(1.5,1,0)$. The endpoint of your vector is not $(3,2,0)$; what is it?

(b) Redraw the picture, including the parallelogram in the plane that shows the sum of the vectors ending at $(1.5,0,1)$ and $(1.5,1,0)$. The endpoint of the sum, on the diagonal, is not $(3,1,1)$; what is it?

1.9 Show that the line segments $\overline{(a_1,a_2)(b_1,b_2)}$ and $\overline{(c_1,c_2)(d_1,d_2)}$ have the same lengths and slopes if $b_1-a_1=d_1-c_1$ and $b_2-a_2=d_2-c_2$. Is that only if?

1.10 How should we define $\mathbb{R}^0$?

? ✓ 1.11 [Math. Mag., Jan. 1957] A person traveling eastward at a rate of 3 miles per hour finds that the wind appears to blow directly from the north. On doubling his speed it appears to come from the north east. What was the wind's velocity?

1.12 Euclid describes a plane as "a surface which lies evenly with the straight lines on itself". Commentators such as Heron have interpreted this to mean, "(A plane surface is) such that, if a straight line pass through two points on it, the line coincides wholly with it at every spot, all ways". (Translations from [Heath], pp. 171-172.) Do planes, as described in this section, have that property? Does this description adequately define planes?

II.2 Length and Angle Measures

We've translated the first section's results about solution sets into geometric terms, to better understand those sets. But we must be careful not to be misled by our own terms — labeling subsets of $\mathbb{R}^k$ of the forms $\{\vec{p}+t\vec{v} \mid t \in \mathbb{R}\}$ and $\{\vec{p}+t\vec{v}+s\vec{w} \mid t, s \in \mathbb{R}\}$ as 'lines' and 'planes' doesn't make them act like the lines and planes of our past experience. Rather, we must ensure that the names suit the sets. While we can't prove that the sets satisfy our intuition — we can't prove anything about intuition — in this subsection we'll observe that a result familiar from $\mathbb{R}^2$ and $\mathbb{R}^3$, when generalized to arbitrary $\mathbb{R}^n$, supports the idea that a line is straight and a plane is flat. Specifically, we'll see how to do Euclidean geometry in a 'plane' by giving a definition of the angle between two $\mathbb{R}^n$ vectors, in the plane that they generate.

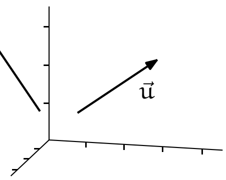
2.1 Definition The *length* of a vector $\vec{v} \in \mathbb{R}^n$ is the square root of the sum of the squares of its components.

$$|\vec{v}\,| = \sqrt{v_1^2 + \cdots + v_n^2}$$

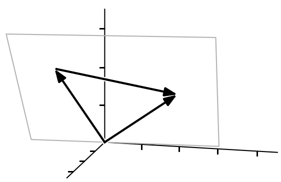
2.2 Remark This is a natural generalization of the Pythagorean Theorem. A classic motivating discussion is in [Polya].

For any nonzero $\vec{v}$, the vector $\vec{v}/|\vec{v}|$ has length one. We say that the second *normalizes* $\vec{v}$ to length one.

We can use that to get a formula for the angle between two vectors. Consider two vectors in $\mathbb{R}^3$ where neither is a multiple of the other



(the special case of multiples will turn out below not to be an exception). They determine a two-dimensional plane — for instance, put them in canonical position and take the plane formed by the origin and the endpoints. In that plane consider the triangle with sides $\vec{u}$, $\vec{v}$, and $\vec{u}-\vec{v}$.



Apply the Law of Cosines: $|\vec{u}-\vec{v}\,|^2 = |\vec{u}\,|^2 + |\vec{v}\,|^2 - 2\,|\vec{u}\,|\,|\vec{v}\,|\cos\theta$ where θ is the angle between the vectors. The left side gives

$$(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2 \\ = (u_1^2 - 2u_1v_1 + v_1^2) + (u_2^2 - 2u_2v_2 + v_2^2) + (u_3^2 - 2u_3v_3 + v_3^2)$$

while the right side gives this.

$$(u_1^2 + u_2^2 + u_3^2) + (v_1^2 + v_2^2 + v_3^2) - 2\,|\vec{u}\,|\,|\vec{v}\,|\cos\theta$$

Canceling squares u_1^2, ..., v_3^2 and dividing by 2 gives a formula for the angle.

$$\theta = \arccos\Big(\frac{u_1v_1 + u_2v_2 + u_3v_3}{|\vec{u}\,|\,|\vec{v}\,|}\Big)$$

In higher dimensions we cannot draw pictures as above but we can instead make the argument analytically. First, the form of the numerator is clear; it comes from the middle terms of $(u_i - v_i)^2$.

2.3 Definition The *dot product* (or *inner product* or *scalar product*) of two n-component real vectors is the linear combination of their components.

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Note that the dot product of two vectors is a real number, not a vector, and that the dot product is only defined if the two vectors have the same number of components. Note also that dot product is related to length: $\vec{u} \cdot \vec{u} = u_1u_1 + \cdots + u_nu_n = |\vec{u}\,|^2$.

2.4 Remark Some authors require that the first vector be a row vector and that the second vector be a column vector. We shall not be that strict and will allow the dot product operation between two column vectors.

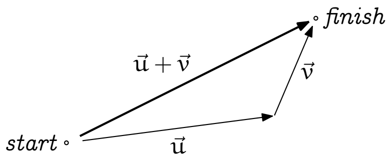
Still reasoning analytically but guided by the pictures, we use the next theorem to argue that the triangle formed by the line segments making the bodies of $\vec{u}$, $\vec{v}$, and $\vec{u} + \vec{v}$ in $\mathbb{R}^n$ lies in the planar subset of $\mathbb{R}^n$ generated by $\vec{u}$ and $\vec{v}$ (see the figure below).

2.5 Theorem (Triangle Inequality) For any $\vec{u}, \vec{v} \in \mathbb{R}^n$,

$$|\vec{u} + \vec{v}\,| \leqslant |\vec{u}\,| + |\vec{v}\,|$$

with equality if and only if one of the vectors is a nonnegative scalar multiple of the other one.

This is the source of the familiar saying, "The shortest distance between two points is in a straight line."



Proof (We'll use some algebraic properties of dot product that we have not yet checked, for instance that $\vec{u}\cdot(\vec{a}+\vec{b})=\vec{u}\cdot\vec{a}+\vec{u}\cdot\vec{b}$ and that $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$. See Exercise 18.) Since all the numbers are positive, the inequality holds if and only if its square holds.

$$\begin{aligned} |\vec{u}+\vec{v}|^2 &\leqslant (\,|\vec{u}\,|+|\vec{v}\,|\,)^2 \\ (\,\vec{u}+\vec{v}\,)\cdot(\,\vec{u}+\vec{v}\,) &\leqslant |\vec{u}\,|^2+2\,|\vec{u}\,|\,|\vec{v}\,|+|\vec{v}\,|^2 \\ \vec{u}\cdot\vec{u}+\vec{u}\cdot\vec{v}+\vec{v}\cdot\vec{u}+\vec{v}\cdot\vec{v} &\leqslant \vec{u}\cdot\vec{u}+2\,|\vec{u}\,|\,|\vec{v}\,|+\vec{v}\cdot\vec{v} \\ 2\,\vec{u}\cdot\vec{v} &\leqslant 2\,|\vec{u}\,|\,|\vec{v}\,| \end{aligned}$$

That, in turn, holds if and only if the relationship obtained by multiplying both sides by the nonnegative numbers $|\vec{u}\,|$ and $|\vec{v}\,|$

$$2\,(\,|\vec{v}\,|\,\vec{u}\,)\cdot(\,|\vec{u}\,|\,\vec{v}\,)\leqslant 2\,|\vec{u}\,|^2\,|\vec{v}\,|^2$$

and rewriting

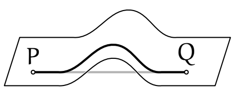
$$0\leqslant |\vec{u}\,|^2\,|\vec{v}\,|^2-2\,(\,|\vec{v}\,|\,\vec{u}\,)\cdot(\,|\vec{u}\,|\,\vec{v}\,)+|\vec{u}\,|^2\,|\vec{v}\,|^2$$

is true. But factoring shows that it is true

$$0\leqslant (\,|\vec{u}\,|\,\vec{v}-|\vec{v}\,|\,\vec{u}\,)\cdot(\,|\vec{u}\,|\,\vec{v}-|\vec{v}\,|\,\vec{u}\,)$$

since it only says that the square of the length of the vector $|\vec{u}\,|\,\vec{v}-|\vec{v}\,|\,\vec{u}$ is not negative. As for equality, it holds when, and only when, $|\vec{u}\,|\,\vec{v}-|\vec{v}\,|\,\vec{u}$ is $\vec{0}$. The check that $|\vec{u}\,|\,\vec{v}=|\vec{v}\,|\,\vec{u}$ if and only if one vector is a nonnegative real scalar multiple of the other is easy. QED

This result supports the intuition that even in higher-dimensional spaces, lines are straight and planes are flat. We can easily check from the definition that linear surfaces have the property that for any two points in that surface, the line segment between them is contained in that surface. But if the linear surface were not flat then that would allow for a shortcut.



Because the Triangle Inequality says that in any $\mathbb{R}^n$ the shortest cut between two endpoints is simply the line segment connecting them, linear surfaces have no bends.

Back to the definition of angle measure. The heart of the Triangle Inequality's proof is the $\vec{u} \cdot \vec{v} \leqslant |\vec{u}\,||\vec{v}\,|$ line. We might wonder if some pairs of vectors satisfy the inequality in this way: while $\vec{u} \cdot \vec{v}$ is a large number, with absolute value bigger than the right-hand side, it is a negative large number. The next result says that does not happen.

2.6 Corollary (Cauchy-Schwarz Inequality) For any $\vec{u}, \vec{v} \in \mathbb{R}^n$,

$$|\,\vec{u} \cdot \vec{v}\,| \leqslant |\,\vec{u}\,||\vec{v}\,|$$

with equality if and only if one vector is a scalar multiple of the other.

PROOF The Triangle Inequality's proof shows that $\vec{u} \cdot \vec{v} \leqslant |\vec{u}\,||\vec{v}\,|$ so if $\vec{u} \cdot \vec{v}$ is positive or zero then we are done. If $\vec{u} \cdot \vec{v}$ is negative then this holds.

$$|\,\vec{u} \cdot \vec{v}\,| = -(\,\vec{u} \cdot \vec{v}\,) = (-\vec{u}\,) \cdot \vec{v} \leqslant |-\vec{u}\,||\vec{v}\,| = |\vec{u}\,||\vec{v}\,|$$

The equality condition is Exercise 19. QED

The Cauchy-Schwarz inequality assures us that the next definition makes sense because the fraction has absolute value less than or equal to one.

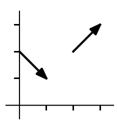
2.7 Definition The *angle* between two nonzero vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is

$$\theta = \arccos(\,\frac{\vec{u} \cdot \vec{v}}{|\vec{u}\,||\vec{v}\,|}\,)$$

(if either is the zero vector, we take the angle to be right).

2.8 Corollary Vectors from $\mathbb{R}^n$ are orthogonal, that is, perpendicular, if and only if their dot product is zero. They are parallel if and only if their dot product equals the product of their lengths.

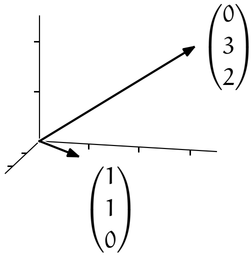
2.9 Example These vectors are orthogonal.



$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

We've drawn the arrows away from canonical position but nevertheless the vectors are orthogonal.

2.10 Example The $\mathbb{R}^3$ angle formula given at the start of this subsection is a special case of the definition. Between these two



the angle is

$$\arccos\left(\frac{(1)(0)+(1)(3)+(0)(2)}{\sqrt{1^2+1^2+0^2}\sqrt{0^2+3^2+2^2}}\right)=\arccos\left(\frac{3}{\sqrt{2}\sqrt{13}}\right)$$

approximately 0.94 radians. Notice that these vectors are not orthogonal. Although the yz-plane may appear to be perpendicular to the xy-plane, in fact the two planes are that way only in the weak sense that there are vectors in each orthogonal to all vectors in the other. Not every vector in each is orthogonal to all vectors in the other.

Exercises

✓ 2.11 Find the length of each vector.

(a) $\begin{pmatrix}3\\1\end{pmatrix}$ **(b)** $\begin{pmatrix}-1\\2\end{pmatrix}$ **(c)** $\begin{pmatrix}4\\1\\1\end{pmatrix}$ **(d)** $\begin{pmatrix}0\\0\\0\end{pmatrix}$ **(e)** $\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}$

✓ 2.12 Find the angle between each two, if it is defined.

(a) $\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}1\\4\end{pmatrix}$ **(b)** $\begin{pmatrix}1\\2\\0\end{pmatrix},\begin{pmatrix}0\\4\\1\end{pmatrix}$ **(c)** $\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}1\\4\\-1\end{pmatrix}$

✓ 2.13 [Ohanian] During maneuvers preceding the Battle of Jutland, the British battle cruiser *Lion* moved as follows (in nautical miles): 1.2 miles north, 6.1 miles 38 degrees east of south, 4.0 miles at 89 degrees east of north, and 6.5 miles at 31 degrees east of north. Find the distance between starting and ending positions. (Ignore the earth's curvature.)

2.14 Find k so that these two vectors are perpendicular.

$$\begin{pmatrix}k\\1\end{pmatrix}\qquad\begin{pmatrix}4\\3\end{pmatrix}$$

2.15 Describe the set of vectors in $\mathbb{R}^3$ orthogonal to this one.

$$\begin{pmatrix}1\\3\\-1\end{pmatrix}$$

✓ 2.16 **(a)** Find the angle between the diagonal of the unit square in $\mathbb{R}^2$ and one of the axes.

(b) Find the angle between the diagonal of the unit cube in $\mathbb{R}^3$ and one of the axes.

(c) Find the angle between the diagonal of the unit cube in $\mathbb{R}^n$ and one of the axes.

(d) What is the limit, as n goes to ∞, of the angle between the diagonal of the unit cube in $\mathbb{R}^n$ and one of the axes?

2.17 Is any vector perpendicular to itself?

✓ 2.18 Describe the algebraic properties of dot product.

(a) Is it right-distributive over addition: $(\vec{u}+\vec{v})\cdot\vec{w}=\vec{u}\cdot\vec{w}+\vec{v}\cdot\vec{w}$?

(b) Is it left-distributive (over addition)?

(c) Does it commute?

(d) Associate?

(e) How does it interact with scalar multiplication?

As always, you must back any assertion with either a proof or an example.

2.19 Verify the equality condition in Corollary 2.6, the Cauchy-Schwarz Inequality.

(a) Show that if $\vec{u}$ is a negative scalar multiple of $\vec{v}$ then $\vec{u}\cdot\vec{v}$ and $\vec{v}\cdot\vec{u}$ are less than or equal to zero.

(b) Show that $|\vec{u}\cdot\vec{v}|=|\vec{u}\,||\vec{v}\,|$ if and only if one vector is a scalar multiple of the other.

2.20 Suppose that $\vec{u}\cdot\vec{v}=\vec{u}\cdot\vec{w}$ and $\vec{u}\neq\vec{0}$. Must $\vec{v}=\vec{w}$?

✓ 2.21 Does any vector have length zero except a zero vector? (If “yes”, produce an example. If “no”, prove it.)

✓ 2.22 Find the midpoint of the line segment connecting (x_1,y_1) with (x_2,y_2) in $\mathbb{R}^2$. Generalize to $\mathbb{R}^n$.

2.23 Show that if $\vec{v}\neq\vec{0}$ then $\vec{v}/|\vec{v}\,|$ has length one. What if $\vec{v}=\vec{0}$?

2.24 Show that if $r\geqslant 0$ then $r\vec{v}$ is r times as long as $\vec{v}$. What if $r<0$?

✓ 2.25 A vector $\vec{v}\in\mathbb{R}^n$ of length one is a *unit* vector. Show that the dot product of two unit vectors has absolute value less than or equal to one. Can ‘less than’ happen? Can ‘equal to’?

2.26 Prove that $|\vec{u}+\vec{v}\,|^2+|\vec{u}-\vec{v}\,|^2=2|\vec{u}\,|^2+2|\vec{v}\,|^2$.

2.27 Show that if $\vec{x}\cdot\vec{y}=0$ for every $\vec{y}$ then $\vec{x}=\vec{0}$.

2.28 Is $|\vec{u}_1+\cdots+\vec{u}_n|\leqslant|\vec{u}_1|+\cdots+|\vec{u}_n|$? If it is true then it would generalize the Triangle Inequality.

2.29 What is the ratio between the sides in the Cauchy-Schwarz inequality?

2.30 Why is the zero vector defined to be perpendicular to every vector?

2.31 Describe the angle between two vectors in $\mathbb{R}^1$.

2.32 Give a simple necessary and sufficient condition to determine whether the angle between two vectors is acute, right, or obtuse.

✓ 2.33 Generalize to $\mathbb{R}^n$ the converse of the Pythagorean Theorem, that if $\vec{u}$ and $\vec{v}$ are perpendicular then $|\vec{u}+\vec{v}\,|^2=|\vec{u}\,|^2+|\vec{v}\,|^2$.

2.34 Show that $|\vec{u}\,|=|\vec{v}\,|$ if and only if $\vec{u}+\vec{v}$ and $\vec{u}-\vec{v}$ are perpendicular. Give an example in $\mathbb{R}^2$.

2.35 Show that if a vector is perpendicular to each of two others then it is perpendicular to each vector in the plane they generate. (*Remark.* They could generate a degenerate plane — a line or a point — but the statement remains true.)

2.36 Prove that, where $\vec{u}, \vec{v} \in \mathbb{R}^n$ are nonzero vectors, the vector

$$\frac{\vec{u}}{|\vec{u}|} + \frac{\vec{v}}{|\vec{v}|}$$

bisects the angle between them. Illustrate in $\mathbb{R}^2$.

2.37 Verify that the definition of angle is dimensionally correct: (1) if $k > 0$ then the cosine of the angle between $k\vec{u}$ and $\vec{v}$ equals the cosine of the angle between $\vec{u}$ and $\vec{v}$, and (2) if $k < 0$ then the cosine of the angle between $k\vec{u}$ and $\vec{v}$ is the negative of the cosine of the angle between $\vec{u}$ and $\vec{v}$.

✓ 2.38 Show that the inner product operation is *linear*: for $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and $k, m \in \mathbb{R}$, $\vec{u} \cdot (k\vec{v} + m\vec{w}) = k(\vec{u} \cdot \vec{v}) + m(\vec{u} \cdot \vec{w})$.

✓ 2.39 The *geometric mean* of two positive reals x, y is $\sqrt{xy}$. It is analogous to the *arithmetic mean* $(x + y)/2$. Use the Cauchy-Schwarz inequality to show that the geometric mean of any $x, y \in \mathbb{R}$ is less than or equal to the arithmetic mean.

? 2.40 [Cleary] Astrologers claim to be able to recognize trends in personality and fortune that depend on an individual's birthday by somehow incorporating where the stars were 2000 years ago, during the Hellenistic period. Suppose that instead of star-gazers coming up with stuff, math teachers who like linear algebra (we'll call them vectologers) had come up with a similar system as follows: Consider your birthday as a row vector $(\text{month} \;\; \text{day})$. For instance, I was born on July 12 so my vector would be $(7 \;\; 12)$. Vectologers have made the rule that how well individuals get along with each other depends on the angle between vectors. The smaller the angle, the more harmonious the relationship.

(a) Compute the angle between your vector and mine, expressing the answer in radians.

(b) Would you get along better with me, or with a professor born on September 19?

(c) For maximum harmony in a relationship, when should the other person be born?

(d) Is there a person with whom you have a "worst case" relationship, i.e., your vector and theirs are orthogonal? If so, what are the birthdate(s) for such people? If not, explain why not.

? 2.41 [Am. Math. Mon., Feb. 1933] A ship is sailing with speed and direction $\vec{v}_1$; the wind blows apparently (judging by the vane on the mast) in the direction of a vector $\vec{a}$; on changing the direction and speed of the ship from $\vec{v}_1$ to $\vec{v}_2$ the apparent wind is in the direction of a vector $\vec{b}$.

Find the vector velocity of the wind.

2.42 Verify the Cauchy-Schwarz inequality by first proving Lagrange's identity:

$$\left(\sum_{1 \leqslant j \leqslant n} a_j b_j\right)^2 = \left(\sum_{1 \leqslant j \leqslant n} a_j^2\right)\left(\sum_{1 \leqslant j \leqslant n} b_j^2\right) - \sum_{1 \leqslant k < j \leqslant n} (a_k b_j - a_j b_k)^2$$

and then noting that the final term is positive. (Recall the meaning

$$\sum_{1 \leqslant j \leqslant n} a_j b_j = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

and

$$\sum_{1 \leqslant j \leqslant n} a_j{}^2 = a_1{}^2 + a_2{}^2 + \cdots + a_n{}^2$$

of the Σ notation.) This result is an improvement over Cauchy-Schwarz because it gives a formula for the difference between the two sides. Interpret that difference in $\mathbb{R}^2$.

III Reduced Echelon Form

After developing the mechanics of Gauss's Method, we observed that it can be done in more than one way. For example, from this matrix

$$\begin{pmatrix} 2 & 2 \\ 4 & 3 \end{pmatrix}$$

we could derive any of these three echelon form matrices.

$$\begin{pmatrix} 2 & 2 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

The first results from $-2\rho_1 + \rho_2$. The second comes from following $(1/2)\rho_1$ with $-4\rho_1 + \rho_2$. The third comes from $-2\rho_1 + \rho_2$ followed by $2\rho_2 + \rho_1$ (after the first row combination the matrix is already in echelon form so the second one is extra work but it is nonetheless a legal row operation).

The fact that echelon form is not unique raises questions. Will any two echelon form versions of a linear system have the same number of free variables? If yes, will the two have exactly the same free variables? In this section we will give a way to decide if one linear system can be derived from another by row operations. The answers to both questions, both "yes," will follow from this.

III.1 Gauss-Jordan Reduction

Here is an extension of Gauss's Method that has some advantages.

1.1 Example To solve

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

we can start as usual by reducing it to echelon form.

$$\xrightarrow{-\rho_1+\rho_3} \left(\begin{array}{ccc|c} 1 & 1 & -2 & -2 \\ 0 & 1 & 3 & 7 \\ 0 & -1 & 1 & 1 \end{array}\right) \xrightarrow{\rho_2+\rho_3} \left(\begin{array}{ccc|c} 1 & 1 & -2 & -2 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 4 & 8 \end{array}\right)$$

We can keep going to a second stage by making the leading entries into 1's

$$\xrightarrow{(1/4)\rho_3} \left(\begin{array}{ccc|c} 1 & 1 & -2 & -2 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

and then to a third stage that uses the leading entries to eliminate all of the other entries in each column by combining upwards.

$$\xrightarrow[2\rho_3+\rho_1]{-3\rho_3+\rho_2} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow{-\rho_2+\rho_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

The answer is $x = 1$, $y = 1$, and $z = 2$.

Using one entry to clear out the rest of a column is *pivoting* on that entry.

Note that the row combination operations in the first stage move left to right, from column one to column three, while the combination operations in the third stage move right to left.

1.2 Example The middle stage operations that turn the leading entries into 1's don't interact, so we can combine multiple ones into a single step.

$$\begin{aligned} \left(\begin{array}{cc|c} 2 & 1 & 7 \\ 4 & -2 & 6 \end{array}\right) &\xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cc|c} 2 & 1 & 7 \\ 0 & -4 & -8 \end{array}\right) \\ &\xrightarrow[(-1/4)\rho_2]{(1/2)\rho_1} \left(\begin{array}{cc|c} 1 & 1/2 & 7/2 \\ 0 & 1 & 2 \end{array}\right) \\ &\xrightarrow{-(1/2)\rho_2+\rho_1} \left(\begin{array}{cc|c} 1 & 0 & 5/2 \\ 0 & 1 & 2 \end{array}\right) \end{aligned}$$

The answer is $x = 5/2$ and $y = 2$.

This extension of Gauss's Method is the *Gauss-Jordan Method* or *Gauss-Jordan reduction*.

1.3 Definition A matrix or linear system is in *reduced echelon form* if, in addition to being in echelon form, each leading entry is a 1 and is the only nonzero entry in its column.

The cost of using Gauss-Jordan reduction to solve a system is the additional arithmetic. The benefit is that we can just read off the solution set description.

More specifically, in any echelon form, reduced or not, we can read off when the system has an empty solution set because there is a contradictory equation. We can read off when the system has a one-element solution set because there is no contradiction and every variable is the leading variable in some row. And, we can read off when the system has an infinite solution set because there is no contradiction and at least one variable is free.

However, in reduced echelon form we can read off not just the size of the solution set but also its description. We have no trouble describing the solution set when it is empty, of course. Example 1.1 and 1.2 show how in a single element solution set case the single element is in the column of constants. The next example shows how to read the parametrization of an infinite solution set.

1.4 Example

$$\left(\begin{array}{cccc|c} 2 & 6 & 1 & 2 & 5 \\ 0 & 3 & 1 & 4 & 1 \\ 0 & 3 & 1 & 2 & 5 \end{array}\right) \xrightarrow{-\rho_2+\rho_3} \left(\begin{array}{cccc|c} 2 & 6 & 1 & 2 & 5 \\ 0 & 3 & 1 & 4 & 1 \\ 0 & 0 & 0 & -2 & 4 \end{array}\right)$$

$$\xrightarrow[\substack{(1/3)\rho_2 \\ -(1/2)\rho_3}]{(1/2)\rho_1} \xrightarrow[-\rho_3+\rho_1]{-(4/3)\rho_3+\rho_2} \xrightarrow{-3\rho_2+\rho_1} \left(\begin{array}{cccc|c} 1 & 0 & -1/2 & 0 & -9/2 \\ 0 & 1 & 1/3 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

As a linear system this is

$$\begin{aligned} x_1 \quad - 1/2x_3 \quad &= -9/2 \\ x_2 + 1/3x_3 \quad &= 3 \\ x_4 &= -2 \end{aligned}$$

so a solution set description is this.

$$S = \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -9/2 \\ 3 \\ 0 \\ -2 \end{pmatrix} + \begin{pmatrix} 1/2 \\ -1/3 \\ 1 \\ 0 \end{pmatrix} x_3 \mid x_3 \in \mathbb{R} \}$$

Thus echelon form isn't some kind of one best form for systems. Other forms, such as reduced echelon form, have advantages and disadvantages. Instead of picturing linear systems (and the associated matrices) as things we operate on, always directed toward the goal of echelon form, we can think of them as interrelated when we can get from one to another by row operations. The rest of this subsection develops this relationship.

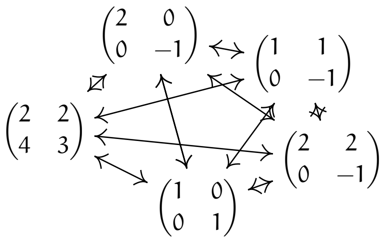
1.5 Lemma Elementary row operations are reversible.

Proof For any matrix A, the effect of swapping rows is reversed by swapping them back, multiplying a row by a nonzero k is undone by multiplying by $1/k$, and adding a multiple of row i to row j (with $i \neq j$) is undone by subtracting the same multiple of row i from row j.

$$A \xrightarrow{\rho_i \leftrightarrow \rho_j} \xrightarrow{\rho_j \leftrightarrow \rho_i} A \qquad A \xrightarrow{k\rho_i} \xrightarrow{(1/k)\rho_i} A \qquad A \xrightarrow{k\rho_i+\rho_j} \xrightarrow{-k\rho_i+\rho_j} A$$

(We need the $i \neq j$ condition; see Exercise 15.) QED

Again, the point of view that we are developing, supported now by the lemma, is that the term ‘reduces to’ is misleading: where $A \longrightarrow B$, we shouldn’t think of B as after A or simpler than A. Instead we should think of the two matrices as interrelated. Below is a picture. It shows the matrices from the start of this section and their reduced echelon form version in a cluster, as inter-reducible.



We say that matrices that reduce to each other are equivalent with respect to the relationship of row reducibility. The next result justifies this, using the definition of an equivalence.*

1.6 Lemma Between matrices, ‘reduces to’ is an equivalence relation.

Proof We must check the conditions (i) reflexivity, that any matrix reduces to itself, (ii) symmetry, that if A reduces to B then B reduces to A, and (iii) transitivity, that if A reduces to B and B reduces to C then A reduces to C.

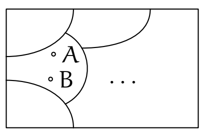
Reflexivity is easy; any matrix reduces to itself in zero-many operations.

The relationship is symmetric by the prior lemma — if A reduces to B by some row operations then also B reduces to A by reversing those operations.

For transitivity, suppose that A reduces to B and that B reduces to C. Following the reduction steps from $A \to \cdots \to B$ with those from $B \to \cdots \to C$ gives a reduction from A to C. QED

1.7 Definition Two matrices that are interreducible by elementary row operations are *row equivalent*.

The diagram below shows the collection of all matrices as a box. Inside that box each matrix lies in a class. Matrices are in the same class if and only if they are interreducible. The classes are disjoint — no matrix is in two distinct classes. We have partitioned the collection of matrices into *row equivalence classes*.†



* More information on equivalence relations is in the appendix.

† More information on partitions and class representatives is in the appendix.

One of the classes is the cluster of interrelated matrices from the start of this section pictured earlier, expanded to include all of the nonsingular 2×2 matrices.

The next subsection proves that the reduced echelon form of a matrix is unique. Rephrased in terms of the row-equivalence relationship, we shall prove that every matrix is row equivalent to one and only one reduced echelon form matrix. In terms of the partition what we shall prove is: every equivalence class contains one and only one reduced echelon form matrix. So each reduced echelon form matrix serves as a representative of its class.

Exercises

✓ **1.8** Use Gauss-Jordan reduction to solve each system.

(a) $\begin{aligned} x+y&=2\\ x-y&=0\end{aligned}$ **(b)** $\begin{aligned} x-z&=4\\ 2x+2y&=1\end{aligned}$ **(c)** $\begin{aligned} 3x-2y&=1\\ 6x+y&=1/2\end{aligned}$

(d) $\begin{aligned} 2x-y&=-1\\ x+3y-z&=5\\ y+2z&=5\end{aligned}$

✓ **1.9** Find the reduced echelon form of each matrix.

(a) $\begin{pmatrix}2&1\\1&3\end{pmatrix}$ **(b)** $\begin{pmatrix}1&3&1\\2&0&4\\-1&-3&-3\end{pmatrix}$ **(c)** $\begin{pmatrix}1&0&3&1&2\\1&4&2&1&5\\3&4&8&1&2\end{pmatrix}$

(d) $\begin{pmatrix}0&1&3&2\\0&0&5&6\\1&5&1&5\end{pmatrix}$

✓ **1.10** Find each solution set by using Gauss-Jordan reduction and then reading off the parametrization.

(a) $\begin{aligned} 2x+y-z&=1\\ 4x-y&=3\end{aligned}$ **(b)** $\begin{aligned} x-z&=1\\ y+2z-w&=3\\ x+2y+3z-w&=7\end{aligned}$ **(c)** $\begin{aligned} x-y+z&=0\\ y+w&=0\\ 3x-2y+3z+w&=0\\ -y-w&=0\end{aligned}$

(d) $\begin{aligned} a+2b+3c+d-e&=1\\ 3a-b+c+d+e&=3\end{aligned}$

1.11 Give two distinct echelon form versions of this matrix.

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

✓ **1.12** List the reduced echelon forms possible for each size.

(a) 2×2 **(b)** 2×3 **(c)** 3×2 **(d)** 3×3

✓ **1.13** What results from applying Gauss-Jordan reduction to a nonsingular matrix?

1.14 [Cleary] Consider the following relationship on the set of 2×2 matrices: we say that A is *sum-what like* B if the sum of all of the entries in A is the same as the sum of all the entries in B. For instance, the zero matrix would be sum-what like the matrix whose first row had two sevens, and whose second row had two negative sevens. Prove or disprove that this is an equivalence relation on the set of 2×2 matrices.

1.15 The proof of Lemma 1.5 contains a reference to the $i \neq j$ condition on the row combination operation.

(a) Write down a $2{\times}2$ matrix with nonzero entries, and show that the $-1 \cdot \rho_1 + \rho_1$ operation is not reversed by $1 \cdot \rho_1 + \rho_1$.

(b) Expand the proof of that lemma to make explicit exactly where it uses the $i \neq j$ condition on combining.

1.16 [Cleary] Consider the set of students in a class. Which of the following relationships are equivalence relations? Explain each answer in at least a sentence.

(a) Two students x, y are related if x has taken at least as many math classes as y.

(b) Students x, y are related if they have names that start with the same letter.

1.17 Show that each of these is an equivalence on the set of $2{\times}2$ matrices. Describe the equivalence classes.

(a) Two matrices are related if they have the same product down the diagonal, that is, if the product of the entries in the upper left and lower right are equal.

(b) Two matrices are related if they both have at least one entry that is a 1, or if neither does.

1.18 Show that each is not an equivalence on the set of $2{\times}2$ matrices.

(a) Two matrices A, B are related if $a_{1,1} = -b_{1,1}$.

(b) Two matrices are related if the sum of their entries are within 5, that is, A is related to B if $|(a_{1,1} + \cdots + a_{2,2}) - (b_{1,1} + \cdots + b_{2,2})| < 5$.

III.2 The Linear Combination Lemma

We will close this chapter by proving that every matrix is row equivalent to one and only one reduced echelon form matrix. The ideas here will reappear, and be further developed, in the next chapter.

The crucial observation concerns how row operations act to transform one matrix into another: they combine the rows linearly, that is, the new rows are linear combinations of the old rows.

2.1 Example Consider this Gauss-Jordan reduction.

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 1 & 3 & 5 \end{array}\right) \xrightarrow{-(1/2)\rho_1+\rho_2} \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 5/2 & 5 \end{array}\right)$$
$$\xrightarrow[(2/5)\rho_2]{(1/2)\rho_1} \left(\begin{array}{cc|c} 1 & 1/2 & 0 \\ 0 & 1 & 2 \end{array}\right)$$
$$\xrightarrow{-(1/2)\rho_2+\rho_1} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array}\right)$$

Denoting those matrices $A \to D \to G \to B$ and writing the rows of A as α_1 and α_2, etc., we have this.

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \xrightarrow{-(1/2)\rho_1+\rho_2} \begin{pmatrix} \delta_1 = \alpha_1 \\ \delta_2 = -(1/2)\alpha_1 + \alpha_2 \end{pmatrix}$$

$$\xrightarrow[(2/5)\rho_2]{(1/2)\rho_1} \begin{pmatrix} \gamma_1 = (1/2)\alpha_1 \\ \gamma_2 = -(1/5)\alpha_1 + (2/5)\alpha_2 \end{pmatrix}$$

$$\xrightarrow{-(1/2)\rho_2+\rho_1} \begin{pmatrix} \beta_1 = (3/5)\alpha_1 - (1/5)\alpha_2 \\ \beta_2 = -(1/5)\alpha_1 + (2/5)\alpha_2 \end{pmatrix}$$

2.2 Example The fact that Gaussian operations combine rows linearly also holds if there is a row swap. With this A, D, G, and B

$$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} \xrightarrow{\rho_1 \leftrightarrow \rho_2} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \xrightarrow{(1/2)\rho_2} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \xrightarrow{-\rho_2+\rho_1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

we get these linear relationships.

$$\begin{pmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \end{pmatrix} \xrightarrow{\rho_1 \leftrightarrow \rho_2} \begin{pmatrix} \vec{\delta}_1 = \vec{\alpha}_2 \\ \vec{\delta}_2 = \vec{\alpha}_1 \end{pmatrix} \xrightarrow{(1/2)\rho_2} \begin{pmatrix} \vec{\gamma}_1 = \vec{\alpha}_2 \\ \vec{\gamma}_2 = (1/2)\vec{\alpha}_1 \end{pmatrix}$$

$$\xrightarrow{-\rho_2+\rho_1} \begin{pmatrix} \vec{\beta}_1 = (-1/2)\vec{\alpha}_1 + 1 \cdot \vec{\alpha}_2 \\ \vec{\beta}_2 = (1/2)\vec{\alpha}_1 \end{pmatrix}$$

In summary, Gauss's Method systematically finds a suitable sequence of linear combinations of the rows.

2.3 Lemma (Linear Combination Lemma) A linear combination of linear combinations is a linear combination.

PROOF Given the set $c_{1,1}x_1 + \cdots + c_{1,n}x_n$ through $c_{m,1}x_1 + \cdots + c_{m,n}x_n$ of linear combinations of the x's, consider a combination of those

$$d_1(c_{1,1}x_1 + \cdots + c_{1,n}x_n) + \cdots + d_m(c_{m,1}x_1 + \cdots + c_{m,n}x_n)$$

where the d's are scalars along with the c's. Distributing those d's and regrouping gives

$$= (d_1c_{1,1} + \cdots + d_mc_{m,1})x_1 + \cdots + (d_1c_{1,n} + \cdots + d_mc_{m,n})x_n$$

which is also a linear combination of the x's. QED

2.4 Corollary Where one matrix reduces to another, each row of the second is a linear combination of the rows of the first.

Proof For any two interreducible matrices A and B there is some minimum number of row operations that will take one to the other. We proceed by induction on that number.

In the base step, that we can go from the first to the second using zero reduction operations, the two matrices are equal. Then each row of B is trivially a combination of A's rows $\vec{\beta}_i = 0 \cdot \vec{\alpha}_1 + \cdots + 1 \cdot \vec{\alpha}_i + \cdots + 0 \cdot \vec{\alpha}_m$.

For the inductive step assume the inductive hypothesis: with $k \geqslant 0$, any matrix that can be derived from A in k or fewer operations has rows that are linear combinations of A's rows. Consider a matrix B such that reducing A to B requires $k+1$ operations. In that reduction there is a next-to-last matrix G, so that $A \longrightarrow \cdots \longrightarrow G \longrightarrow B$. The inductive hypothesis applies to this G because it is only k steps away from A. That is, each row of G is a linear combination of the rows of A.

We will verify that the rows of B are linear combinations of the rows of G. Then the Linear Combination Lemma, Lemma 2.3, applies to show that the rows of B are linear combinations of the rows of A.

If the row operation taking G to B is a swap then the rows of B are just the rows of G reordered and each row of B is a linear combination of the rows of G. If the operation taking G to B is multiplication of a row by a scalar $c\rho_i$ then $\vec{\beta}_i = c\vec{\gamma}_i$ and the other rows are unchanged. Finally, if the row operation is adding a multiple of one row to another $r\rho_i + \rho_j$ then only row j of B differs from the matching row of G, and $\vec{\beta}_j = r\gamma_i + \gamma_j$, which is indeed a linear combinations of the rows of G.

Because we have proved both a base step and an inductive step, the proposition follows by the principle of mathematical induction. QED

We now have the insight that Gauss's Method builds linear combinations of the rows. But of course its goal is to end in echelon form, since that is a particularly basic version of a linear system, as it has isolated the variables. For instance, in this matrix

$$R = \begin{pmatrix} 2 & 3 & 7 & 8 & 0 & 0 \\ 0 & 0 & 1 & 5 & 1 & 1 \\ 0 & 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

x_1 has been removed from x_5's equation. That is, Gauss's Method has made x_5's row in some way independent of x_1's row.

The following result makes this intuition precise. What Gauss's Method eliminates is linear relationships among the rows.

2.5 Lemma In an echelon form matrix, no nonzero row is a linear combination of the other nonzero rows.

Proof Let R be an echelon form matrix and consider its non-$\vec{0}$ rows. First observe that if we have a row written as a combination of the others $\vec{\rho}_i = c_1\vec{\rho}_1 + \cdots + c_{i-1}\vec{\rho}_{i-1} + c_{i+1}\vec{\rho}_{i+1} + \cdots + c_m\vec{\rho}_m$ then we can rewrite that equation as

$$\vec{0} = c_1\vec{\rho}_1 + \cdots + c_{i-1}\vec{\rho}_{i-1} + c_i\vec{\rho}_i + c_{i+1}\vec{\rho}_{i+1} + \cdots + c_m\vec{\rho}_m \qquad (*)$$

where not all the coefficients are zero; specifically, $c_i = -1$. The converse holds also: given equation $(*)$ where some $c_i \neq 0$ we could express $\vec{\rho}_i$ as a combination of the other rows by moving $c_i\vec{\rho}_i$ to the left and dividing by $-c_i$. Therefore we will have proved the theorem if we show that in $(*)$ all of the coefficients are 0. For that we use induction on the row number i.

The base case is the first row $i = 1$ (if there is no such nonzero row, so that R is the zero matrix, then the lemma holds vacuously). Let ℓ_i be the column number of the leading entry in row i. Consider the entry of each row that is in column ℓ_1. Equation $(*)$ gives this.

$$0 = c_1 r_{1,\ell_1} + c_2 r_{2,\ell_1} + \cdots + c_m r_{m,\ell_1} \qquad (**)$$

The matrix is in echelon form so every row after the first has a zero entry in that column $r_{2,\ell_1} = \cdots = r_{m,\ell_1} = 0$. Thus equation $(**)$ shows that $c_1 = 0$, because $r_{1,\ell_1} \neq 0$ as it leads the row.

The inductive step is much the same as the base step. Again consider equation $(*)$. We will prove that if the coefficient c_i is 0 for each row index $i \in \{1, \ldots, k\}$ then c_{k+1} is also 0. We focus on the entries from column ℓ_{k+1}.

$$0 = c_1 r_{1,\ell_{k+1}} + \cdots + c_{k+1} r_{k+1,\ell_{k+1}} + \cdots + c_m r_{m,\ell_{k+1}}$$

By the inductive hypothesis $c_1, \ldots c_k$ are all 0 so this reduces to the equation $0 = c_{k+1} r_{k+1,\ell_{k+1}} + \cdots + c_m r_{m,\ell_{k+1}}$. The matrix is in echelon form so the entries $r_{k+2,\ell_{k+1}}, \ldots, r_{m,\ell_{k+1}}$ are all 0. Thus $c_{k+1} = 0$, because $r_{k+1,\ell_{k+1}} \neq 0$ as it is the leading entry. QED

2.6 Theorem Each matrix is row equivalent to a unique reduced echelon form matrix.

Proof [Yuster] Fix a number of rows m. We will proceed by induction on the number of columns n.

The base case is that the matrix has $n = 1$ column. If this is the zero matrix then its echelon form is the zero matrix. If instead it has any nonzero entries then when the matrix is brought to reduced echelon form it must have at least one nonzero entry, which must be a 1 in the first row. Either way, its reduced echelon form is unique.

For the inductive step we assume that $n > 1$ and that all m row matrices having fewer than n columns have a unique reduced echelon form. Consider an $m \times n$ matrix A and suppose that B and C are two reduced echelon form matrices derived from A. We will show that these two must be equal.

Let $\hat{A}$ be the matrix consisting of the first $n-1$ columns of A. Observe that any sequence of row operations that bring A to reduced echelon form will also bring $\hat{A}$ to reduced echelon form. By the inductive hypothesis this reduced echelon form of $\hat{A}$ is unique, so if B and C differ then the difference must occur in column n.

We finish the inductive step, and the argument, by showing that the two cannot differ only in that column. Consider a homogeneous system of equations for which A is the matrix of coefficients.

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= 0 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= 0 \end{aligned} \qquad (*)$$

By Theorem One.I.1.5 the set of solutions to that system is the same as the set of solutions to B's system

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + \cdots + b_{1,n}x_n &= 0 \\ b_{2,1}x_1 + b_{2,2}x_2 + \cdots + b_{2,n}x_n &= 0 \\ &\vdots \\ b_{m,1}x_1 + b_{m,2}x_2 + \cdots + b_{m,n}x_n &= 0 \end{aligned} \qquad (**)$$

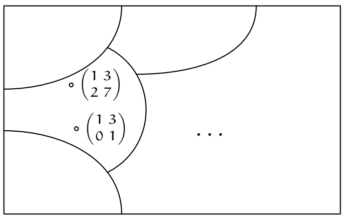
and to C's.

$$\begin{aligned} c_{1,1}x_1 + c_{1,2}x_2 + \cdots + c_{1,n}x_n &= 0 \\ c_{2,1}x_1 + c_{2,2}x_2 + \cdots + c_{2,n}x_n &= 0 \\ &\vdots \\ c_{m,1}x_1 + c_{m,2}x_2 + \cdots + c_{m,n}x_n &= 0 \end{aligned} \qquad (***)$$

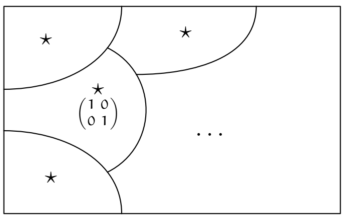
With B and C different only in column n, suppose that they differ in row i. Subtract row i of $(***)$ from row i of $(**)$ to get the equation $(b_{i,n} - c_{i,n}) \cdot x_n = 0$. We've assumed that $b_{i,n} \neq c_{i,n}$ so $x_n = 0$. Thus in $(**)$ and $(***)$ the n-th column contains a leading entry, or else the variable x_n would be free. That's a contradiction because with B and C equal on the first $n-1$ columns, the leading entries in the n-th column would have to be in the same row, and with both matrices in reduced echelon form, both leading entries would have to be 1, and would have to be the only nonzero entries in that column. So $B = C$. QED

That result answers the two questions from this section's introduction: do any two echelon form versions of a linear system have the same number of free variables, and if so are they exactly the same variables? We get from any echelon form version to the reduced echelon form by eliminating up, so any echelon form version of a system has the same free variables as the reduced echelon form, and therefore uniqueness of reduced echelon form gives that the same variables are free in all echelon form version of a system. Thus both questions are answered "yes." There is no linear system and no combination of row operations such that, say, we could solve the system one way and get y and z free but solve it another way and get y and w free.

We close with a recap. In Gauss's Method we start with a matrix and then derive a sequence of other matrices. We defined two matrices to be related if we can derive one from the other. That relation is an equivalence relation, called row equivalence, and so partitions the set of all matrices into row equivalence classes.



(There are infinitely many matrices in the pictured class, but we've only got room to show two.) We have proved there is one and only one reduced echelon form matrix in each row equivalence class. So the reduced echelon form is a canonical form* for row equivalence: the reduced echelon form matrices are representatives of the classes.



The idea here is that one way to understand a mathematical situation is by being able to classify the cases that can happen. This is a theme in this book and we have seen this several times already. We classified solution sets of linear systems into the no-elements, one-element, and infinitely-many elements

* More information on canonical representatives is in the appendix.

cases. We also classified linear systems with the same number of equations as unknowns into the nonsingular and singular cases.

Here, where we are investigating row equivalence, we know that the set of all matrices breaks into the row equivalence classes and we now have a way to put our finger on each of those classes — we can think of the matrices in a class as derived by row operations from the unique reduced echelon form matrix in that class.

Put in more operational terms, uniqueness of reduced echelon form lets us answer questions about the classes by translating them into questions about the representatives. For instance, as promised in this section's opening, we now can decide whether one matrix can be derived from another by row reduction. We apply the Gauss-Jordan procedure to both and see if they yield the same reduced echelon form.

2.7 Example These matrices are not row equivalent

$$\begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & -3 \\ -2 & 5 \end{pmatrix}$$

because their reduced echelon forms are not equal.

$$\begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2.8 Example Any nonsingular 3×3 matrix Gauss-Jordan reduces to this.

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

2.9 Example We can describe all the classes by listing all possible reduced echelon form matrices. Any 2×2 matrix lies in one of these: the class of matrices row equivalent to this,

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

the infinitely many classes of matrices row equivalent to one of this type

$$\begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix}$$

where $a \in \mathbb{R}$ (including $a = 0$), the class of matrices row equivalent to this,

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

and the class of matrices row equivalent to this

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

(this is the class of nonsingular 2×2 matrices).

Exercises

✓ 2.10 Decide if the matrices are row equivalent.

(a) $\begin{pmatrix}1&2\\4&8\end{pmatrix},\begin{pmatrix}0&1\\1&2\end{pmatrix}$ (b) $\begin{pmatrix}1&0&2\\3&-1&1\\5&-1&5\end{pmatrix},\begin{pmatrix}1&0&2\\0&2&10\\2&0&4\end{pmatrix}$

(c) $\begin{pmatrix}2&1&-1\\1&1&0\\4&3&-1\end{pmatrix},\begin{pmatrix}1&0&2\\0&2&10\end{pmatrix}$ (d) $\begin{pmatrix}1&1&1\\-1&2&2\end{pmatrix},\begin{pmatrix}0&3&-1\\2&2&5\end{pmatrix}$

(e) $\begin{pmatrix}1&1&1\\0&0&3\end{pmatrix},\begin{pmatrix}0&1&2\\1&-1&1\end{pmatrix}$

2.11 Describe the matrices in each of the classes represented in Example 2.9.

2.12 Describe all matrices in the row equivalence class of these.

(a) $\begin{pmatrix}1&0\\0&0\end{pmatrix}$ (b) $\begin{pmatrix}1&2\\2&4\end{pmatrix}$ (c) $\begin{pmatrix}1&1\\1&3\end{pmatrix}$

2.13 How many row equivalence classes are there?

2.14 Can row equivalence classes contain different-sized matrices?

2.15 How big are the row equivalence classes?

(a) Show that for any matrix of all zeros, the class is finite.

(b) Do any other classes contain only finitely many members?

✓ 2.16 Give two reduced echelon form matrices that have their leading entries in the same columns, but that are not row equivalent.

✓ 2.17 Show that any two $n\times n$ nonsingular matrices are row equivalent. Are any two singular matrices row equivalent?

✓ 2.18 Describe all of the row equivalence classes containing these.

(a) 2×2 matrices (b) 2×3 matrices (c) 3×2 matrices

(d) 3×3 matrices

2.19 (a) Show that a vector $\vec{\beta}_0$ is a linear combination of members of the set $\{\vec{\beta}_1,\ldots,\vec{\beta}_n\}$ if and only if there is a linear relationship $\vec{0}=c_0\vec{\beta}_0+\cdots+c_n\vec{\beta}_n$ where c_0 is not zero. (*Hint.* Watch out for the $\vec{\beta}_0=\vec{0}$ case.)

(b) Use that to simplify the proof of Lemma 2.5.

✓ 2.20 [Trono] Three truck drivers went into a roadside cafe. One truck driver purchased four sandwiches, a cup of coffee, and ten doughnuts for \$8.45. Another driver purchased three sandwiches, a cup of coffee, and seven doughnuts for \$6.30. What did the third truck driver pay for a sandwich, a cup of coffee, and a doughnut?

✓ 2.21 The Linear Combination Lemma says which equations can be gotten from Gaussian reduction of a given linear system.

(1) Produce an equation not implied by this system.

$$\begin{aligned} 3x + 4y &= 8 \\ 2x + y &= 3 \end{aligned}$$

(2) Can any equation be derived from an inconsistent system?

2.22 [Hoffman & Kunze] Extend the definition of row equivalence to linear systems. Under your definition, do equivalent systems have the same solution set?

✓ 2.23 In this matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 0 & 3 \\ 1 & 4 & 5 \end{pmatrix}$$

the first and second columns add to the third.

(a) Show that remains true under any row operation.

(b) Make a conjecture.

(c) Prove that it holds.

Topic

Computer Algebra Systems

The linear systems in this chapter are small enough that their solution by hand is easy. For large systems, including those involving thousands of equations, we need a computer. There are special purpose programs such as LINPACK for this. Also popular are general purpose computer algebra systems including *Maple*, *Mathematica*, or *MATLAB*, and *Sage*.

For example, in the Topic on Networks, we need to solve this.

$$\begin{array}{rcrcrcrcrcrcrcl} i_0 & - & i_1 & - & i_2 & & & & & & & & & = & 0 \\ & & i_1 & & & - & i_3 & & & - & i_5 & & & = & 0 \\ & & & & i_2 & & & - & i_4 & + & i_5 & & & = & 0 \\ & & & & & & i_3 & + & i_4 & & & - & i_6 & = & 0 \\ & & 5i_1 & & & + & 10i_3 & & & & & & & = & 10 \\ & & & & 2i_2 & & & + & 4i_4 & & & & & = & 10 \\ & & 5i_1 & - & 2i_2 & & & & & + & 50i_5 & & & = & 0 \end{array}$$

Doing this by hand would take time and be error-prone. A computer is better.

Here is that system solved with *Sage*. (There are many ways to do this; the one here has the advantage of simplicity.)

```
sage: var('i0,i1,i2,i3,i4,i5,i6')
(i0, i1, i2, i3, i4, i5, i6)
sage: network_system=[i0-i1-i2==0, i1-i3-i5==0,
....:       i2-i4+i5==0,, i3+i4-i6==0, 5*i1+10*i3==10,
....:       2*i2+4*i4==10, 5*i1-2*i2+50*i5==0]
sage: solve(network_system, i0,i1,i2,i3,i4,i5,i6)
[[i0 == (7/3), i1 == (2/3), i2 == (5/3), i3 == (2/3),
      i4 == (5/3), i5 == 0, i6 == (7/3)]]
```

Magic.

Here is the same system solved under Maple. We enter the array of coefficients and the vector of constants, and then we get the solution.

```
> A:=array( [[1,-1,-1,0,0,0,0],
            [0,1,0,-1,0,-1,0],
            [0,0,1,0,-1,1,0],
            [0,0,0,1,1,0,-1],
            [0,5,0,10,0,0,0],
```

```
              [0,0,2,0,4,0,0],
              [0,5,-2,0,0,50,0]] );
> u:=array( [0,0,0,0,10,10,0] );
> linsolve(A,u);
      7  2  5  2  5     7
    [ -, -, -, -, -, 0, - ]
      3  3  3  3  3     3
```

If a system has infinitely many solutions then the program will return a parametrization.

Exercises

1 Use the computer to solve the two problems that opened this chapter.

(a) This is the Statics problem.

$$\begin{aligned} 40h + 15c &= 100 \\ 25c &= 50 + 50h \end{aligned}$$

(b) This is the Chemistry problem.

$$\begin{aligned} 7h &= 7j \\ 8h + 1i &= 5j + 2k \\ 1i &= 3j \\ 3i &= 6j + 1k \end{aligned}$$

2 Use the computer to solve these systems from the first subsection, or conclude 'many solutions' or 'no solutions'.

(a) $\begin{aligned} 2x + 2y &= 5 \\ x - 4y &= 0 \end{aligned}$ (b) $\begin{aligned} -x + y &= 1 \\ x + y &= 2 \end{aligned}$ (c) $\begin{aligned} x - 3y + z &= 1 \\ x + y + 2z &= 14 \end{aligned}$ (d) $\begin{aligned} -x - y &= 1 \\ -3x - 3y &= 2 \end{aligned}$

(e) $\begin{aligned} 4y + z &= 20 \\ 2x - 2y + z &= 0 \\ x + z &= 5 \\ x + y - z &= 10 \end{aligned}$ (f) $\begin{aligned} 2x + z + w &= 5 \\ y - w &= -1 \\ 3x - z - w &= 0 \\ 4x + y + 2z + w &= 9 \end{aligned}$

3 Use the computer to solve these systems from the second subsection.

(a) $\begin{aligned} 3x + 6y &= 18 \\ x + 2y &= 6 \end{aligned}$ (b) $\begin{aligned} x + y &= 1 \\ x - y &= -1 \end{aligned}$ (c) $\begin{aligned} x_1 + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 5 \\ 4x_1 - x_2 + 5x_3 &= 17 \end{aligned}$

(d) $\begin{aligned} 2a + b - c &= 2 \\ 2a + c &= 3 \\ a - b &= 0 \end{aligned}$ (e) $\begin{aligned} x + 2y - z &= 3 \\ 2x + y + w &= 4 \\ x - y + z + w &= 1 \end{aligned}$ (f) $\begin{aligned} x + z + w &= 4 \\ 2x + y - w &= 2 \\ 3x + y + z &= 7 \end{aligned}$

4 What does the computer give for the solution of the general 2×2 system?

$$\begin{aligned} ax + cy &= p \\ bx + dy &= q \end{aligned}$$

Topic

Accuracy of Computations

Gauss's Method lends itself to computerization. The code below illustrates. It operates on an $n \times n$ matrix named a, doing row combinations using the first row, then the second row, etc.

```
for(row=1; row<=n-1; row++){
  for(row_below=row+1; row_below<=n; row_below++){
    multiplier=a[row_below,row]/a[row,row];
    for(col=row; col<=n; col++){
      a[row_below,col]-=multiplier*a[row,col];
    }
  }
}
```

This is in the C language. The `for(row=1; row<=n-1; row++){ .. }` loop initializes `row` at 1 and then iterates while `row` is less than or equal to $n-1$, each time through incrementing `row` by one with the `++` operation. The other non-obvious language construct is that the `-=` in the innermost loop has the effect of `a[row_below,col]=-1*multiplier*a[row,col]+a[row_below,col]`.

While that code is a first take on mechanizing Gauss's Method, it is naive. For one thing, it assumes that the entry in the `row,row` position is nonzero. So one way that it needs to be extended is to cover the case where finding a zero in that location leads to a row swap or to the conclusion that the matrix is singular.

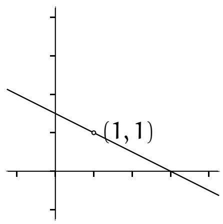
We could add some `if` statements to cover those cases but we will instead consider another way in which this code is naive. It is prone to pitfalls arising from the computer's reliance on floating point arithmetic.

For example, above we have seen that we must handle a singular system as a separate case. But systems that are nearly singular also require care. Consider this one (the extra digits are in the ninth significant place).

$$\begin{aligned} x + 2y &= 3 \\ 1.000\,000\,01x + 2y &= 3.000\,000\,01 \end{aligned} \qquad (*)$$

By eye we easily spot the solution $x = 1$, $y = 1$. A computer has more trouble. If it represents real numbers to eight significant places, called *single precision*, then it will represent the second equation internally as $1.000\,000\,0x+2y=3.000\,000\,0$, losing the digits in the ninth place. Instead of reporting the correct solution, this computer will think that the two equations are equal and it will report that the system is singular.

For some intuition about how the computer could come up with something that far off, consider this graph of the system.



At the scale of this drawing we cannot tell the two lines apart. This system is nearly singular in the sense that the two lines are nearly the same line.

Near-singularity gives the system $(*)$ the property that a small change in the system can cause a large change in its solution. For instance, changing the $3.000\,000\,01$ to $3.000\,000\,03$ changes the intersection point from $(1,1)$ to $(3,0)$. The solution changes radically depending on the ninth digit, which explains why an eight-place computer has trouble. A problem that is very sensitive to inaccuracy or uncertainties in the input values is *ill-conditioned*.

The above example gives one way in which a system can be difficult to solve on a computer. It has the advantage that the picture of nearly-equal lines gives a memorable insight into one way for numerical difficulties to happen. Unfortunately this insight isn't useful when we wish to solve some large system. We typically will not understand the geometry of an arbitrary large system.

There are other ways that a computer's results may be unreliable, besides that the angle between some of the linear surfaces is small. For example, consider this system (from [Hamming]).

$$\begin{aligned} 0.001x+y&=1\\ x-y&=0 \end{aligned} \qquad (**)$$

The second equation gives $x=y$, so $x=y=1/1.001$ and thus both variables have values that are just less than 1. A computer using two digits represents the system internally in this way (we will do this example in two-digit floating point arithmetic for clarity but inventing a similar one with eight or more digits is easy).

$$\begin{aligned} (1.0\times10^{-3})\cdot x+(1.0\times10^{0})\cdot y&=1.0\times10^{0}\\ (1.0\times10^{0})\cdot x-(1.0\times10^{0})\cdot y&=0.0\times10^{0} \end{aligned}$$

The row reduction step $-1000\rho_1 + \rho_2$ produces a second equation $-1001y = -1000$, which this computer rounds to two places as $(-1.0 \times 10^3)y = -1.0 \times 10^3$. The computer decides from the second equation that $y = 1$ and with that it concludes from the first equation that $x = 0$. The y value is close but the x is bad — the ratio of the actual answer to the computer's answer is infinite. In short, another cause of unreliable output is the computer's reliance on floating point arithmetic when the system-solving code leads to using leading entries that are small.

An experienced programmer may respond by using *double precision*, which retains sixteen significant digits, or perhaps using some even larger size. This will indeed solve many problems. However, double precision has greater memory requirements and besides we can obviously tweak the above to give the same trouble in the seventeenth digit, so double precision isn't a panacea. We need a strategy to minimize numerical trouble as well as some guidance about how far we can trust the reported solutions.

A basic improvement on the naive code above is to not determine the factor to use for row combinations by simply taking the entry in the `row,row` position, but rather to look at all of the entries in the `row` column below the `row,row` entry and take one that is likely to give reliable results because it is not too small. This is *partial pivoting*.

For example, to solve the troublesome system $(**)$ above we start by looking at both equations for a best entry to use, and take the 1 in the second equation as more likely to give good results. The combination step of $-.001\rho_2 + \rho_1$ gives a first equation of $1.001y = 1$, which the computer will represent as $(1.0 \times 10^0)y = 1.0 \times 10^0$, leading to the conclusion that $y = 1$ and, after back-substitution, that $x = 1$, both of which are close to right. We can adapt the code from above to do this.

```
for(row=1; row<=n-1; row++){
/* find the largest entry in this column (in row max) */
  max=row;
  for(row_below=row+1; row_below<=n; row_below++){
    if (abs(a[row_below,row]) > abs(a[max,row]));
      max = row_below;
  }
/* swap rows to move that best entry up */
  for(col=row; col<=n; col++){
    temp=a[row,col];
    a[row,col]=a[max,col];
    a[max,col]=temp;
  }
/* proceed as before */
  for(row_below=row+1; row_below<=n; row_below++){
    multiplier=a[row_below,row]/a[row,row];
     for(col=row; col<=n; col++){
       a[row_below,col]-=multiplier*a[row,col];
    }
  }
}
```

A full analysis of the best way to implement Gauss's Method is beyond the scope of this book (see [Wilkinson 1965]), but the method recommended by most experts first finds the best entry among the candidates and then scales it to a number that is less likely to give trouble. This is *scaled partial pivoting*.

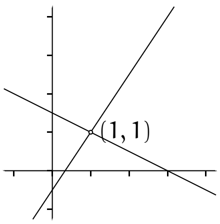
In addition to returning a result that is likely to be reliable, most well-done code will return a *conditioning number* that describes the factor by which uncertainties in the input numbers could be magnified to become inaccuracies in the results returned (see [Rice]).

The lesson is that just because Gauss's Method always works in theory, and just because computer code correctly implements that method, doesn't mean that the answer is reliable. In practice, always use a package where experts have worked hard to counter what can go wrong.

Exercises

1 Using two decimal places, add 253 and 2/3.

2 This intersect-the-lines problem contrasts with the example discussed above.



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

Illustrate that in this system some small change in the numbers will produce only a small change in the solution by changing the constant in the bottom equation to 1.008 and solving. Compare it to the solution of the unchanged system.

3 Solve this system by hand ([Rice]).

$$\begin{aligned} 0.0003x + 1.556y &= 1.569 \\ 0.3454x - 2.346y &= 1.018 \end{aligned}$$

(a) Solve it accurately, by hand. **(b)** Solve it by rounding at each step to four significant digits.

4 Rounding inside the computer often has an effect on the result. Assume that your machine has eight significant digits.

(a) Show that the machine will compute $(2/3) + ((2/3) - (1/3))$ as unequal to $((2/3) + (2/3)) - (1/3)$. Thus, computer arithmetic is not associative.

(b) Compare the computer's version of $(1/3)x + y = 0$ and $(2/3)x + 2y = 0$. Is twice the first equation the same as the second?

5 Ill-conditioning is not only dependent on the matrix of coefficients. This example [Hamming] shows that it can arise from an interaction between the left and right sides of the system. Let ε be a small real.

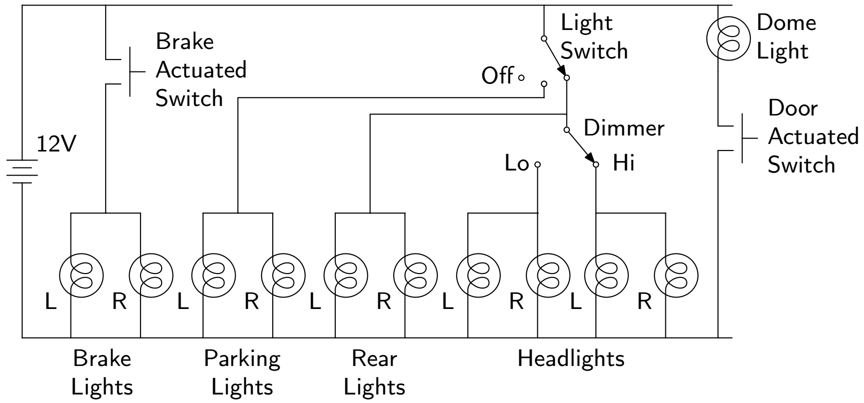
$$\begin{aligned} 3x + 2y + z &= 6 \\ 2x + 2\varepsilon y + 2\varepsilon z &= 2 + 4\varepsilon \\ x + 2\varepsilon y - \varepsilon z &= 1 + \varepsilon \end{aligned}$$

(a) Solve the system by hand. Notice that the ε's divide out only because there is an exact cancellation of the integer parts on the right side as well as on the left.

(b) Solve the system by hand, rounding to two decimal places, and with $\varepsilon = 0.001$.

Analyzing Networks

The diagram below shows some of a car's electrical network. The battery is on the left, drawn as stacked line segments. The wires are lines, shown straight and with sharp right angles for neatness. Each light is a circle enclosing a loop.



The designer of such a network needs to answer questions such as: how much electricity flows when both the hi-beam headlights and the brake lights are on? We will use linear systems to analyze simple electrical networks.

For the analysis we need two facts about electricity and two facts about electrical networks.

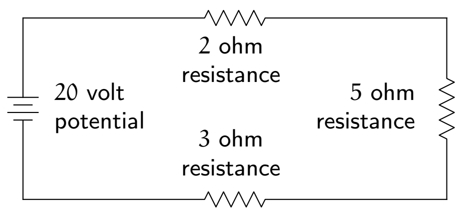
The first fact is that a battery is like a pump, providing a force impelling the electricity to flow, if there is a path. We say that the battery provides a *potential*. For instance, when the driver steps on the brake then the switch makes contact and so makes a circuit on the left side of the diagram, which includes the brake lights. Once the circuit exists, the battery's force creates a current flowing through that circuit, lighting the lights.

The second electrical fact is that in some kinds of network components the amount of flow is proportional to the force provided by the battery. That is, for each such component there is a number, it's *resistance*, such that the potential is equal to the flow times the resistance. Potential is measured in *volts*, the rate of flow is in *amperes*, and resistance to the flow is in *ohms*; these units are defined so that volts = amperes · ohms.

Components with this property, that the voltage-amperage response curve is a line through the origin, are *resistors*. For example, if a resistor measures 2 ohms then wiring it to a 12 volt battery results in a flow of 6 amperes. Conversely, if electrical current of 2 amperes flows through that resistor then there must be a 4 volt potential difference between it's ends. This is the *voltage drop* across the resistor. One way to think of the electrical circuits that we consider here is that the battery provides a voltage rise while the other components are voltage drops.

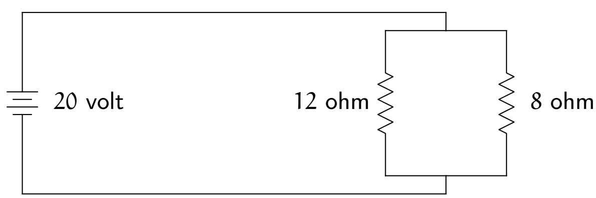
The facts that we need about networks are *Kirchoff's Current Law*, that for any point in a network the flow in equals the flow out and *Kirchoff's Voltage Law*, that around any circuit the total drop equals the total rise.

We start with the network below. It has a battery that provides the potential to flow and three resistors, shown as zig-zags. When components are wired one after another, as here, they are in *series*.

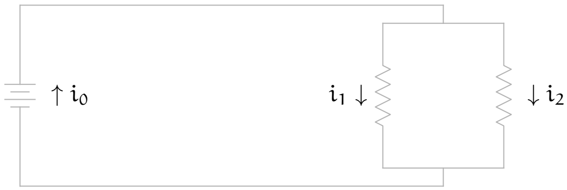


By Kirchoff's Voltage Law, because the voltage rise is 20 volts, the total voltage drop must also be 20 volts. Since the resistance from start to finish is 10 ohms (the resistance of the wire connecting the components is negligible), the current is $(20/10) = 2$ amperes. Now, by Kirchhoff's Current Law, there are 2 amperes through each resistor. Therefore the voltage drops are: 4 volts across the 2 ohm resistor, 10 volts across the 5 ohm resistor, and 6 volts across the 3 ohm resistor.

The prior network is simple enough that we didn't use a linear system but the next one is more complicated. Here the resistors are in *parallel*.



We begin by labeling the branches as below. Let the current through the left branch of the parallel portion be i_1 and that through the right branch be i_2, and also let the current through the battery be i_0. Note that we don't need to know the actual direction of flow — if current flows in the direction opposite to our arrow then we will get a negative number in the solution.

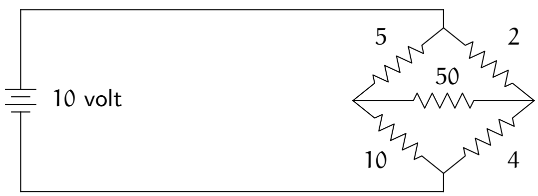


The Current Law, applied to the split point in the upper right, gives that $i_0 = i_1 + i_2$. Applied to the split point lower right it gives $i_1 + i_2 = i_0$. In the circuit that loops out of the top of the battery, down the left branch of the parallel portion, and back into the bottom of the battery, the voltage rise is 20 while the voltage drop is $i_1 \cdot 12$, so the Voltage Law gives that $12i_1 = 20$. Similarly, the circuit from the battery to the right branch and back to the battery gives that $8i_2 = 20$. And, in the circuit that simply loops around in the left and right branches of the parallel portion (we arbitrarily take the direction of clockwise), there is a voltage rise of 0 and a voltage drop of $8i_2 - 12i_1$ so $8i_2 - 12i_1 = 0$.

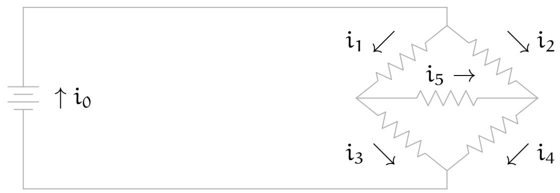
$$\begin{aligned}
i_0 - \quad i_1 - \ i_2 &= 0 \\
-i_0 + \quad i_1 + \ i_2 &= 0 \\
12i_1 \qquad\quad &= 20 \\
8i_2 &= 20 \\
-12i_1 + 8i_2 &= 0
\end{aligned}$$

The solution is $i_0 = 25/6$, $i_1 = 5/3$, and $i_2 = 5/2$, all in amperes. (Incidentally, this illustrates that redundant equations can arise in practice.)

Kirchhoff's laws can establish the electrical properties of very complex networks. The next diagram shows five resistors, whose values are in ohms, wired in *series-parallel*.



This is a *Wheatstone bridge* (see Exercise 3). To analyze it, we can place the arrows in this way.



Kirchhoff's Current Law, applied to the top node, the left node, the right node, and the bottom node gives these.

$$\begin{aligned} i_0 &= i_1 + i_2 \\ i_1 &= i_3 + i_5 \\ i_2 + i_5 &= i_4 \\ i_3 + i_4 &= i_0 \end{aligned}$$

Kirchhoff's Voltage Law, applied to the inside loop (the i_0 to i_1 to i_3 to i_0 loop), the outside loop, and the upper loop not involving the battery, gives these.

$$\begin{aligned} 5i_1 + 10i_3 &= 10 \\ 2i_2 + 4i_4 &= 10 \\ 5i_1 + 50i_5 - 2i_2 &= 0 \end{aligned}$$

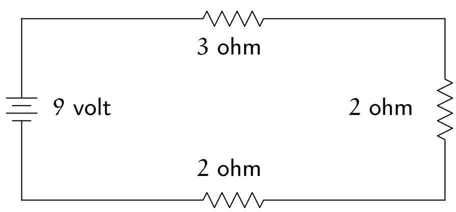
Those suffice to determine the solution $i_0 = 7/3$, $i_1 = 2/3$, $i_2 = 5/3$, $i_3 = 2/3$, $i_4 = 5/3$, and $i_5 = 0$.

We can understand many kinds of networks in this way. For instance, the exercises analyze some networks of streets.

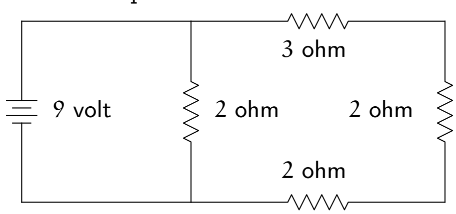
Exercises

1 Calculate the amperages in each part of each network.

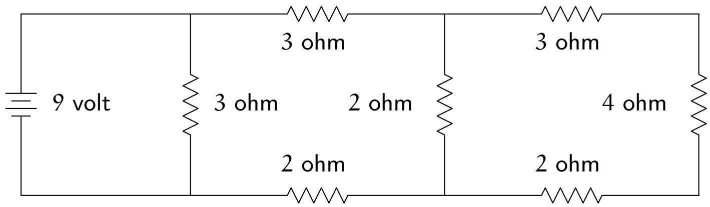
(a) This is a simple network.



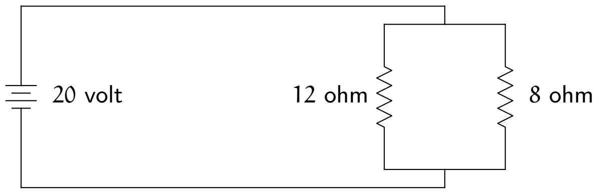
(b) Compare this one with the parallel case discussed above.



(c) This is a reasonably complicated network.



2 In the first network that we analyzed, with the three resistors in series, we just added to get that they acted together like a single resistor of 10 ohms. We can do a similar thing for parallel circuits. In the second circuit analyzed,



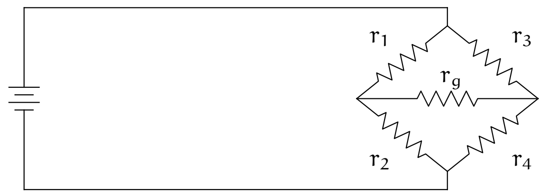
the electric current through the battery is $25/6$ amperes. Thus, the parallel portion is *equivalent* to a single resistor of $20/(25/6) = 4.8$ ohms.

(a) What is the equivalent resistance if we change the 12 ohm resistor to 5 ohms?

(b) What is the equivalent resistance if the two are each 8 ohms?

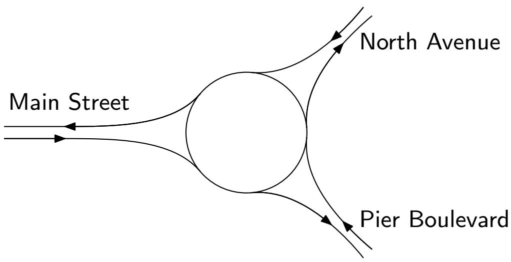
(c) Find the formula for the equivalent resistance if the two resistors in parallel are r_1 ohms and r_2 ohms.

3 A *Wheatstone bridge* is used to measure resistance.



Show that in this circuit if the current flowing through r_g is zero then $r_4 = r_2 r_3 / r_1$. (To operate the device, put the unknown resistance at r_4. At r_g is a meter that shows the current. We vary the three resistances r_1, r_2, and r_3 — typically they each have a calibrated knob — until the current in the middle reads 0. Then the equation gives the value of r_4.)

4 Consider this traffic circle.



This is the traffic volume, in units of cars per five minutes.

	North	*Pier*	*Main*
into	100	150	25
out of	75	150	50

We can set up equations to model how the traffic flows.

(a) Adapt Kirchhoff's Current Law to this circumstance. Is it a reasonable modeling assumption?

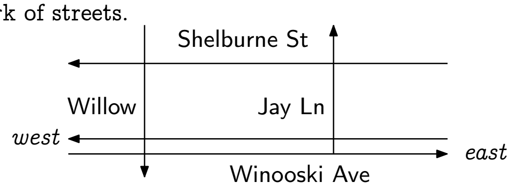
(b) Label the three between-road arcs in the circle with a variable. Using the (adapted) Current Law, for each of the three in-out intersections state an equation describing the traffic flow at that node.

(c) Solve that system.

(d) Interpret your solution.

(e) Restate the Voltage Law for this circumstance. How reasonable is it?

5 This is a network of streets.



We can observe the hourly flow of cars into this network's entrances, and out of its exits.

	east Winooski	*west Winooski*	*Willow*	*Jay*	*Shelburne*
into	80	50	65	–	40
out of	30	5	70	55	75

(Note that to reach Jay a car must enter the network via some other road first, which is why there is no 'into Jay' entry in the table. Note also that over a long period of time, the total in must approximately equal the total out, which is why both rows add to 235 cars.) Once inside the network, the traffic may flow in different ways, perhaps filling Willow and leaving Jay mostly empty, or perhaps flowing in some other way. Kirchhoff's Laws give the limits on that freedom.

(a) Determine the restrictions on the flow inside this network of streets by setting up a variable for each block, establishing the equations, and solving them. Notice that some streets are one-way only. (*Hint:* this will not yield a unique solution, since traffic can flow through this network in various ways; you should get at least one free variable.)

(b) Suppose that someone proposes construction for Winooski Avenue East between Willow and Jay, and traffic on that block will be reduced. What is the least amount of traffic flow that can we can allow on that block without disrupting the hourly flow into and out of the network?