

Section 1 A chatty tour thru Gaussian elimination.

The simplest linear equation is $Ax = b$, for example, $5x = 3$. It has one and only one solution, $x = 3/5$. Another one is $0x = 3$. This one has NO solutions. Another is $0x = 0$. Here, x may be *any* number, so this equation has infinitely many solutions. The equation is linear because the unknown x , appears to first degree, not as a divisor, nor multiplied times any other unknown.

The equation $2x + 3y = 4$ describes a line in the plane - hence the term *linear* to denote first degree. Since a line has infinitely many points on it, the equation has to have infinitely many solutions. We can solve for y and get a formula for y as a function of x : $y = 2 - (2/3)x$. But x can have any value. This is as far as we can go without further restrictions on the solution. For example, we might want to find the point on the line that is closest to $(0,0)$. This problem is not "linear" (solution: $(12,18)/13$), but does arise in the study of linear equations!

We are often interested in solving several equations simultaneously. Here are 2: find x and y so that $2x + 3y = 4$, and $3x + 5y = 7$. Geometrically this represents the set of points where 2 lines meet. We expect to find that they meet at 1 point, not at all, or that they coincide, and so meet at infinitely many points. This corresponds to the "simplest" examples, mentioned before!

The *idea* for solving a system of linear equations (several equations, true simultaneously) is simple: eliminate one unknown at a time. Here's the same one: Solve: $2x + 3y = 4$, and $3x + 5y = 7$. From the first equation, $x = (4 - 3y)/2$. Put this in place of x in the next equation: $3(4 - 3y)/2 + 5y = 7$. Simplify and solve for y : $12 - 9y + 10y = 14$, so $y = 2$. Thus, from the formula for x in terms of y , $x = (4 - 3y)/2 = -1$. Check: $2(-1) + 3 \cdot 2 = 4$ (\checkmark); $3(-1) + 5 \cdot 2 = 7$ (\checkmark). There is only this one solution. Why? We'll see later that one answer is: "The determinant is not zero." For now, we can mention the geometric interpretation: The 2 lines meet at $(-1, 2)$, and we know that different lines meet at one point or none.

When we deal with a system of linear equations, solve means to *describe completely the solution-set of the equations, assuming them to hold simultaneously*. The solution-set might be empty, or might be a single point, a line, a plane, or a higher-dimensional "linear object". The goal of our study is to acquaint you with useful ways of dealing with systems of linear equations and with other questions that have arisen in the study of linear systems that many people, including some mathematicians, have had to answer. The first main topic will be the method of Gaussian elimination, important because of its efficiency. It also introduces matrices as a way to minimize the arithmetic work and to help in error avoidance and correction. It will also give, as a by-product, handy ways to completely describe the solution-set of a linear system.

Here's another system of 3 linear equations in 3 unknowns:

$$\begin{aligned} 2x + 3y + 4z &= 5, \\ 3x + 5y + 7z &= 9, \\ 4x + 7y + 10z &= 13. \end{aligned}$$

Solve the first equation for x , in terms of y and z , and replace x in all the other equations by the solution formula. This reduces the number of unknowns by one in the second two equations. Repeat. If there is just one solution, you will eventually get to an equation with just 1 unknown, that is solvable. If there are none, you'll get an unsolvable equation. If there are LOTS of solutions, you'll get $0z = 0$. Please stop now and try to solve them before reading further!

The point of asking you to stop and solve that system using simple methods was to get you to experience some of the tedium that copying and simplifying lead to. Now imagine working on a system with hundreds of equations in hundreds of unknowns! Nowadays your reaction might be to put the problem on a computer. Would you be able to understand the output? It would be presented in a form that assumes you know a certain amount of terminology - one of the further

goals of this introduction is to provide you with this terminology. Let us begin. Here is a slightly different system:

$$\begin{aligned}2x + 3y + 4z &= 5, \\3x + 5y + 7z &= 9, \\4x + 7y + 11z &= 18.\end{aligned}$$

An unknown serves only to connect a number, its coefficient, with the unknown it goes with. In the first equation, the coefficient of x is 2, that of y is 3, and so on. I choose the x, y, z ordering of the unknowns. You may choose any one you like, as long as you make clear somehow what the ordering is. Let's write down an array, or table, or matrix that has definite rows and col-umns associated with the equations and their unknowns. Each row "contains" 1 equation's co-efficients, in the same order as the unknowns. Each column contains all the coefficients that be-long with a particular one of the unknowns, except the last column. The last column contains the desired total for each combination of unknowns and coefficients. Here is the matrix:

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

I don't know why the parentheses are there. Perhaps they make us think of the array as a single mathematical object. It is. At any rate, the parentheses are customary, but not, at this point, need-ed. One caution: if one of the equations does not contain an unknown that some other equation does contain, then the row for that equation must contain a 0 in the column for that missing un-known. Here is an example:

$$\begin{aligned}2x + 4z &= 5, \\3x + 5y + 7z &= 9, \\4x + 7y + 11z &= 18.\end{aligned}$$

The first equation does not contain y , but the second and third ones do contain y . So we have to enter a 0 in row 1 and column 2. The matrix for the system is

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

Now we need to realize that we can go the other way. Given a matrix, we can construct a corresponding system of linear equations. Thus,

$$\begin{pmatrix} 2 & 0 & 4 & 5 & 3 \\ 5 & 7 & 9 & 4 & 7 \\ 11 & 18 & 21 & 28 & 31 \end{pmatrix}.$$

is the matrix for the system

$$\begin{aligned}2x + 4z + 5w &= 3, \\5x + 7y + 9z + 4w &= 7, \\11x + 18y + 21z + 28w &= 31.\end{aligned}$$

It is also the matrix for the system

$$\begin{aligned}2a + 4c + 5d &= 3, \\9c + 7b + 5a + 4d &= 7, \\18b + 11a + 28d + 21c &= 31.\end{aligned}$$

The point is, it does not matter what the names of the unknowns are. It matters very much, though, where their coefficients go in the matrix! The wise plan would be to rewrite the equations so that the unknowns all appear in the same order in each equation.

Exercises: Given a system, write a matrix for it. Given a matrix, write a system for it.

1.
$$\begin{aligned} x + y + z &= 1, \\ x - y + 2z &= 0, \\ -x + 3y - 5z &= 3. \end{aligned}$$

2.
$$\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 \\ 2 & 8 & 3 & 4 \end{pmatrix}.$$

3.
$$\begin{pmatrix} -2 & 0 & 4 & 5 & 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 & 2 & 8 & 3 & 4 \end{pmatrix}$$

4.
$$\begin{aligned} x + y + 2z &= w, \\ x - y + 3z &= 0, \\ x + 4y - 6z &= 3w; \end{aligned} \quad \text{unknowns: } x, y, z, w.$$

5.
$$\begin{aligned} x + y + z &= 1, \\ x - y + 2z &= 0, \\ -x + 3y - 5z &= 3, \\ x - 3y + 5z &= 3. \end{aligned}$$

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

We now have several examples of systems of equations. We have seen how to go back and forth between a matrix and a system of equations. Let's start on Gaussian elimination. Solve

$$\begin{aligned} 2x + 3y + 4z &= 5, \\ 3x + 5y + 7z &= 9, \\ 4x + 7y + 11z &= 18. \end{aligned}$$

STEP 1: Set up its matrix,

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

The first entry in the first row is 2 ($\neq 0$). Multiply each entry in the first row by $3/2$, and subtract the products, one by one, from the corresponding entries in the second row. This yields the matrix

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 1/2 & 1 & 3/2 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

What's the point? To see the point, let's write down the system of equations that corresponds to this matrix, using the original unknowns, x, y, z , *in that order*:

$$\begin{aligned} 2x + 3y + 4z &= 5, \\ 0x + 1/2y + 1z &= 3/2, \\ 4x + 7y + 11z &= 18. \end{aligned}$$

It is exactly what we get when we subtract $3/2$ times the first equation from the (original) second one. The new equation has no x . It is also exactly what we get if we solve the first equation for x in terms of y and z , and then plug the formula for x into the second equation. That is, x is eliminated from the second equation. Please stop now and check it out!

Don't particularly like working with fractions? I don't. We can do a variation on Gauss elimination by first multiplying the second row (or the second equation) by 2, and then multiply-and-subtract using 3 as the multiplier in place of $3/2$:

The original
$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

Row 2 times 2
$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 6 & 10 & 14 & 18 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

Subtract 3 times row 1 from row 2
$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 4 & 7 & 11 & 18 \end{pmatrix}.$$

Why does this work? Well, multiplying a row is the "same" as multiplying the corresponding equation in a corresponding system. The solution-set of an equation is not changed if we multiply the equation by a *non-zero* number. The solution-set of "all the equations holding simultaneously" is the intersection of the three solution-sets of "each equation, holding by itself." Then, we subtracted 3 times the first row from the second row, the same thing as subtracting 3 times the first equation from the second equation. I don't know how the solution-set of the modified equation compares with the solution-set of the unmodified equation. But just suppose that a triple (x^*, y^*, z^*) is a solution of all 3 equations. Then all 3 equations are true, so when I subtract 3 times the first equation from the second equation, the new second equation is still true. Therefore, (x^*, y^*, z^*) is a solution of all 3 *new* equations. What does this show? It shows that the solution-set of the original system is contained in the solution-set of the new system. Now suppose that (u, v, w) is a solution of the new system, the one we got by subtracting 3 times the first equation from the second equation. Then, if we *add* 3 times the first equation to the second equation, we get what? The original second equation! What does this show? It shows that the solution-set of the new system is contained in the solution-set of the original system. This is a standard trick used in math to show 2 sets are equal: show that each of the 2 sets is contained in the other one. This is the basic argument for showing that adding (*subtracting*) a multiple of one equation to (*from*) another equation does not change the solution set of the system.

What does it all mean? In terms of equations, we have eliminated x from the second equation. In terms of the matrix, we have subtracted the right multiple of row 1 from row 2 that causes the new row 2 to have a zero as its first entry. Working with the matrix is quicker than working with equations, because no variable (i.e., "unknown") symbols need to be written, and no equal signs. We have to be careful that the matrix is set up right!

We keep going. To get a zero at the start of the third row, we'll

subtract 2 times row 1 from row 3
$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 8 \end{pmatrix}.$$

Now we are done with column 1, and, for the moment, with row 1. So now we go to row 2, and start with column 2. We want to keep going, so we subtract row 2 from row 3, because doing so will put a zero in the second place in row 3. The new matrix is

$$\text{(subtracted row 2 from row 3)} \quad \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 5 \end{pmatrix}.$$

Let's convert this back to a system of equations to see that we're almost done:

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

So, now we know z must be 5 , and we plug this back into equation 2, and solve *it*:

$y + 2 \cdot 5 = 3$, so y must be -7 . Plug *both* values into equation 1 and solve *that*:

$2x + 3 \cdot (-7) + 4 \cdot 5 = 5$, so x must be 3 . This is called backsolving. The final step is to check the solution, namely $(x, y, z) = (5, -7, 3)$, by substituting it into the ORIGINAL system of equations. Please stop now and do that. I think it's correct. The solution is: $x = 3, y = -7, z = 5$. There is only one solution, because we wrote down the equations, *assuming* them to be true, and performed operations on them that *did not change the solution-set*, and found single, specific values for the unknowns. In this case, then, the solution-set is: $\{ (5, -7, 3) \}$.

Here is an example, done without comment. Try R. L. Moore's Method here: use a piece of paper to cover all but the system following, and predict the next step by writing down what you think it will be. Uncover the text, a line at a time - you'll see that you probably got it right.

Example: Solve

$$\begin{aligned} 2x + 4z &= 5, \\ 3x + 5y + 7z &= 9, \\ 4x + 7y + 11z &= 18. \end{aligned}$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 11 & 18 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 6 & 10 & 14 & 18 \\ 4 & 7 & 11 & 18 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 0 & 10 & 2 & 3 \\ 4 & 7 & 11 & 18 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 0 & 10 & 2 & 3 \\ 0 & 7 & 3 & 8 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 0 & 10 & 2 & 3 \\ 0 & 70 & 30 & 80 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 4 & 5 \\ 0 & 10 & 2 & 3 \\ 0 & 0 & 16 & 59 \end{pmatrix},$$

$$\begin{aligned}2x + 4z &= 5, \\10y + 2z &= 3, \\16z &= 59.\end{aligned}$$

$$z = 59/16;$$

$$y = (3 - 2z)/10 = (16 \cdot 3 - 2 \cdot 59)/160 = (48 - 118)/160 = -70/160 = -7/16;$$

$$x = (5 - 4z)/2 = (16 \cdot 5 - 4 \cdot 59)/32 = (80 - 236)/32 = -156/32 = -78/16.$$

Check: $2x + 4z = 5$?

$$2(-78/16) + 4(59/16) = (-156 + 236)/16 = 80/16 = 5, (\checkmark);$$

$$3x + 5y + 7z = 9?$$

$$3(-78/16) + 5(-7/16) + 7(59/16) = (-234 - 35 + 413)/16 = (-269 + 413)/16 = 144/16 = 9 (\checkmark);$$

$$4x + 7y + 11z = 18?$$

$$4(-78/16) + 7(-7/16) + 11(59/16) = (-312 - 49 + 649)/16 = (-361 + 649)/16 = 288/16 = 18, (\checkmark).$$

Solution: $(x, y, z) = (-78/16, -7/16, 59/16) = (-78, -7, 59)/16$. It is unique.

Next comes an example starting with a matrix - when the time comes to backsolve, we'll put in unknowns.

Example: Apply Gaussian elimination to the following matrix:

$$\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 \\ 2 & 8 & 3 & 4 \end{pmatrix}.$$

In this example, I'll keep track of what I do by writing down what I did to various rows, and where I put the result. A row will be denoted by its number, in brackets. For example, the first thing I'll do is add row 1 to row 4 to get a row starting with 0. The description I'll give, in shorthand, is $[1] + [4] \rightarrow [4]$, meaning "add row 1 to row 4, and put the result in row 4, replacing the original row 4." This example illustrates a "wandering around" approach that does the "easy" stuff first, and does some "extra" stuff to avoid fractions.

Steps to get from this $\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 \\ 2 & 8 & 3 & 4 \end{pmatrix}$, matrix to the next are listed "below," left

$[1] + [4] \rightarrow [4]$ $\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 \\ 0 & 8 & 7 & 9 \end{pmatrix}$, I saw it was easy - no multiplying

$2[1] + [3] \rightarrow [3]$ $\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 0 & 7 & 7 & 18 \\ 0 & 8 & 7 & 9 \end{pmatrix}$,

$[1] + [2] \rightarrow [2]$ $\begin{pmatrix} -2 & 0 & 4 & 5 \\ 1 & 5 & 13 & 12 \\ 0 & 7 & 7 & 18 \\ 0 & 8 & 7 & 9 \end{pmatrix}$, to get a 1 in row 2, to avoid fractions!

$$[1] + 2[2] \rightarrow [1] \quad \begin{pmatrix} 0 & 10 & 30 & 29 \\ 1 & 5 & 13 & 12 \\ 0 & 7 & 7 & 18 \\ 0 & 8 & 7 & 9 \end{pmatrix},$$

$$[2] \leftrightarrow [1] \quad \begin{pmatrix} 1 & 5 & 13 & 12 \\ 0 & 10 & 30 & 29 \\ 0 & 7 & 7 & 18 \\ 0 & 8 & 7 & 9 \end{pmatrix}, \quad \text{i.e. exchange rows 1 and 2,}$$

$$[3] - [4] \rightarrow [3] \quad \begin{pmatrix} 1 & 5 & 13 & 12 \\ 0 & 10 & 30 & 29 \\ 0 & -1 & 0 & 9 \\ 0 & 8 & 7 & 9 \end{pmatrix}, \quad \text{to get a } -1 \text{ in row 3, to avoid fractions!}$$

$$\begin{array}{l} 8[3]+[4] \rightarrow [4] \text{ \&} \\ 10[3]+[2] \rightarrow [2] \end{array} \quad \begin{pmatrix} 1 & 5 & 13 & 12 \\ 0 & 0 & 30 & 119 \\ 0 & -1 & 0 & 9 \\ 0 & 0 & 7 & 81 \end{pmatrix}, \quad \text{2 steps at once O.K. - but don't change [3]!}$$

$$[3] \leftrightarrow [2] \quad \begin{pmatrix} 1 & 5 & 13 & 12 \\ 0 & -1 & 0 & 9 \\ 0 & 0 & 30 & 119 \\ 0 & 0 & 7 & 81 \end{pmatrix},$$

$$7[3] - 30[4] \rightarrow [4] \quad \begin{pmatrix} 1 & 5 & 13 & 12 \\ 0 & -1 & 0 & 9 \\ 0 & 0 & 30 & 109 \\ 0 & 0 & 0 & -1597 \end{pmatrix}.$$

The last step is O.K.; imagine that we *first* did the step $-30[4] \rightarrow [4]$, *then* did $7[3] + [4] \rightarrow [4]$. It would be doing the same thing. Please check it out!

Notice what the last row says! If we choose x, y, z as our unknowns, it says $0z = -1597$, and this equation has no solution! At the start we actually assumed all the equations were true for the same choices of x, y, z throughout. We found that was NOT TRUE. That is, there is NO solution.

Solution: There is no triple x, y, z that makes all the equations true. We could say, the solution is that there is no solution! But it does solve the problem because we *have* completely described the solution-set of the system of equations: it is the empty set.

Comments: This system had more equations than unknowns. Usually, such systems don't have any solutions - too many conditions for too few unknowns. But, sometimes they do have a solution, even infinitely many! And sometimes there is no solution when there are the same number of equations and unknowns - for example, 2 parallel lines in the plane that don't meet can each be described by its own equation, and, they can't have the same pair (x, y) as a solution. Please make up an example, now!

Example: Let's apply Gaussian elimination to the system I asked you to solve the "direct" way:

$$\begin{array}{l} 2x + 3y + 4z = 5, \\ 3x + 5y + 7z = 9, \\ 4x + 7y + 10z = 13. \end{array}$$

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 4 & 7 & 10 & 13 \end{pmatrix},$$

$$-2[1] + [3] \rightarrow [3] \quad \begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 5 & 7 & 9 \\ 0 & 1 & 2 & 3 \end{pmatrix},$$

$$-3[1] + 2[2] \rightarrow [2] \quad \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

We can see, without doing it, that the next step will yield a last row that is all 0's. This corresponds to an equation with x and y eliminated, and having the form $0z = 0$. Since any z at all satisfies this equation, we treat this equation as having the solution $z = t$, a parameter. Now we can backsolve. We have these 2 equations (recall that the last one "disappeared"):

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

Then $y = 3 - 2z = 3 - 2t$, and $2x + 3(3 - 2z) + 4z = 5$ implies $x = -2 + z = -2 + t$.

Check:

$$\begin{aligned} 2x + 3y + 4z &= 5? \\ 2(-2 + z) + 3(3 - 2z) + 4z &= 5 \quad (\text{Yes! The } z\text{'s cancel!}) \end{aligned}$$

$$\begin{aligned} 3x + 5y + 7z &= 9? \\ 3(-2 + z) + 5(3 - 2z) + 7z &= -6 + 15 + (3 - 10 + 7)z = 9 \quad (\checkmark) \end{aligned}$$

$$\begin{aligned} 4x + 7y + 10z &= 13? \\ 4(-2 + z) + 7(3 - 2z) + 10z &= -8 + 21 + (4 - 14 + 10)z = 13 \quad (\checkmark). \end{aligned}$$

Solution: There are infinitely many solutions, describable in terms of one parameter, t : $x = -2 + t$, $y = 3 - 2t$, $z = t$, OR $(x, y, z) = (-2, 3, 0) + t(1, -2, 1)$, t an arbitrary real number. The solution-set has dimension 1, since there is 1 parameter.

Please choose 3 different values of t , now, and check that each of the 3 sets of corresponding x , y and z do indeed solve the equation!

Comment: When there are more unknowns than equations, it turns out (as we'll see later) that there are usually infinitely many solutions, and always if the right-hand-side is all 0's.

Exercises: Given a system, solve it. Given a matrix, apply Gaussian elimination to it.

$$1. \quad \begin{pmatrix} 2 & 0 & 4 & 5 & 3 \\ 5 & 7 & 9 & 4 & 7 \\ 11 & 18 & 21 & 28 & 31 \end{pmatrix}.$$

$$2. \quad \begin{aligned} 2x + 4z + 5w &= 3, \\ 5x + 7y + 9z + 4w &= 7, \\ 11x + 18y + 21z + 28w &= 31. \end{aligned}$$

$$3. \quad \begin{aligned} 2a + 4c + 5d &= 3, \\ 9c + 7b + 5a + 4d &= 7, \\ 18b + 11a + 28d + 21c &= 31. \end{aligned}$$

4.
$$\begin{aligned}x + y + z &= 1, \\x - y + 2z &= 0, \\-x + 3y - 5z &= 3.\end{aligned}$$

5.
$$\begin{pmatrix} -2 & 0 & 4 & 5 \\ 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 \\ 2 & 8 & 3 & 4 \end{pmatrix}.$$

6.
$$\begin{pmatrix} -2 & 0 & 4 & 5 & 3 & 5 & 9 & 7 \\ 4 & 7 & -1 & 8 & 2 & 8 & 3 & 4 \end{pmatrix}.$$

7.
$$\begin{aligned}x + y + 2z &= w, \\x - y + 3z &= 0, \\x + 4y - 6z &= 3w; \text{ the unknowns are } x, y, z, w.\end{aligned}$$

8.
$$\begin{aligned}x + y + z &= 1, \\x - y + 2z &= 0, \\-x + 3y - 5z &= 3, \\x - 3y + 5z &= 3.\end{aligned}$$

9.
$$\begin{pmatrix} 1 & 1 & 1 & 5 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & -1 & 8 \end{pmatrix}.$$

In this section, you have seen by example how to solve a system of linear equations by setting up its matrix, applying Gaussian elimination (with variations to avoid fractions), converting back to simpler equations (having the same solution-set), backsolving, and checking. You have seen examples of solution-sets that contain one point, no points, and infinitely many points (one parameter). Here is a list of the underlined terms, listed by columns in their order of appearance in the text:

linear equation
 solution
 infinitely many
 linear
 unknown
 system of linear equations
 solution-set
 dimension
 linear system
 Gaussian elimination
 matrices
 coefficient
 matrix
 row
 column
 enter
 entry
 variable
 main diagonal
 backsolving

unique
parameter
infinite

1.
$$\begin{aligned} -11x + 14y + 6z + 9p - 14q &= 15, \\ -9x - 9y + 9z + 9p - 6q &= 10, \\ 12x - 3y + 2z + 16p + 13q &= 14, \\ 16x + 14y + 16z + 7p - q &= 14. \end{aligned}$$
2.
$$\begin{aligned} 10x + 14y + 7z + 10p + 10q + 13r + 9s &= 12, \\ 7x + y + 11z + 13p &= 0, \\ 3x + 16y + z - 7p + 12q + 8r - 7s &= 12. \end{aligned}$$
3.
$$\begin{aligned} 12x + 13y + 9z + 11p - 4q &= 9, \\ 2x + 15y - z - 9p + 5q &= 13, \\ 2x + 3y + 5z - 3p + 9q &= 7, \\ 16x + 13y - 13z + 10p + 12q &= 3, \\ -16x + 10y + 6z + 6p + 16q &= 6, \\ 6x + 5y + 15z + 8p + 16q &= 2, \\ 8x + 3y + 11z + 15p + 10q &= -13, \\ 8x + y + 7z + 10p - 11q &= 16, \\ 12x + 14y + 6z + 8p - 15q &= 6. \end{aligned}$$
4.
$$\begin{aligned} 9x + 14y + z - 2p + 10q + 10r + 5s &= 12, \\ -8x + 7y + 15z + 9p - 16q + 16r + 5s &= -2, \\ 16x - 13y + 13z + 11p + 7q &= 0, \\ 5x + 10y - 4z - 11p - 13q + r - 14s &= 15. \end{aligned}$$
5.
$$\begin{aligned} -15x - 11y + 16z + 7p + 6q + 6r + 2s &= 6, \\ -7x + 15y + 3z + 6p + 5q + 4r + 9s &= 7, \\ 7x + 5y - 8z + 12p + 12q + 2r + 11s &= 8. \end{aligned}$$
6.
$$\begin{aligned} -3x + 12y + 5z + 11p - 8q &= 11, \\ 5x + 4y + 7z + 5p &= 0, \\ x + 16y + 13z + 8p - 3q &= 2, \\ 12x + 10y + 5z + 15p + 11q &= 11. \end{aligned}$$
7.
$$\begin{aligned} -3x + 2y + 10z + 7p + 2q + 8r + s &= -16, \\ 8x - 2y + 12z + 10p + 16q + r - 10s &= 0, \\ 14x + 4y + 5z + 8p - 9q + 12r - 2s &= 12, \\ -5x + 9y + z + 7p + 11q + 10r + s &= 5, \\ 8x + 2y + 8z + 14p + 10q + 6r + 2s &= 2, \\ -12x + y + 2z &= 0, \\ 15x + 13y - 15z + 13p + 12q + 4r + 7s &= -10. \end{aligned}$$
8.
$$\begin{aligned} 16x + 3y &= 2, \\ -x + 13y &= 3. \end{aligned}$$
9.
$$\begin{aligned} -4x - 12y + 11z + 11p + 10q - r + 4s + 11t + 8u &= -6, \\ -2x + 10y + 16z + 8p + 13q + 6r - 14s &= 0, \\ 14x + 15y + 16z + 8p + q + 16r + 7s + 12t + 15u &= 1, \\ 10x + 2y - 12z + 7p + 7q + 5r + 2s &= 0, \\ 11x + 16y &= 0. \end{aligned}$$
10.
$$\begin{aligned} 5x + 7y &= 16, \\ 5x + 4y &= 8. \end{aligned}$$
11.
$$\begin{aligned} 4x + 13y + 11z + 7p + 14q &= 13, \\ 2x + 12y + 3z + 12p + 5q &= -2, \\ 12x - 14y - 2z - 3p + 14q &= -5. \end{aligned}$$
12.
$$\begin{aligned} 11x - 2y &= 1, \\ 12x + 13y &= 15, \\ -14x - 5y &= 3, \end{aligned}$$

$$\begin{aligned}x + 3y &= 0, \\7x + 8y &= 9, \\13x + 10y &= 2, \\x + 8y &= -4, \\-11x + 15y &= 11.\end{aligned}$$

13.
$$\begin{aligned}-16x + 2y + 2z + 15p &= 0, \\15x + 6y + 15z + 15p &= -3, \\9x - 12y + 6z + 4p &= -15, \\-7x + 14y + 12z + 15p &= 4, \\-10x + 13y + 5z - 4p &= 3, \\6x + 9y + 14z - 13p &= 4, \\5x + 3y - 3z + 9p &= 12.\end{aligned}$$
14.
$$\begin{aligned}-7x + 10y + 8z + 16p &= -13, \\13x - 8y + 3z - 12p &= -16, \\-5x + y + 15z - 10p &= 7, \\6x + 2y + 8z + 6p &= 14, \\12x - 14y + 13z + 7p &= 7.\end{aligned}$$
15.
$$\begin{aligned}-4x + 5y - 15z - 15p &= 16, \\-2x + y &= 0, \\12x - 11y &= 0, \\11x + 3y + 8z + 14p &= 16.\end{aligned}$$
16.
$$\begin{aligned}15x + 6y + 8z + 15p + 6q &= 16, \\-11x + 9y + 10z + 16p &= 0, \\8x + 3y + 5z + 6p + 10q &= 0, \\-10x - 15y + 7z + 10p + 9q &= 0, \\15x + 8y &= 0, \\x + 13y - 2z - 11p + 15q &= -11, \\11x + 10y + z + 11p + 12q &= 7, \\-6x + 5y + 9z + 14p + 11q &= -13, \\15x + 6y + 10z + 12p + 6q &= 10, \\-2x - 4y + 6z + 5p + 4q &= 8.\end{aligned}$$
17.
$$\begin{aligned}-10x + 3y &= 13, \\-14x - 16y &= 10, \\15x + 15y &= -14, \\4x + 4y &= -15, \\5x + 12y &= 3, \\4x - 2y &= 13, \\7x - 6y &= 8.\end{aligned}$$
18.
$$\begin{aligned}15x - 5y + 6z &= 13, \\-4x + 15y + 13z &= 7.\end{aligned}$$
19.
$$\begin{aligned}11x - 15y + 2z + p + 16q &= 6, \\4x + 11y + 3z + 8p - 13q &= 13, \\3x + 11y + 7z + 6p + 15q &= 6.\end{aligned}$$
20.
$$\begin{aligned}2x - 16y &= 0, \\-3x + 16y + 10z + 15p + 15q &= -3, \\-5x - 4y - 11z - 5p + 14q &= 0, \\4x - 13y + 8z + 16p + 7q &= 12, \\-14x - 16y - 14z + 3p + 11q &= 11, \\2x - 9y - 16z - 11p + 14q &= 10, \\-6x - 3y + 8z + 16p + 4q &= 7, \\15x + 15y - 6z + 6p + 2q &= -5, \\-7x + 8y - 10z + p + 10q &= 4,\end{aligned}$$

- $$7x + 10y + 14z + 15p + 3q = 8.$$
21. $15x + 7y + z + 4p + 6q = 12,$
 $12x - 15y + 5z + 12p + 2q = 5,$
 $-x + 13y + 11z + 11p + 11q = 15.$
22. $4x - y + z = 7,$
 $16x + 2y + 13z = -5,$
 $7x - 16y + 4z = 12,$
 $13x + y + 15z = 13,$
 $15x - 13y - 3z = 0,$
 $-x + 10y - 11z = 16,$
 $-12x + 12y + 3z = 16.$
23. $3x + 10y - 15z + 8p - 10q + 3r + 13s + 14t - 6u + v = -12,$
 $6x + 9y - 5z + 16p + 9q + 13r - 4s + 13t + 16u - 11v = 0,$
 $x + 15y = 0,$
 $-10x + 4y = 0,$
 $-9x + 2y + 13z + 6p = 0,$
 $3x + 2y = 0,$
 $14x + 14y - 15z + 16p + 10q + 16r + 9s + 5t + 8u + 12v = 4,$
 $-3x + 5y = 0.$
24. $5x + 10y + 15z + 8p = 14,$
 $-14x + y - 9z + 3p = 5,$
 $5x - 11y + 10z + 2p = 8,$
 $-14x + 16y - 14z - p = 10,$
 $5x + 11y + 6z + 16p = 4,$
 $8x + 16y + 2z + 15p = 10,$
 $2x + 4y + 5z - 4p = 6.$
25. $11x + 10y + 5z + 8p + 8q + 10r + 7s = 12,$
 $x + 16y + 14z + 7p + 5q + 2r + 9s = 2,$
 $13x - 15y - 6z + 16p + 9q - 6r - 9s = 14,$
 $6x - 7y - 14z + 3p + 9q + 15r + 9s = -9,$
 $14x + y + 9z + 2p + 2q + 16r + 7s = 12,$
 $14x + 2y + 2z + 6p + 15q = 0,$
 $11x + 12y - z + 12p + 6q - 12r - 13s = -3,$
 $10x + 8y + 14z + p + 16q + 7r - 8s = 4,$
 $4x + 16y - 13z + 14p + 9q - 16r = 0,$
 $7x + 8y + 10z + 12p - 4q + 4r + 13s = 0.$
26. $8x + 10y + 6z + 8p + 10q - 10r + 6s + t = 8,$
 $16x - 11y + 5z + 13p + 15q + 5r + 9s = 0,$
 $8x + 12y - 7z + 10p + 11q - 12r + 11s + 7t = 7,$
 $13x + 10y = 0.$
27. $12x - 3y + 8z + 2p + 13q + 9r + 4s = -11,$
 $9x + 5y + 16z + 12p + 8q + 4r + s = 5$
28. $9x - 3y + 14z - 9p + 8q + r + 3s + 10t + 7u + 9v = 1,$
 $16x + 6y + 7z + 5p + 11q + 12r - 6s + t + 15u - 10v = 9,$
 $-14x + 12y + 9z + 16p + 13q + r + 8s + 9t + 6u + 15v = -2,$
 $16x + 8y + 15z + 8p + 11q + 9r - 7s + 13t - 5u + v = 12,$
 $15x + 2y + 14z + 7p + 12q - 3r - 11s + 8t + 3u + 6v = -8,$
 $4x + 4y + 2z - 11p + q + 9r + 5s + 9t + 15u + 5v = 8,$
 $-12x + 11y + 3z - 4p + 13q - 7r + 15s + 3t + 10u - 5v = 14)$
29. $1x + 2y + 1z + 2p + 8q + 13r + 1s$

- $$\begin{aligned}
 &8x + 3y + 16z + 4p + 1q = 0, \\
 &-8x + 11y + 13z - 6p + 5q + 7r + 12s + 15t = 16, \\
 &7x + 7y + 12z + 16p + 6q + 12r - 4s - 12t = 3, \\
 &9x + 1y + 6z - 15p + 12q + 16r + 4s + 3t = 14, \\
 &6x + 6y - 1z + 7p + 10q - 6r + 8s - 9t = 13, \\
 &3x - 5y + 13z + 2p + 13q + 14r - 8s + 3t = 6 \\
 30. &2x + 8y + 16z + 13p + 15q + r + 16s + 10t + 11u = -4, \\
 &12x - 3y + 6z + 15p - 13q - 13r + 8s + 2t \\
 &10x + 12y - 14z - 9p + 6q - 12r + 15s + 12t - 9u = 15, \\
 &2x + 6y + 16z + 13p + 6q + 5r \\
 &2x + 4y + 8z + 13p + 15q + 7r - 10s - 7t + 2u = 9, \\
 &11x + 13y + 11z + p - 7q + 8r - 11s + 15t + 16u = 13, \\
 &10x + 14y + 13z - p + 4q + 13r + 2s - 14t + 13u = 3 \\
 31. &-7x + 13y + 15z - 6p = 4, \\
 &13x + 7y - 3z + 8p = 1. \\
 32. &12x + 10y + 14z - 5p + 12q + 3r + 14s = 2, \\
 &10x + 3y - 16z + 7p + 11q + 6r + 15s = -12, \\
 &7x + 13y + 5z + 4p = 0, \\
 &9x + 14y + 16z - 7p + q - 12r + 6s = 9 \\
 33. &2x + y + 5z + 12p + 8q = 16, \\
 &9x - 6y + 14z + 9p - 2q = 9, \\
 &9x + 2y + 13z + 14p + 9q = -16, \\
 &-9x + 16y + 10z + p + 6q = 13.
 \end{aligned}$$

Section 2 Solving a system of linear equations by Gaussian elimination and backsolving (skip to G. E. defined, next page, for “what to do!”)

In this section, Gaussian elimination will be described technically. I recommend that you read it, noting where terms are defined (underlining is used to indicate a term being defined), then *carefully* (*italics* are used for emphasis) follow the examples, referring to the formal material as needed. The examples given in this section, and those already given in Section 1, illustrate a wide range of “things that can happen” when we try to solve a system of linear equations.

A system of m linear equations in n unknowns consists of m equations of the form

$$(2.1) \quad a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i, \quad i = 1, \dots, m.$$

The quantities a_{ij} are called the coefficients of the system. They are considered to be “known,” and are usually numbers, treated as being constant. The quantities b_i are called the constants of the system. They are also considered to be “known,” and are usually numbers, treated as being constant. Each a_{ij} is associated with the unknown x_j . Thus, each subscript on the a_{ij} ’s has a definite function: the first one, i , tells the equation a_{ij} came from; the second one, j , tells the unknown that a_{ij} goes with (by multiplication). If we are given a system of equations in numerical form, such as the examples in Section 1, and one or more of the unknowns is absent in the i^{th} equation, the corresponding a_{ij} are 0. The x_j , $j = 1, \dots, n$, are the unknowns of the system, and our objective is to “find” all possible n -tuples (x_1, \dots, x_n) of numbers that satisfy the equations, that is, “make *all* m of the equations true when substituted in.” If (x_1, \dots, x_n) does satisfy the equations, it is called a solution (of the system). Section 1 contains some examples. In this section, the method of solving linear equations by Gaussian elimination and backsolving will be described in detail.

The matrix associated with the equations (2.1) is an arrangement of the coefficients a_{ij} and constants b_i into into a “table” with m rows and $n + 1$ columns. The i^{th} row contains, in order,

the numbers a_{ij} followed by the number b_j . When a number is in a matrix, it is called an entry. To locate an entry, we call the entry that is in row i and in column j the ij -th entry. Here is (a picture of) the matrix associated with equations (2.1):

$$(2.2) \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

We also say that (2.1) is the system of the matrix (2.2), or that (2.1) and (2.2) “correspond to each other.” The use of a matrix here is for convenience. Notice that the last column is somehow different than the others. Later, matrices will be used systematically, and the one we use here will be replaced by two of them, one containing just the a_{ij} 's the other the b_j 's. And a third will be added that contains the x_j 's.

The object of Gaussian elimination is to use row operations to change a (system's) matrix into one that has, in each row, a string of consecutive zeroes (leading zeroes) longer by at least one than the string in the row just above it. The first row usually has no leading zeroes - a string of length zero. Such a matrix is in upper triangular form, because there are only zero entries below the main diagonal, and it looks like:

$$(2.3) \quad \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1,n-1} & r_{1n} & c_1 \\ 0 & r_{22} & \dots & r_{2,n-1} & r_{2n} & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & r_{m-1,n-1} & r_{m-1,n} & c_{m-1} \\ 0 & 0 & \dots & 0 & r_{mn} & c_m \end{pmatrix}.$$

Why do this work? Because this new, reduced matrix corresponds to a reduced system:

$$(2.4) \quad \begin{aligned} r_{11}x_1 + r_{12}x_2 + \dots + r_{1,n-1}x_{n-1} + r_{1n}x_n &= c_1, \\ r_{22}x_2 + \dots + r_{2,n-1}x_{n-1} + r_{2n}x_n &= c_2, \\ \dots, \dots, \dots, \dots, \dots, \\ r_{m-1,n-1}x_{n-1} + r_{m-1,n}x_n &= c_{m-1}, \\ r_{mn}x_n &= c_m \end{aligned}$$

A reduced system can be easily “solved” by backsolving. This means, solve the last equation first, then use the solution to solve the next-to-last one, and so on. Section 1 has some examples. It is here that we discover whether there is just one solution, no solutions, or infinitely many. The reduced system has exactly the same solution-set as the original one. This statement requires proof, and will be proved later, perhaps by you! When we backsolve we always get one of 3 outcomes: a single solution, the information that no solution exists, or infinitely many solutions describable in terms of one or more parameters, the number of which is called the dimension of the solution-set.

Gaussian elimination is a rather mechanical process. It can be done straightforwardly on a computer. Usually, clever variations are written into computer programs to increase speed and to reduce round-off error. We will use these 3 basic row operations:

- (1) Interchange 2 rows (corresponding entries are interchanged),
- (2) Multiply a row by a *non-zero* constant (each entry in the row is multiplied by the constant),
- (3) Add a multiple of one row to *another* row. (to add c times row i to row k , replace each entry a_{kj} in row k by $a_{kj} + ca_{ij}$, for $j = 1, \dots, n$, and replace the entry b_k in row k by $b_k + cb_i$).

Each row operation corresponds to an operation that can be performed on a system of equations: (1) interchange 2 equations, (2) multiply an equation by a non-zero number, and (3) add a multiple of one equation to another equation.

It is extremely useful to me to have a way to record the row operations I use, on a particular matrix, so that I can fix errors. There are many handy ways to keep track. One way was introduced in Section 1. We will learn others later, and you can invent ways that others might find useful.

Gaussian elimination defined: This is where what you actually DO is described!

Step 1 Given an $m \times (n+1)$ matrix, find the first row that has a non-zero entry in column 1. If that row is the first row, do nothing. Otherwise, interchange it with row 1. If no entry of column 1 is non-zero, apply this step in succession to the matrices obtained by ignoring zero columns (i.e. columns containing only 0 entries). Treat the first column with a non-zero entry as though it were column 1. There is one exception: if the first n columns are zero columns, then Gaussian elimination is done. Thus, if there is a non-zero column, when Step 1 is done, a matrix is presented to the doer of Step 2 that has a non-zero entry in “row 1,” and “column 1;” this is the entry called a_{11} by the doer of Step 2.

If this zero column happens right away, then the equations the matrix presumably came from may as well not contain the variable corresponding to column 1, for that variable can take any value. This won't happen in practice right off, but does happen when Step 1 is repeated in the course of working on a given matrix! It could also happen immediately in a computer program's environment, so it is necessary to be aware of the possibility.

Step 2 For each row below the first, say the i^{th} row, multiply each entry in row 1 by a_{i1}/a_{11} , and subtract the result from the corresponding entry in row i . The result of this subtraction replaces that corresponding element in row i . In particular, the first entry of row i is replaced by 0.

Step 3 If the row currently being called the first row, is actually where the last row of the original matrix was, or if the column currently being called the first column, is actually where the next-to-last column of the original matrix was, then Gaussian elimination is done. Otherwise, the part of the matrix that is below the row currently being called the first row, and to the right of the column currently being called the first column, is presented to the doer of Step 1 as an $m \times (n+1)$ matrix with new values of m and n that are each one less than they just were, in Step 2.

Steps 4–7 Write down the system of equations that corresponds to the reduced matrix, and back-solve.

Perhaps it's not quite right to call Steps 4–7 part of Gaussian elimination, but many people do, and that's why we *do* Gaussian elimination, when we have a system of equations to solve, so that we *can* backsolve. Steps 4, 5, 6 and 7 will follow shortly.

Backsolving, and presenting the solution

The 3 desired outcomes of solving a system of linear equations are:

- (1) finding a unique set of numbers that make all the equations true when they are used as values of the unknowns,
- (2) finding that no solution exists,
- (3) finding that infinitely many solutions exist, and expressing them in terms of parameters.

The next steps say how to handle these 3 possibilities, and to know why there are no others. Examples and more comments will follow the description of the steps.

Step 4 Check whether m , the number of equations, is at least as large as n , the number of unknowns, *and* whether each of the first n rows of the reduced matrix has a non-zero entry on the

main diagonal (being the entries whose row- and column- subscripts are the same), *and* whether all the remaining rows are completely filled with zeroes, including their last entries.

If the answer is YES, do the following, which is called **backsolving**:

Write down the system of equations that corresponds to the first n rows of the reduced matrix. Solve the last equation, $r_{nn}x_n = c_n$ for $x_n = c_n/r_{nn}$, substitute this value of x_n into the preceding equation, and solve for

$x_{n-1} = (c_{n-1} - (r_{n-1,n}c_n/r_{nn}))/r_{n-1,n-1}$, and so on through the rest of the equations. In this way, n numbers are found, and these numbers, when used in place of the corresponding unknowns, make all the equations true - assuming all the arithmetic was rightly done! No matter how the elimination is done, if it is done correctly, the same numbers will arise when the reduced equations are solved in this manner.

This is so because each operation we performed was one of the 3 basic row operations, and we know the solution-set is not changed when one of these operations is performed. Since our reduced equations lead to specific values for each of the unknowns, the solution-set is precisely (the set containing) (x_1, \dots, x_n) .

The equations are now solved. No more steps need to be done.

If the answer is NO, meaning one of the 3 things you checked is false, go to Step 5.

Step 5 There are 3 ways Step 4 can yield a NO answer, and all of them may occur at once. Let us check them in the order that leads most quickly to the end of the procedure.

If the reduced matrix has a row that has all zeroes except for the last entry (this was the last possibility mentioned in the first part of Step 4), then the reduced system has an equation of the form $0x_1 + 0x_2 + \dots + 0x_n = c \neq 0$, which has no solution. The equations are now solved. *Note* the fact that there are no solutions. *No more steps need to be done.* For other applications, you might want to notice which of the original equations this was, or whether there are other "bad" rows in the reduced matrix. This is where it's handy to keep track of what your row operations were.

If the reduced matrix has *no* row that has all zeroes except for the last entry, *and* if the reduced matrix has 2 consecutive rows whose first non-zero entries are *not* in adjacent columns, the system has infinitely many solutions. *Note all* columns that never get to contain the first non-zero (leading) entry in some row, and go on to Step 6.

If you get this far, it must be true that there are no all-zero rows, the main diagonal has no zeroes on it, *and* there are more unknowns than equations. *Note* that columns $m+1, \dots, n$ are the ones that never get to contain the leading entry in some row. Go on to Step 6.

Step 6 Write down (almost) the system of equations corresponding to the reduced matrix, with this modification: subtract, from both sides of each equation, all the terms that contain one of the unknowns corresponding to a column that didn't get to contain the leading entry in some row. Then replace each of these unknowns, which now appear only on the right-hand side of one of the equations, if they appear at all, by parameters. Use the same parameter for each occurrence of the same unknown. Use different parameters for each of these unknowns. Most people use r, s, t , or t_1, t_2, t_3 , and so on as names for their parameters. (Optional in Step 6) write down one new equation for each of these supplanted unknowns: $x_{i_1} = t_1, x_{i_2} = t_2, x_{i_3} = t_3$, and so on, where $x_{i_1}, x_{i_2}, x_{i_3}$, and so on, are the unknowns that were replaced by parameters. Go to Step 7.

Step 7 Backsolve the modified reduced equations that still involve unknowns *not* replaced by parameters (see the YES part of Step 4 for how to backsolve). This can be done because all unknowns that *could not* be solved for are *assigned* values (their parameters) that can be chosen later; the remaining ones have a non-zero coefficient in at least one equation, and each equation, beginning with the last one, can be solved uniquely for one unknown in terms of the entries in the reduced matrix, your parameters, and previously found unknowns (which are already expressed in

terms of the entries in the reduced matrix and your parameters). Now write down (simplified) equations for the unknowns, $x_i = d_i + e_{i1}t_1 + e_{i2}t_2 + e_{i3}t_3 + \dots$ and so on. The system is solved. There are no more steps. Note that some of the equations will have the form $x_i = t_j$. These are the equations that were optional in Step 6.

A word to the wise: Unless you find that there are no solutions, always check your solution by substituting it into the original system, even if it contains parameters! You might consider it to be:

Step 8 (unofficial) check your answer! We'll see later a way to check your answer even if it is that there is no solution. If you *do* get a solution, the best way to check it will always be to substitute it into the original equations.

Example 1

$$\begin{aligned}x + y + z &= 1, \\x - y + 2z &= 0, \\-x + 3y - 5z &= 3, \\x - 3y + 5z &= 3.\end{aligned}$$

After Step 1,

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

After Step 2,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \\ 0 & 4 & -4 & 4 \\ 0 & -4 & 4 & 2 \end{pmatrix}.$$

Step 3 says, "Look at

$$\begin{pmatrix} * & * & * & * \\ * & -2 & 1 & -1 \\ * & 4 & -4 & 4 \\ * & -4 & 4 & 2 \end{pmatrix}."$$

After Step 1 and Step 2 for the new matrix,

$$\begin{pmatrix} * & * & * & * \\ * & -2 & 1 & -1 \\ * & 0 & -2 & 2 \\ * & 0 & 2 & 4 \end{pmatrix}.$$

Step 3 says, "Look at

$$\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & -2 & 2 \\ * & * & 2 & 4 \end{pmatrix}."$$

After Step 1 and Step 2 for the new matrix,

$$\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & -2 & 2 \\ * & * & 0 & 6 \end{pmatrix}.$$

It's done. Even without going through the motions (as directed in Step 4 and Step 5) of writing down the reduced matrix, we can see that the last equation is going to be $0z = 6$, so there is no solution. As I mentioned before, we'll find a way to check that later on. Right now, please check the arithmetic to see if you agree!

Solution to Example 1: *No solutions exist.*

Comments: Please go back and look at the original equations. If we simply add the last 2, we get $0 = 6$, which can't be true. So we could have seen at once that there was no solution, with no doubting based on the possibility of arithmetic error. But, as I said, the procedure is mechanical, and this but one example. In most cases it's not easy to spot equations that contradict each other. One rule of thumb is that when there are more equations than unknowns we expect that there are no solutions - *but there can be*, as a later example will show.

Example 2

$$(2.5) \quad \begin{aligned} -2x_1 - 4x_3 + 5x_4 + 5x_5 &= 3, \\ 5x_1 + 9x_2 + 10x_3 + 4x_4 &= 7, \\ -x_1 + 8x_2 - 2x_3 + 8x_4 - x_5 &= 3. \end{aligned}$$

This system has more unknowns than equations. We know at once that it has infinitely many solutions if it has any, but we do not know how many parameters will be needed, nor do we know which unknowns are free to take any value whatever, or whether any solutions exist. So, I'll apply Gaussian elimination, recording steps the way it was done in Section 1:

$$\begin{aligned} & \begin{pmatrix} -2 & 0 & -4 & 5 & 5 & 3 \\ 5 & 9 & 10 & 4 & 0 & 7 \\ -1 & 8 & -2 & 8 & -1 & 3 \end{pmatrix}, \\ 5[1]+2[2] \rightarrow [2] \ \& \quad \begin{pmatrix} -2 & 0 & -4 & 5 & 5 & 3 \\ 0 & 18 & 0 & 33 & 25 & 29 \\ 0 & 16 & 0 & 11 & -7 & 3 \end{pmatrix}, \\ -[1]+2[3] \rightarrow [3] \ \text{give} & \\ [2]-[3] \rightarrow [2] \ \& \quad \begin{pmatrix} -2 & 0 & -4 & 5 & 5 & 3 \\ 0 & 2 & 0 & 22 & 32 & 26 \\ 0 & 0 & 0 & -165 & -263 & -205 \end{pmatrix}, \\ -8[2]+[3] \rightarrow [3] \ \text{give} & \end{aligned}$$

The reduced system is:

$$\begin{aligned} -2x_1 - 4x_3 + 5x_4 + 5x_5 &= 3, \\ x_2 + 11x_4 + 16x_5 &= 13, \quad \text{2nd eqn divided by 2,} \\ -165x_4 - 263x_5 &= -205. \end{aligned}$$

Follow Step 6: noticing that x_3 and x_5 never get to start an equation:

$$\begin{aligned} -2x_1 + 5x_4 &= 3 + 4x_3 - 5x_5, \\ x_2 + 11x_4 + 16x_5 &= 13, \\ -165x_4 &= -205 + 263x_5. \end{aligned}$$

Let's use s, t as the parameters, replace x_3 by s , x_5 by t , and add 2 new equations (this is the Optional part of Step 6):

$$\begin{aligned} -2x_1 + 5x_4 &= 3 + 4s - 5t, \\ x_2 + 11x_4 &= 13 - 32t, \\ -165x_4 &= -205 + 263t, \\ x_3 &= s, \end{aligned}$$

$$x_5 = t.$$

This completes Step 6. Now, on to Step 7: The last 2 equations are already “solved.” The third equation has the solution

$$x_4 = (205/165) - (263/165)t = 41/33 - (263/165)t,$$

and this leads the second equation to

$$x_2 = -11(41/33 - (263/165)t) + 13/2 - 16t, \text{ or}$$

$$x_2 = -41/3 + (263/15)t + 13/2 - 16t, \text{ or}$$

$$x_2 = -2/3 + (23/15)t.$$

The first equation yields (after similar arithmetic I hope you will check)

$$-2x_1 + 5x_4 = 3 + 4s - 5t, \text{ or}$$

$$x_1 = 53/33 - 2s - (49/33)t.$$

Solution to Example 2:

$$x_1 = 53/33 - 2s - (49/33)t,$$

$$x_2 = -2/3 + (23/15)t,$$

$$x_3 = s,$$

$$x_4 = 41/33 - (263/165)t,$$

$$x_5 = t.$$

Comment: Often such formulas are written in the form

$$(x_1, x_2, x_3, x_4, x_5) =$$

$$(53/33, -2/3, 0, 41/33, 0) + s(-2, 0, 1, 0, 0) + t(-49/33, 23/15, 0, -263/165, 1).$$

Here, $s(-2, 0, 1, 0, 0)$ means $(-2s, 0, s, 0, 0)$. In general, $c(x_1, \dots, x_n)$ means (cx_1, \dots, cx_n) , where each and every x_j is multiplied by c . This is called scalar multiplication.

Example 3

$$x + y + z = 1,$$

$$x - y + 2z = 0,$$

$$x - 3y + 5z = 3,$$

$$x + 7y - 4z = 0,$$

$$x + 9y - 9z = -7.$$

After Step 1,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ 1 & -3 & 5 & 3 \\ 1 & 7 & -4 & 0 \\ 1 & 9 & -9 & -7 \end{pmatrix}.$$

The remaining matrix steps follow, without comment:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \\ 0 & -4 & 4 & 2 \\ 0 & 6 & -5 & -1 \\ 0 & 8 & -10 & -8 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & -2 & -4 \\ 0 & 0 & -6 & -12 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\begin{aligned} x + y + z &= 1, \\ -2y + z &= -1, \\ 2z &= 4, \\ 0 &= 0, \\ 0 &= 0. \end{aligned}$$

In Step 4 we get a YES answer. The number of unknowns, n , for our matrix is 3. So we backsolve the first 3 equations:

$$2z = 4, \text{ so } z = 2; \quad -2y + 2 = -1, \text{ so } y = 3/2; \quad x + 3/2 + 2 = 1, \text{ so } x = -5/2.$$

Having begun with the assumption that the equations all were true for some triple (x, y, z) as yet unknown we have found that there is indeed a triple, namely $(-5/2, 3/2, 2)$, that solves the equations. The solution is now seen to be unique, assuming the arithmetic done correctly. For, were there others, we would have found parametric formulas for them, and we did not.

Solution to Example 3: $(x, y, z) = (-5/2, 3/2, 2)$ is the only solution.

Comment: Even though there are more equations than unknowns, a unique solution exists. The last 2 equations seem to have disappeared. In a way that is true, but they are still in effect; the Gaussian elimination showed that they are consistent with the first three equations. In Example 1, the last equation was inconsistent with the others, and so the solution-set of the last equation, and the solution-set of the others-holding-all-at-once, have no points in common.

Example 4: What else can happen?

$$\begin{aligned} x + y + z &= 1, \\ x - y + 2z &= 0, \\ x + 3y &= 2, \\ x + 7y - 2z &= 4, \\ x + 11y - 4z &= 6. \end{aligned}$$

The remaining steps follow, with little comment:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ 1 & 3 & 0 & 2 \\ 1 & 7 & -2 & 4 \\ 1 & 11 & -4 & 6 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \\ 0 & 2 & -1 & 1 \\ 0 & 6 & -3 & 3 \\ 0 & 10 & -5 & 5 \end{pmatrix};$$

each of the last 3 rows is a multiple of the second one, so we next get

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

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

z will be replaced by a parameter (why?), say t :

$$\begin{aligned} x + y &= 1 - t, \\ -2y &= -1 - t, \\ z &= t; \end{aligned}$$

backsolve:

$$y = (1 + t)/2, \quad x = (1 - 3t)/2, \quad z = t.$$

Solution to Example 4: $(x, y, z) = (1/2, 1/2, 0) + t(-3/2, 1/2, 1)$; the solution-set is a line. It is described with one parameter, hence has dimension 1.

Comment: Here, there are more equations than unknowns, yet the system has infinitely many solutions!

Example 5: What *else* can happen?

$$(2.6) \quad \begin{aligned} -2x_1 - 4x_3 + 5x_4 + 5x_5 &= 3, \\ 5x_1 + 9x_2 + 10x_3 + 4x_4 &= 7, \\ -x_1 + 9x_2 - 2x_3 + 19x_4 + 15x_5 &= 11. \end{aligned}$$

$$\begin{pmatrix} -2 & 0 & -4 & 5 & 5 & 3 \\ 5 & 9 & 10 & 4 & 0 & 7 \\ -1 & 9 & -2 & 19 & 15 & 11 \end{pmatrix},$$

$$\begin{aligned} 5[1]+2[2] \rightarrow [2] \text{ \& } \\ -[1]+2[3] \rightarrow [3] \text{ give} \end{aligned} \quad \begin{pmatrix} -2 & 0 & -4 & 5 & 5 & 3 \\ 0 & 18 & 0 & 33 & 25 & 29 \\ 0 & 18 & 0 & 33 & 25 & 19 \end{pmatrix}.$$

The next step is to subtract row 2 from row 3. The result is a reduced matrix with a row that has all zeroes except for the last entry. There is no solution.

Solution to Example 5: *No solutions exist.*

Comment: Here we had more unknowns than equations, and yet had *no solutions!*

Lead-in: The preceding examples have all had specific numbers in them. Sometimes the structure of a system, and simple conditions, allow us to recognize at once that any system with the same form has a solution, without doing *any* work. If that is all you need to know, then knowing such “theoretical” things is quite practical: *it can save you work!*

Example 6: Show that, if a , b , and c are all different, then for any d , e , f the system

$$\begin{aligned}x + y + z &= d, \\ax + by + cz &= e, \\a^2x + b^2y + c^2z &= f,\end{aligned}$$

has a unique solution. Once this is known, and you encounter such a system, you'll know it has a unique solution. And this kind of system does come up, for example, in solving ordinary differential equations.

$$\begin{pmatrix} 1 & 1 & 1 & d \\ a & b & c & e \\ a^2 & b^2 & c^2 & f \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & d \\ 0 & b-a & c-a & e-ad \\ 0 & b^2-a^2 & c^2-a^2 & f-a^2d \end{pmatrix}.$$

The next step would *not* be done if $b^2 = a^2$:

$$\begin{pmatrix} 1 & 1 & 1 & d \\ 0 & b-a & c-a & e-ad \\ 0 & 0 & c^2-a^2-(c-a)(b+a) & f-a^2d-(e-ad)(b+a) \end{pmatrix}.$$

Since $c^2 - a^2 - (c - a)(b + a) = (c - a)((c + a) - (b + a)) = (c - a)(c - b) \neq 0$, and since $b - a$ is also non-zero, we know, from Step 4, that the system has a unique solution. This is all that needs to be done to complete Example 6. If you want the solution-formula, just continue the algebraic steps to arrive at it.

Exercises: Solve the following linear systems. Cleverness is encouraged!

1.
$$\begin{aligned}2x + 4z + 5w &= 3, \\5x + 7y + 9z + 4w &= 7, \\x + 4y + 3z + 20w &= 21, \\x - 4y + z + 3w &= 217.\end{aligned}$$

2.
$$\begin{aligned}2a + 4c + 5d &= 3, \\9c + 7b + 5a + 4d &= 7, \\8b + a + 8d + 2c &= 3.\end{aligned}$$

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

4.
$$\begin{aligned}x + y + 2z &= w, \\x - y + 3z &= 0, \\x + 4y - 6z &= 3w.\end{aligned}$$

5.
$$\begin{aligned}x \cos\theta - y \sin\theta &= \cos\omega, \\x \sin\theta + y \cos\theta &= \sin\omega.\end{aligned}$$

6.
$$\begin{aligned}x + y &= 0, \\2x + 2y + z &= 3, \\x - y + 2z &= 4.\end{aligned}$$

7.
$$\begin{aligned}r + 2s + 2t + u &= 0, \\2r + 6s + t + 2u &= 5, \\r + 4s - t + u &= 5, \\r + 5t + u &= -5.\end{aligned}$$

8.
$$\begin{aligned}r + 2s + 2t + u &= 3, \\2r + 6s + t + 2u &= 2, \\r + 4s - t + u &= -1, \\r + 5t + u &= 5.\end{aligned}$$

9.
$$\begin{aligned}r + 2s + 2t + u &= 1, \\2r + 6s + t + 2u &= -1, \\r + 4s - t + u &= -2, \\r + 5t + u &= 4.\end{aligned}$$

$$\begin{aligned}
 10. \quad & r + 2s + 2t + u = 2, \\
 & 2r + 6s + t + 2u = 3, \\
 & r + 4s - t + u = 1, \\
 & r + 5t + u = 1.
 \end{aligned}$$

11. Suppose $AD - BC = 1$. Solve

$$\begin{aligned}
 Ax + By &= u, \\
 Cx + Dy &= v.
 \end{aligned}$$

Note that your solution-formula is a linear system!

12. (a) Solve for x , y , and z :

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

(b) Substitute your values of x , y , and z for the constants of the system below and solve it:

$$\begin{aligned}
 r + s + t &= x, \\
 r - s + tz &= y, \\
 r + s + tz &= z.
 \end{aligned}$$

13. In Exercise 12, substitute the expressions for x , y , and z (in terms of r , s , and t , part (b)) into the system of part (a), simplify, then solve the resulting system for r , s , and t , and compare with the values of r , s , and t you got in Exercise 12.

14. This exercise illustrates the rule of thumb, that a “random” system of equations with the same number of unknowns as equations usually has a unique solution, a system with more equations than unknowns usually has a no solutions, and a system with more unknowns than equations usually has infinitely many. Get a quarter, a dime, a nickel, and a penny. Shake them up well and drop them on a table. Find the value of $2(4Q + 2D + N + 1)(P - 1/2) + 1$, where, each time you toss, $Q = 1$ if the quarter is Heads-up, otherwise $Q = 0$, and so on, with the other coins being treated the same way. Thus, you’d find $2(4 \cdot 0 + 2 \cdot 1 + 1 + 1)(1 - 1/2) + 1 = 5$ if the quarter shows Tails, and each of the others, Heads. If you get 9, change it to 1. If you get 8, ignore it – try again. Do this 12 times, and enter your numbers as the *coefficients* and *constants of the system* in the following system, one at a time, from left to right, starting with the first row and continuing to the second and then the third:

$$\text{____}X + \text{____}Y + \text{____}Z = \text{____},$$

$$\text{____}X + \text{____}Y + \text{____}Z = \text{____},$$

$$\text{____}X + \text{____}Y + \text{____}Z = \text{____}.$$

Now, using the same list of numbers, enter them, in the same way, into this system:

$$\text{____}X + \text{____}Y = \text{____},$$

$$\text{____}X + \text{____}Y = \text{____},$$

$$\text{____}X + \text{____}Y = \text{____},$$

$$\text{____}X + \text{____}Y = \text{____}.$$

Solve both systems, and bring your results to class on the appropriate day. Assuming that your way of shaking up your coins makes every one of the 16 possible outcomes equally likely, your systems are “random.” When I did the experiment I got these numbers: 5, -7, 4, 2, 6, -2, 9, -4, 2, 5, -2, -5.

Conclusion: This section gives you technical details of a step-by-step way to solve systems of linear equations - Gaussian elimination and backsolving. It develops, both descriptively and by examples, the way to present the “solution” of a system - a single n -tuple when the solution is

unique, the assertion that the system has no solution when it does not, and a list of parametric solution-formulas when there are infinitely many solutions. Here are the underlined terms:

system of m linear equations in n unknowns
 coefficients
 constants of the system
 matrix
 entry
 ij -th entry
 leading zeroes
 upper triangular form
 reduced matrix
 reduced system
 backsolving
 dimension
 row operations
Gaussian elimination
 main diagonal
 leading
 parameter
 scalar multiplication
 inconsistent
 consistent

Section 3 Matrices, matrix operations, and their basic properties

Definition of a matrix, and some nomenclature

An m -by- n (notation: $m \times n$, called the size of a) (rectangular) matrix is an arrangement of mn numbers a_{ij} (or other objects - but more on this much later) into a "table" or "array" with m rows and n columns. The i^{th} row contains, in order, the numbers a_{ij} , $j = 1, \dots, n$. The j^{th} column contains, in order, the numbers a_{ij} , $i = 1, \dots, m$. When a number is in a matrix, it is called an entry. To locate an entry, we call the entry that is in row i and in column j the ij -th entry. Here is (a picture of) an $m \times n$ rectangular matrix :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{pmatrix}.$$

We'll use a compact notation, $A = (a_{ij})$, to denote a matrix. When we want to indicate its size, we'll use one of the symbols $(a_{ij})_{m \times n}$ or $A^{m \times n}$. The entries that have equal subscripts, namely those of the form a_{ii} , lie on the main diagonal. The other diagonals parallel to the main diagonal are characterized by constancy of the *difference* between their subscripts. The diagonals perpendicular to the main one don't get used much. They are characterized by constancy of the *sum* of their subscripts. An entry a_{ij} is below the main diagonal if its row number, namely i , is larger than its column number, j . If $i - j < 0$, then a_{ij} lies above the main diagonal. Each entry in a matrix can be thought of as a 1×1 matrix. This is sometimes useful! Each row of an $m \times n$ matrix is itself a matrix - a $1 \times n$ matrix. Thus, row 1 is $(a_{11} \ a_{12} \ \dots \ a_{1n}) = (a_{1j})_{1 \times n}$. Notice that there are no commas separating the entries listed in the row. On the other hand, column 1, also a matrix, is

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}.$$

Since it takes up a lot of space on the page to illustrate a column this way, an

alternative way to do it is to write down a *row* of numbers, *separated by commas*, to denote a column: $(a_{11}, a_{21}, \dots, a_{n1}) = (a_{i1})_{m \times 1}$. It is important to notice that a $1 \times n$ matrix has a *different* size than an $n \times 1$ matrix!

Definition of matrix equality

Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are equal if they have the same size *and* if corresponding entries are equal. That is, they have to have the same numbers of rows and columns, and, for every allowable i and every allowable j , $a_{ij} = b_{ij}$.

Definition of scalar multiplication

A matrix can always be multiplied by a scalar, the word given to describe a number in linear algebra. If $A = (a_{ij})$ is a matrix, and c is a scalar, then cA is the matrix formed by multiplying every entry of A by c . This gives $cA = (ca_{ij})$. Sometimes we write Ac for cA .

Definition of matrix addition

Two matrices (the plural) can be added together *sometimes* - when they have the same size. In that case (and only then!) we define the sum of (a_{ij}) and (b_{ij}) , denoted $(a_{ij}) + (b_{ij})$, by adding corresponding entries. The result: $(a_{ij} + b_{ij})$. The reason for the requirement that the matrices have the same size is that then each entry in each matrix has a corresponding entry in the other one to be added to.

Definition of zero matrix

For each matrix size, there is a special matrix, $0_{m \times n}$, each entry of which is 0. If A is an $m \times n$ matrix, then $A + 0_{m \times n} = A$.

Definition of a matrix multiplication - motivation

For some of you, matrix multiplication is a new concept! It is complicated, but not too much, once you get the idea. The idea is based on something called the "dot product of 2 vectors of the same size." In this context, a vector is an ordered list of numbers. Its size is the number of numbers in the list. If $a = (a_1, a_2, \dots, a_n)$, and $b = (b_1, b_2, \dots, b_n)$ are 2 vectors of size n , their dot product, denoted $a \bullet b$, is the *number* obtained by multiplying corresponding numbers in the list together, and then adding these products. The *formula* for the dot product $a \bullet b$ is

$$a \bullet b = \sum_{i=1}^n a_i b_i.$$

I'll begin with showing how matrix multiplication comes up naturally. In 2 earlier problems, (Exercises 12 and 13, Section 2; look them up NOW, please!) you were asked to solve a system

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = c_i, \quad i = 1, \dots, m,$$

and then to solve a system

$$b_{i1}u_1 + b_{i2}u_2 + \dots + b_{ik}u_k = x_i, \quad i = 1, \dots, n,$$

in which the unknowns in the first system, after being found, become the constants of the system in the second system of equations. Do you agree that those problems fit the description just given?

The point of those problems was to show that you could do such a thing with less work if you first substitute the expressions for the x_i , given by the equations in the second system, into the first equations:

$$a_{i1}(b_{11}u_1 + b_{12}u_2 + \dots + b_{1k}u_k) + a_{i2}(b_{21}u_1 + b_{22}u_2 + \dots + b_{2k}u_k) + \dots \\ \dots + a_{in}(b_{n1}u_1 + b_{n2}u_2 + \dots + b_{nk}u_k) = c_i, \quad i = 1, \dots, m.$$

Then you simplify, which is a straightforward process, and get a new system with new coefficients, and involving only the unknowns u_j . So only one Gaussian elimination needs to be done! Of course, the arithmetic is tedious, and involves a lot of unnecessary writing.

In the new system, what are the coefficients of the ultimate unknowns, u_j ? Let's find the coefficient of u_1 first. In the l^{th} equation, it is the sum of $a_{l1}b_{11}$, $a_{l2}b_{21}$, and so on, with last summand $a_{ln}b_{n1}$. In compact summation notation this is $\sum_{j=1}^n a_{lj}b_{j1}$. Notice that this is a sum of products, the products of the entries in *row* l of the matrix $(a_{ij})_{m \times n}$, and *column* 1 of the matrix $(b_{ij})_{n \times k}$. Thus there is a systematic way to combine the *coefficients* of the 2 systems, without having to write down so many symbols for unknowns, and it forms the basis for the definition of the matrix product. **Important!** Notice that the number of *columns* of (a_{ij}) , the matrix that came first, is the same as the number of *rows* of (b_{ij}) ; it came second. This procedure for finding the coefficients of the new system did not require the writing of the unknowns u_j . They served really as "placeholders." This is the way the product of matrices is calculated.

Definition of matrix multiplication product: in words -

Two matrices can be multiplied together *sometimes* - when they have sizes that fit right. They might fit right in only one order, so it might be possible to form the product AB but not the product BA . The condition is this: the number of columns of the first factor must be equal to the number of rows of the second factor. In that case (and only then!) we define the matrix product of $(a_{ij})_{m \times n}$ and $(b_{ij})_{n \times k}$, denoted $(a_{ij})(b_{ij})$, by forming, in all mk possible ways, the sum of the n products of the entries in a *row* of (a_{ij}) with the *corresponding* n entries of a *column* of (b_{ij}) . There are mk ways to do this, one for each pair of: a row from (a_{ij}) and a column from (b_{ij}) . The result of forming the sum of the products of the entries in row p of (a_{ij}) with the *corresponding* entries of column q of (b_{ij}) becomes the pq^{th} entry of the product matrix, which thus has m rows and k columns. Notice that the pq^{th} entry of the product matrix is the dot product of row p of (a_{ij}) with column q of (b_{ij}) ! The n is "used up" in the process of forming the product, though it still has influence, as we'll see, later.

Definition of identity matrix, square matrix

There is a special matrix associated with square matrices, namely those with the same number of rows as columns. It has 1's on the main diagonal, and 0's elsewhere, and is called the $(n \times n)$ identity matrix, denoted I (or $I_{n \times n}$), because for every $m \times n$ matrix A ,

$$I_{m \times m}A = A, \quad \text{and} \quad AI_{n \times n} = A.$$

Definition of the matrix product: in formulas -

Here is a formula version of the definition of matrix product:

If $A = (a_{ij})$ has size $m \times n$, and $B = (b_{ij})$ has size $n \times p$, then the product, $C = (c_{ij}) = (a_{ij})(b_{ij}) = AB$ exists, has size $m \times p$, and the entry c_{ij} , in row i and column j of $C = AB$, is given by

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}, \quad i = 1, \dots, m; \quad j = 1, \dots, p.$$

Definition of the matrix product: what you DO when their sizes are small-

To *form* the product AB of $A = (a_{ij})$ and $B = (b_{ij})$, first check that the number of columns of A agrees with the number of rows of B , then form the dot product of the i^{th} row of A with the j^{th} column of B , and put the result in the new matrix as its ij^{th} entry. I usually imagine picking up the i^{th} row of A , turning it into a column, and match that column with the j^{th} column of B ,

multiply the matching numbers together, and add them up as I go. This is the usual way people multiply matrices. But there are other useful ways to do the work; more on this later!

Example 1 Find the product

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

Work: The first factor has 4 columns, the second 4 rows, so the product can be formed. It will have size 3×2 . Row 1 is all 1's, so when I match corresponding entries, multiply and add, it just amounts to adding up the entries in column 1: I get $c_{11} = 0$. Next I'll do c_{12} , just add the numbers in column 2, to get $c_{12} = 6$. Working with row 2 of the first factor is not as simple, but since it has a 0 as its last entry, I only have to match, multiply, and add the top 3 entries in the columns of the second factor. I get $c_{21} = 1(-1) + (-1)(-5) + 2 \cdot 1 = 6$, and $c_{22} = 1 \cdot 3 + (-1)3 + 2(-3) = -6$. Row 3 is the most tedious. But I can see some patterns - to find c_{31} I notice the 3's get multiplied times 5's of opposite sign, so they'll cancel, leaving $c_{31} = -4$. Agree? I get $c_{32} = 30$. Agree? If my arithmetic is correct, this gives

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ -1 & 3 & -5 & 3 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ -5 & 3 \\ 1 & -3 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 6 \\ 6 & -6 \\ -4 & 30 \end{pmatrix}.$$

Example 2 Find the product

$$\begin{pmatrix} -2 & 0 & -4 \\ 5 & 5 & 3 \\ 5 & 9 & 10 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix}.$$

This time, the product will be done differently. The new way illustrates a very important way to look at matrix multiplication. Notice that, were I to form the dot products as described before, the entries in column 1 all would get multiplied by 4, those in column 2 would all get multiplied by 3, and those in column 3 by 7. This suggests another way to carry out the matrix multiplication: Multiply column 1 by 4, multiply column 2 by 3 and add that to 4(column 1), then finally multiply column 3 by 7 and add that to the previous total to get the answer. **Work:**

$$4 \begin{pmatrix} -2 \\ 5 \\ 5 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 5 \\ 9 \end{pmatrix} + 7 \begin{pmatrix} -4 \\ 3 \\ 10 \end{pmatrix} = \begin{pmatrix} -8 \\ 20 \\ 20 \end{pmatrix} + \begin{pmatrix} 0 \\ 15 \\ 27 \end{pmatrix} + \begin{pmatrix} -28 \\ 21 \\ 70 \end{pmatrix} = \begin{pmatrix} -36 \\ 56 \\ 117 \end{pmatrix}.$$

You should do this the original way to see that it works. This approach to matrix multiplication is of theoretical importance, but is also useful, for example, in case the column whose entries provide the scalars has some nice pattern, such as, all 1's, in which case, you'd just add the columns of the first factor. A sum of products of matrices by scalars is called a linear combination. This term will be used a lot! In particular, we can say "Ax is the linear combination of the columns of A with coefficients from x."

More terminology A special term is used to describe matrices with more than one entry that have only one row or column: vector. A matrix with one row (and more than one column) is also called a row vector; a matrix with one column (and more than one row) is also called a column vector. The standard "vector" for us is going to be a column vector. A matrix with 1 row and 1 column we treat as a scalar. We *could* treat scalars as 1×1 matrices. But we don't.

Last major matrix operation: transpose We can change a row vector into a column vector, and vice versa, by reversing the row and column subscripts. This can be done to any matrix. Thus, if $A = (a_{ij})$ is an $m \times n$ matrix, $(a_{ji})^t$ is an $n \times m$ matrix, denoted A^t , and read “A transpose.” Although there are other matrix operations (such as the matrix inverse, which *sometimes* exists), all can be described in terms of scalar multiplication, sum, product, and transpose. The basic operations themselves are called scalar multiplication, (matrix) addition, (matrix) multiplication, and transpose.

Example 3 Form the transpose of

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

Answer:

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

Notice that the entries on the main diagonal don't change place. The entries are “reflected” in the diagonal (often used instead of “main diagonal”).

Example 4(sort of)

Let's look at how matrix algebra might be useful. Imagine m things that you can make out of n ingredients according to specific directions - so much of this, so much of that, and so on.

Suppose:

To make 1 unit of thing i , you need a_{ij} units of ingredient j , $i = 1, \dots, m$, $j = 1, \dots, n$;

The cost per unit of ingredient j is c_j , $j = 1, \dots, n$;

The value to you, per unit of thing i , is v_i , $i = 1, \dots, m$.

Suppose you decide to make x_i units of thing i , $i = 1, \dots, m$.

Here are some questions and their answers.

How much of each ingredient will you need?

To make x_i units of thing i , you need $x_i a_{ij}$ units of ingredient j . Thus, the total amount of ingredient j you need is $\sum_{i=1}^m x_i a_{ij}$. This can be expressed as $(x_1 \ x_2 \ \dots \ x_m)(a_{ij})$, the product of a row vector of length m , and an $m \times n$ matrix, and it results in a row vector of length n .

What will be the total cost of the ingredients to be used?

You know, from the answer to the preceding question, how many units of each ingredient you need. So you multiply each amount by its cost per unit and add. The answer is

$$\begin{aligned} \sum_{j=1}^n c_j \left(\sum_{i=1}^m x_i a_{ij} \right) &= \sum_{j=1}^n \left(\sum_{i=1}^m x_i a_{ij} c_j \right) = \\ &= \sum_{j=1}^n \sum_{i=1}^m x_i a_{ij} c_j = \sum_{i=1}^m \sum_{j=1}^n x_i a_{ij} c_j. \end{aligned}$$

The “double sums” are interpreted as though the inner ones had parentheses around them, like the ones in the first line. We can do the adding in any order and get the same answer because only a finite number of terms are involved. In terms of matrix operations, this can be expressed as

$$(x_1 \ x_2 \ \dots \ x_m)(a_{ij}) \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix}.$$

What will be the total value to you of the things made?

Multiply the amount of each thing you make by its value to you per unit, and add. The total is $\sum_{i=1}^m v_i x_i$. This looks like a dot product. In a way, it is. What it “really” is, in this part of the

course, is a matrix product of 2 vectors. In order to make it work as a matrix product, rather than as a dot product, make one of the vectors be a row vector and the other one a column vector, so their sizes match up right. Since we already have treated $(x_1 \ x_2 \ \dots \ x_m)$ as a row vector - note the *absence* of commas - we may as well make a column vector out of the v_i 's: (v_1, v_2, \dots, v_m) - note the *presence* of commas, to express this as a column vector, written horizontally. Then we write $\sum_{i=1}^m v_i x_i = (x_1 \ x_2 \ \dots \ x_m)(v_1, v_2, \dots, v_m)$. The sizes match right - the first factor is $1 \times m$, the second one is $m \times 1$, so the result is a 1×1 matrix, which we regard as a number, oops, a *scalar*, that is.

So far, this is just a possibly handy way to keep things and amounts straight. But here's a different question: Suppose you have y_j units of ingredient j available, $j = 1, \dots, n$. For example, your ingredients might be the vegetables you produce in your garden - about 500 tomatos, 90 big zucchini, a bushel of basil leaves, 2 pounds of garlic, 4 bushels of green beans, 75 beets, and so on. Let's say this is what it seems like you have left over by the time your relatives, friends, neighbors and co-workers won't take any more! So you decide to make up batches of various recipes, of which you have m , and make x_i of each recipe, $i = 1, \dots, m$. The question is, how much of each recipe to make in order to use up all your ingredients? Well, make x_i units of each one, where x_i is an unknown! Recall that $(x_1 \ x_2 \ \dots \ x_m)(a_{ij})$ expresses how much of each ingredient you need. You want to use up all your ingredients, so you write down an equation:

$$(x_1 \ x_2 \ \dots \ x_m)(a_{ij}) = (y_1 \ y_2 \ \dots \ y_n).$$

This can be written more compactly: $x^t A = y^t$.

The equation says that if you make x_i units of each recipe, it will take y_j units - what you have - of each ingredient. When all the matrix operations are carried out, you'll get n equations

$$\sum_{i=1}^m x_i a_{ij} = y_j, \quad j = 1, \dots, n.$$

They don't look exactly like the linear systems we've worked on, but only because of the different order of the factors, so this really is a system of linear equations. So now we could use Gaussian elimination to try to solve the equations. In fact, if we transpose both sides of the matrix equation, we'll get a system in standard form (except that row and column subscripts are reversed):

$$(a_{ji})(x_1, x_2, \dots, x_m) = (y_1, y_2, \dots, y_n).$$

This can be written more compactly: $A^t x = y$. Notice that the subscripts in the matrix got reversed, the row vectors became column vectors, and the order of the factors changed. This new equation has exactly the same solution-set as the other one does, but we need a theorem to justify saying that the transpose of a matrix product is the product of their transposes, *in reverse order*. That will come soon. Right now, all I want you to get is a first glimpse at how a “practical” problem can lead to a system of linear equations. And don't look too closely at the example! It fails to take various other ingredients into consideration: what about salt, pots and pans, storage containers and method of storage, and what about your *time*? Can the system be solved?

Basic properties of the matrix operations

Addition, when it can be done, is commutative and associative, just as it is for numbers. That is, $A + B = B + A$, and $A + (B + C) = (A + B) + C$. This is proved by noticing that the operations are performed entry by entry, and they only involve the addition of numbers. Scalar multiplication also

works as you might expect: $s(tA) = (st)A$, again, because the operation only involves working with one entry at a time. Multiplication, when it can be done, is associative, but is not, in general, commutative. That is, $A(BC) = (AB)C$ always, but $AB \neq BA$, except in special circumstances.

An example of this is seen in the last set of exercises (see #13). Let's see why matrix multiplication, when it can be done, is associative. To do this as clearly as possible, I hope, let's let the ij^{th} entry of a matrix (name of matrix) be denoted $(\text{name of matrix})_{ij}$. Let's suppose that A has size $m \times n$, B has size $n \times p$, and C has size $p \times q$. Then BC has size $n \times q$, so

$$(A(BC))_{ij} = \sum_{k=1}^n A_{ik}(BC)_{kj} = \sum_{k=1}^n A_{ik} \sum_{l=1}^p B_{kl}C_{lj} = \sum_{k=1}^n A_{ik} \sum_{l=1}^p B_{kl}C_{lj} =$$

$$\sum_{k=1}^n \sum_{l=1}^p A_{ik}B_{kl}C_{lj} = \sum_{l=1}^p \sum_{k=1}^n A_{ik}B_{kl}C_{lj}, \text{ because the order of adding up a finite number}$$

of quantities is immaterial. By definition, the inner sum, $\sum_{k=1}^n A_{ik}B_{kl}C_{lj}$, is equal to $(AB)_{il}C_{lj}$,

because C_{lj} is a constant as far as the sum in k is concerned. Thus the whole sum is $\sum_{l=1}^p (AB)_{il}C_{lj}$

. This is, in turn, the definition of $((AB)C)_{ij}$. Since i and j were arbitrary, the matrices $A(BC)$ and $(AB)C$ are equal.

Summary of basic properties of matrix operations:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}, \quad (\text{matrix addition is commutative}).$$

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}, \quad (\text{matrix addition is associative}).$$

$$s(\mathbf{tA}) = (\mathbf{st})\mathbf{A}, \quad (\text{scalar multiplication is associative}).$$

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}, \quad (\text{matrix multiplication is associative}).$$

$$(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t, \quad (\text{transpose of a product is reversed-order product of transposes}).$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}, \quad (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}, \quad (\text{multiplication is distributive over addition}).$$

$$s(\mathbf{A} + \mathbf{B}) = s\mathbf{A} + s\mathbf{B}, \quad (s + t)\mathbf{A} = s\mathbf{A} + t\mathbf{A}, \quad (\text{scalar multiplication is distributive}).$$

The first 4 of these we've checked. The fifth one, $(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t$, has been mentioned, but not checked. Why is it true? The idea of the proof is to compare the expressions defining each side of the equation, entry by entry, and show they are equal. This will then satisfy the definition of matrix equality. Here are the steps, assuming A is $m \times n$ and B is $n \times p$:

$$(\mathbf{AB})_{ij}^t = (\mathbf{AB})_{ji} = \sum_{k=1}^n a_{jk}b_{ki}, \quad i = 1, \dots, p; \quad j = 1, \dots, m.$$

Now $a_{jk} = A_{kj}^t$, and $b_{ki} = B_{ik}^t$, where I've used the convention that putting the subscripts ij on the name of a matrix signifies the ij^{th} entry of that matrix. Therefore,

$$(\mathbf{AB})_{ij}^t = \sum_{k=1}^n a_{jk}b_{ki} = \sum_{k=1}^n A_{kj}^t B_{ik}^t = \sum_{k=1}^n B_{ik}^t A_{kj}^t = (\mathbf{B}^t\mathbf{A}^t)_{ij}.$$

Therefore, corresponding entries of $(\mathbf{AB})_{ij}^t$ and $(\mathbf{B}^t\mathbf{A}^t)_{ij}$ are equal, so by the definition of matrix equality, $(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t$.

Some properties matrix operations DON'T have:

In general,

$\mathbf{AB} \neq \mathbf{BA}$, even if both products can be formed; *sometimes* they are equal;

$\mathbf{AB} = \mathbf{0}$ does NOT imply that one or both of A and B is $\mathbf{0}$; *sometimes* it does;

$\mathbf{AB} = \mathbf{AC}$ does NOT imply that $\mathbf{B} = \mathbf{C}$; *sometimes* it does - same for $\mathbf{BA} = \mathbf{CA}$;

$\mathbf{AX} = \mathbf{B}$ does not always have a solution, X ; *sometimes* it does.

We will see later how to tell when these are and are not true.

There is a very useful way of looking at matrices, that we will need later.

Matrices in block form, and multiplication of matrices in block form

A matrix can be written in block form by partitioning the matrix into blocks, by drawing horizontal or vertical lines, or both, in an array, that separate the entries into rectangular blocks, each of which is then treated as a matrix of the appropriate size. Here is an example:

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 & 5 & 4 \\ 0 & 1 & 7 & 8 & 1 & 6 \\ 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Now, I'll mark off the first 2 rows and columns because it'll emphasize the 2×2 identity matrix there, and do the same to the last 3 rows and columns, for the same reason:

$$\begin{array}{cc|c|ccc} 1 & 0 & 2 & 3 & 5 & 4 \\ 0 & 1 & 7 & 8 & 1 & 6 \\ \hline 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \quad \left(\begin{array}{c} \\ \\ \\ \\ \end{array} \right)$$

This also makes the third column into a pair of blocks. Next, I'll define some matrices, and use them as the entries in another matrix. I do have to make sure there are no "holes" in the finished matrix! This means that each row of matrices has to consist of matrices that all have the same number of rows, and each column of matrices has to consist of matrices that all have the same number of columns. Let

$$A_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, \quad A_{13} = \begin{pmatrix} 3 & 5 & 4 \\ 8 & 1 & 6 \end{pmatrix};$$

$$A_{21} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}, \quad A_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Now we can write $A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}$. Notice that each matrix in row 1 has the same number

of rows. The same can be said of *each* row and column. If $B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{pmatrix}$, and the size of B_{ij} is the same as A_{ij} for each pair i, j , then we can add the matrices by adding corresponding

blocks. No big deal. But now suppose $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{pmatrix}$, and that for each j , A_{ij} has the same

number of rows as B_{jk} . Notice that it doesn't matter what i is, or what k is. We do take for granted that the rules for forming the block matrices A and B are satisfied. Then, the product $A_{ij}B_{jk}$ can be formed. If i and k are fixed, then each of these matrices has the same size, no matter what j is, because the rules for forming the block matrices A and B are satisfied. Thus, the matrices can be added over j (meaning: added as j takes on each of its values, one after the

other), and they form a matrix C_{ik} . We will get a matrix $C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$, and the big deal is

that this matrix, if put back into ordinary form by removing the block boundaries, is precisely the product AB ! Try it with an example, and you'll be convinced. If not, let me know, and I'll go thru the proof, which amounts to making good use of notation for partitioning sets of subscripts; the argument is an exercise in bookkeeping. We will find this block multiplication business useful in talking about solving matrix equations.

Exercises

In 1 - 10, let $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ -1 & 3 & -5 & 3 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 & -4 \\ 5 & 5 & 3 \\ 0 & 18 & 0 \\ 13 & 7 & 11 \end{pmatrix}$.

Perform each operation given in 1 – 10, if it can be done. If not, say why not.

1. $A + B$, 2. $A^t + B$, 3. $A + B^t$, 4. $A^t + B^t$,
5. AB , 6. BA , 7. A^tB , 8. AB^t , 9. B^tA , 10. BA^t .

In 11 - 20, let $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Perform each operation given in 11 – 18.

11. AB , 12. BA , 13. $AB - BA$, 14. A^tB , 15. $A^2 - 4A - 5I$,
16. $B^2 - 4B - 5I$, 17. $A^2 - 5A - 2I$, 18. $B^2 - 5B - 2I$,

In exercises 19 and 20, use the results of some of Exercises 15 - 18.

19. Find a matrix C such that $AC = I$. 20. Find a matrix D such that $DB = I$.

21. Let $x = (1 \ 2 \ 3)$, $y = (4, 5, 6)$. Find xy and y^tx^t .

22. You have 7 bags of cashews and 11 bags of peanuts. Your deluxe mix has twice as much, by weight, of cashews as it has of peanuts. Your regular mix has three times as much, by weight, of peanuts as it has of cashews. How much of each mix should you make to use up your ingredients?

23. Find the product:
$$\begin{pmatrix} a & b & 0 & 0 & 0 \\ c & d & 0 & 0 & 0 \\ 1 & 0 & A & B & C \\ 0 & 1 & D & E & F \end{pmatrix} \begin{pmatrix} d & -b & a & c \\ -c & a & b & d \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

Use block multiplication, and choose the “blocking” to make it as easy as you can.

24. Let $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \\ -1 & 3 & 2 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 2 & 1 \\ -1 & 0 & -3 \\ 0 & -1 & 1 \end{pmatrix}$. Find AB and BA .

25. Show $A(B + C) = AB + AC$, $(A + B)C = AC + BC$, (multiplication is distributive over addition),

26. Show $s(A + B) = sA + sB$, $(s + t)A = sA + tA$, (scalar multiplication is distributive).

size
(rectangular) matrix
row
column
entry
ij-th entry
main diagonal
below
row number
column number
above
equal
corresponding entries
scalar
matrices
sum
dot product
matrix product
($n \times n$) identity matrix
linear combination
vector
row vector
column vector
transpose
block form
block
block multiplication

Section 4: Describing linear systems in terms of matrix equations

The discussion of this material is well-illustrated by a chatty example:

Example 0 Here is the first of many systems of linear equations that will all have the same coefficients, but have different “constants of the system:”

$$x + y + z = 1,$$

$$x - y + 2z = 0,$$

$$-x + 3y - 5z = 3.$$

Recall that to carry out Gaussian elimination we write down

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

To describe this system as a matrix equation, notice that the system asserts that 2 column vectors are equal:

$$\begin{pmatrix} x + y + z \\ x - y + 2z \\ -x + 3y - 5z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

Recall that the vector on the left can be expressed as a linear combination of columns, one associated with each of the unknowns, so the equation becomes:

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} y + \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix} z = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

It is customary to write the scalars on the left. The equation becomes:

$$x \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + y \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} + z \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

Recall that, if A is a matrix, and v is a column vector, then Av is a linear combination of the columns of A . This allows us to see that the following matrix equation is the same as the original equations:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 3 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

Please stop now and check, by doing the multiplication, that this is correct.

This gives us an equation $Ax = b$, where $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 3 & -5 \end{pmatrix}$, $x = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, and $b = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$. This

equation does not, in itself, help us find the solution(s) of the system! If we want to know just this system's solution-set, we'd use Gaussian elimination. But it sometimes happens that we have to solve $Ax = b$ for several b 's. Suppose we also needed to know solutions to $Ax = b_k$, for each $k = 0, 1, 2, 3$, where $b_k = \begin{pmatrix} 1 + k \\ k^2 \\ 3 - k \end{pmatrix}$. Then we'd have 4 Gaussian eliminations to do, and it would

involve a lot of repetitive work, unless we keep track of what we do at each step and just perform those steps on the different vectors b_k . An even better way is to write down a bigger matrix for Gauss elimination, having 4 columns for the constants of the systems:

$$(4.1) \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 3 & 4 \\ 1 & -1 & 2 & 0 & 1 & 4 & 9 \\ -1 & 3 & -5 & 3 & 2 & 1 & 0 \end{pmatrix}.$$

Please check that columns 4, 5, 6, and 7 of this matrix are $b_0, b_1, b_2,$ and b_3 . Now just one Gaussian elimination *could* be done, and all the reduced equations would be easy to backsolve - but it would still take a while! If you disagree, try it and let me know! There is a variation of Gauss elimination, *reduction to row echelon form*, that would essentially do all the backsolving “in” the matrix. We’ll go over it later. But then, I could ask you to find a formula for the solutions for *all* k . I have a better way to suggest, and it arises *because* we have expressed the system in the form of a matrix equation $Ax = b$! Just divide both sides of this equation by A . Oops! We have not defined matrix division! And since multiplication can’t always be done, neither can division. It can *sometimes* be done, but it can’t be done just because sizes match! Well, you might say, maybe it can be done if A is not the 0 matrix with the same size as A . That would be nice, but it can’t be true, because then we could solve any system, and we know some can’t be solved. But it can still be done, in a slightly tricky way, sometimes. In fact, it can be done in this example. We can form a matrix C that makes this equation true: $CA = I$, where I is the 3×3 identity matrix. Then, we *multiply* both sides of the equation $Ax = b_k$ by C : $CAX = Cb_k$. Since $CA = I$, we get $Ix = Cb_k$. Since $Ix = x$, we finally get: $x = Cb_k$. This is the tricky way of dividing by A : multiply by A^{-1} . The matrix just now called C is usually denoted A^{-1} , called the (matrix) inverse of A . We could call it the “reciprocal” of A , but people don’t, despite my thinking it would help to do so! It exists *sometimes*, if the matrix A is square - that is, if A is an $n \times n$ matrix. It has this property: $AA^{-1} = I = A^{-1}A$, where I is the $n \times n$ identity matrix. If you stop and ponder awhile, you’ll agree that this is the meaning of $1/x$: $1/x$ is the solution y of the equation $xy = 1$. Since the matrix product is not commutative, I put down 2 equations to define the matrix inverse: $AA^{-1} = I$, and $A^{-1}A = I$. Here is the matrix $C = A^{-1}$ for this example (you’ll learn soon how to find inverses, if they exist, and how to tell if they exist):

$$A^{-1} = (1/4) \begin{pmatrix} -1 & 8 & 3 \\ 3 & -4 & -1 \\ 2 & -4 & -2 \end{pmatrix}.$$

Please stop now and check that this matrix satisfies both of the equations that define matrix inverse:

$AA^{-1} = I = A^{-1}A$. Remember, the $(1/4)$ in front means to multiply *each* entry of the matrix by $(1/4)$. It's easiest to do this multiplication as your last step.

Now we can get the formula for x_k , the solution of the equation $Ax = b_k$. It is

$$x_k = A^{-1}b_k = (1/4) \begin{pmatrix} -1 & 8 & 3 \\ 3 & -4 & -1 \\ 2 & -4 & -2 \end{pmatrix} \begin{pmatrix} 1+k \\ k^2 \\ 3-k \end{pmatrix} = \begin{pmatrix} 2k^2 - k + 2 \\ -k^2 + k \\ -k^2 + k - 1 \end{pmatrix}.$$

Please check the arithmetic! I was surprised when the answer I got had no fractions! The solution formula can be written as a linear combination that separates the roles of k^2 and k from the part of the solution that does not depend on k :

$$x_k = k^2 \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} + k \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Comments: In summary, to convert a system of linear equations to a matrix equation, we write down, in one matrix, say, A , the coefficients of the system, in a column vector, say, x , the unknowns, equal in number to the number of columns in the matrix, and in another column vector, say b , the constants of the system, equal in number to the number of rows in the matrix. Then we write the equation $Ax = b$. Conversely, given a matrix equation of this sort, we can write down a system of equations, or write down an augmented matrix that includes the coefficients of the system, then the constants of the system, as we do when carrying out a Gauss elimination.

More kinds of matrix equations

We could be given other matrix equations, such as $AX = B$, where A is $m \times n$, B is $m \times p$, so that means X would have to be $n \times p$ to make the multiplication do-able. In fact, this is what the example, with the different b_k 's, amounts to (see (4.1)):

$$(4.2) \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 3 & -5 \end{pmatrix} X = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & 9 \\ 3 & 2 & 1 & 0 \end{pmatrix}.$$

Then we found formulas for x_k , for any k . Let's make a matrix out of 4 of the x 's:

$$(4.3) \quad X = \begin{pmatrix} x_0 & x_1 & x_2 & x_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 & 17 \\ 0 & 0 & -2 & 6 \\ -1 & -1 & -3 & 5 \end{pmatrix},$$

where the first array is meant to suggest a row of column vectors (it's an example of when the entries in a matrix might be something besides a number - a vector, in this case).

Please check now that the matrix X solves the equation (4.2).

We could be given equations in the form $XA = B$. We might think, let's do column operations on the augmented-below matrix (A, B) . That should work! But we can use a trick: the equation $XA = B$ is true if, and only if, $A^t X^t = B^t$. Therefore, all we have to do is solve $A^t Y = B^t$ for Y , and set $X = Y^t$.

What if we're given an equation such as $AXC = B$? Well, first we can see whether $AY=B$ has a solution, then seek a solution to $Y = XC$.

We could be given an equation such as: find X such that $AX = XB$. I don't know a nice way to do this one! But telling you about it gives me the chance to introduce you to something called the Kronecker delta, named after a 19th-century mathematician. This is usually written δ_{ij} , and is defined this way: $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ otherwise. So, if X is $m \times n$, A has to have m columns, and B has to have n rows. Then AX has size $? \times n$, XB has size $m \times ?$. Hence, $? = m$, and $? = n$, so A must be $m \times m$, and B must be $n \times n$. $AX = XB$ means the two matrices must be equal entry-by-entry, so $(AX)_{ij} = (XB)_{ij}$, $1 \leq i \leq m$; $1 \leq j \leq n$. From the definition of matrix multiplication (please stop now, and write down the definition; if you can't, look up the definition, and then write it down), this means $\sum_{k=1}^m a_{ik} x_{kj} = \sum_{k=1}^n x_{ik} b_{kj}$, for $i = 1, \dots, m$; $j = 1, \dots, n$. This gives mn equations in the mn unknowns x_{ij} . There is no neat way to write them down, at least no way I think is neat. But the Kronecker delta gives me a formula that could be put onto a computer:

$$\text{equation } ij \text{ is: } \sum_{r=1}^m \sum_{s=1}^n (a_{ir} \delta_{sj} - b_{sj} \delta_{ir}) x_{rs} = 0.$$

Notice: this is a homogeneous system! It either has only the trivial solution, or has infinitely many non-trivial solutions.

Example: $\begin{pmatrix} 4 & 3 & 2 \\ 4 & 5 & 0 \\ 1 & 3 & 2 \end{pmatrix} X = X \begin{pmatrix} 1 & 3 & 2 \\ 1 & 6 & 8 \\ 0 & 2 & 9 \end{pmatrix}$. This equation implies X is 3×3 , so it

amounts to 9 equations in 9 unknowns. Equation 32, for example, is

$$x_{12} + 3x_{22} + 2x_{32} - 3x_{31} - 6x_{32} - 2x_{33} = 0.$$

Notice that, in **all** the equations discussed, the unknown, whether it is a number or a matrix, appears to the first power, and is not multiplied by any unknown! So these are all linear equations.

In an equation such as $X^t AX = B$, the unknown appears to the first power, but it is in an expression multiplied by X^t , also an unknown, so this equation is not linear. It is quadratic, and is much harder to solve.

Using matrix operations to describe what we do when we do Gaussian elimination

Every time we do a matrix product, say AB , we can look at it in at least 2 ways: A is doing something to B , or B is doing something to A . The way A acts on B here is that each row of A provides n coefficients that multiply corresponding *rows* of B and then add; this can be said in this way:

row i of AB is the $1 \times n$ matrix formed by multiplying row 1 of B by a_{i1} , row 2 of B by a_{i2} , row 3 of B by a_{i3} , and so on, and then adding these rows together, and putting the result in row i of the output matrix AB .

The way B acts on A here is that each column of B provides n coefficients that multiply corresponding *columns* of A and then add; this can be said in this way:

column i of BA is the $n \times 1$ matrix formed by multiplying column 1 of A by b_{1j} , column 2 of A by b_{2j} , column 3 of A by b_{3j} , and so on, and then adding these columns together, and putting the result in column i of the output matrix BA .

We can say this all more compactly by saying that the rows of AB are linear combinations of the rows of B , with coefficients from the respective rows of A , and the columns of AB are linear combinations of the columns of A , with coefficients from the respective columns of B .

Let's build matrices that "do" the three basic row operations when they "act" by multiplication on the left. Please stop now, and write down the definitions of the three basic row operations; if you can't, look up the definitions, and then write them down. Suppose we want to add row 1 to row 3 in the following matrix, and we want a computer to do it, and all the computer can do is matrix operations:

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

We can do that by multiplying on the left by $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Try it! Notice that this matrix has the

formula $I + e_3 e_1^t$, where e_3 stands for the 3×1 matrix that has a 1 in its third row, and zeroes elsewhere, and e_1^t stands for the 1×3 matrix that has a 1 in its first column, and zeroes elsewhere (it is also the transpose of the 3×1 matrix that has a 1 in its third row, and zeroes elsewhere).

Next, suppose we want to subtract the first row from the second one. We can arrange that by multiplying on the left by $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Try it! This matrix has the formula $I - e_2 e_1^t$.

The $m \times 1$ matrices that have exactly one entry of 1 and all the rest 0 are very useful. They are called the standard unit vectors.

Definition of the standard unit vectors

The standard unit vector, e_r , is the $m \times 1$ matrix that has an entry 1 in row r and has all its other entries 0. There is a little ambiguity here! Part of the “standard” meaning is the implicit requirement that $1 \leq r \leq m$. Context, or where it is used in an expression will usually let you know what the size of the unit vector is. We can use the Kronecker delta too: $e_r = (\delta_{ir})_{m \times 1}$ is a compact way to express e_r .

Row operation matrix of type III: add a multiple of one row to another row

To add a times row r to row s of a matrix A , multiply A on the left by $I + a e_s e_r^t$. In words, this matrix has 1's on its main diagonal, and 0's elsewhere, except for its sr entry, which is a .

Row operation matrix of type II: multiply a row by a non-zero scalar

To multiply row i of A by $r \neq 0$, multiply A on the left by $I + (r-1)e_i e_i^t$. In words, this matrix has 1's on its main diagonal, except for its ii entry, which is r , and has 0's elsewhere.

Row operation matrix of type I: interchange two rows

To interchange rows i and j of A , multiply A on the left by $I - e_i e_i^t - e_j e_j^t + e_j e_i^t + e_i e_j^t$. Try it!

Exercises

In 1 - 10, let $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ -1 & 3 & -5 & 3 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 & -4 \\ 5 & 5 & 3 \\ 0 & 18 & 0 \\ 13 & 7 & 11 \end{pmatrix}$.

1. Express the system as a matrix equation:

$$x + y + z = 1,$$

$$x - y + 2z = 0,$$

$$-x + 3y - 5z = 3.$$

2. Express the systems as a single matrix equation:

$$\begin{aligned} x + y + z &= 1, & x + y + z &= 4, \\ x - y + 2z &= 2, & x - y + 2z &= 7, \\ -x + 3y - 5z &= 3; & -x + 3y - 5z &= 9. \end{aligned}$$

3. Express the matrix equation

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ -1 & 3 & -5 & 3 \end{pmatrix} \mathbf{X} = \begin{pmatrix} -2 & 0 & -4 & 5 \\ 5 & 3 & 0 & 18 \\ 0 & 13 & 7 & 11 \end{pmatrix}$$

as systems of equations. Don't write them all out; write one out and say how to get the rest!

4. Find all 2×2 matrices A that satisfy the equation $A^2 = A$.

5. Show that the set of all 2×2 matrices Z of the form $Z = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$ is "nice" in the sense that $Z_1 Z_2 = Z_2 Z_1$ for all such Z 's, that, if $Z = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \neq 0$ (matrix), then $W = (x^2 + y^2)^{-1} \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$

is the matrix inverse of Z . By the way, these 2×2 matrices "are" the complex numbers...

6. (See Exercise 5) Find *two* 2×2 matrices Z , of the form $Z = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$,

such that $Z^2 = \begin{pmatrix} -3 & -4 \\ 4 & -3 \end{pmatrix}$. Begin by finding Z^2 in terms of the unknown entries x and y . Then set up the matrix equation that asserts your matrix is equal to $\begin{pmatrix} -3 & -4 \\ 4 & -3 \end{pmatrix}$. Finally, solve the

equations that you need to, using the quadratic formula. Be sure to choose the signs appropriately!

7. Solve Exercise 6 for an arbitrary non-zero matrix $\begin{pmatrix} u & -v \\ v & u \end{pmatrix}$ in place of $\begin{pmatrix} -3 & -4 \\ 4 & -3 \end{pmatrix}$.

Section 5: How to solve matrix equations $AX = B$, and find determinants

The next thing we'll do is to see how to set up and solve matrix equations, using Gauss elimination (and variations on it). Gauss elimination is *the* tool for doing an actual calculation in linear algebra!

But matrix algebra - especially multiplication - will be useful! We will also see how to find determinants. Gaussian elimination can be used there too!

Given: a matrix equation $AX = B$ (Examples we have seen: $Ax = b$, where A is $m \times n$, and x, b are (column) vectors of size n, m respectively ($n \times 1, m \times 1$, resp.); $CA = I$, and $AC = I$, where A is $n \times n$ and so are C and I (the equations defining the matrix inverse); equation (4.2)). The matrices A and B must have the same numbers of rows. They can have different numbers of columns. The

matrix X has to have as many rows as A has columns, and as many columns as B has, in order for the equation to have any chance of being true. If A is a square matrix, there is a *chance* that A^{-1} exists, and if it does exist, the solution is “just” $X = A^{-1}B$. Then you have to find A^{-1} and multiply it times B *on the left*. You find A^{-1} , if it exists, by augmenting A with an identity matrix of the same size as A , (augment means that you add columns to A) and then you do row operations on it until the part where A started out turns into the identity matrix. And then, the part where you added the identity matrix contains A^{-1} . More on this soon! Or, ...

To start the general procedure write down an augmented matrix, beginning with A , to which you add the columns of B , in order. The augmented matrix looks like this: $(A \mid B)$, schematically. Then do Gauss elimination on the augmented matrix. When you are done, you'll have a reduced matrix, as usual. The reduced matrix looks like $(EA \mid EB)$, where E stands for a matrix that has the effect of carrying out all your row operations “at once.” EA will have zeroes below the main diagonal, and perhaps some all-zero rows. EA occupies the part of the matrix where A started out. I'll call this the “A-block” of the augmented matrix. The other part, where B started out, I'll call the “B-block.” If we start with C , not A , I'll refer to the C -block, etc.

If you get an all-zero row in the A-block, and there is a non-zero entry in the same row of the B-block, then the equation $AX = B$ has no solutions, and this is the end of the procedure.

Example 0: Solve $\begin{pmatrix} 1 & 3 & 6 \\ 1 & 2 & 4 \\ 1 & 5 & 10 \end{pmatrix} X = \begin{pmatrix} 11 & 2 & 1 & 7 \\ 5 & 3 & 9 & 8 \\ 6 & 0 & 10 & 4 \end{pmatrix}$. The augmented matrix is $\begin{pmatrix} 1 & 3 & 6 & 11 & 2 & 1 & 7 \\ 1 & 2 & 4 & 5 & 3 & 9 & 8 \\ 1 & 5 & 10 & 6 & 0 & 10 & 4 \end{pmatrix}$. I'll do this combined step: $-3[1] + 2[2] + [3] \rightarrow [3]$; the result is (please check it!): $\begin{pmatrix} 1 & 3 & 6 & 11 & 2 & 1 & 7 \\ 1 & 2 & 4 & 5 & 3 & 9 & 8 \\ 0 & 0 & 0 & -22 & 0 & 15 & -5 \end{pmatrix}$. The A-block, $\begin{pmatrix} 1 & 3 & 6 \\ 1 & 2 & 4 \\ 0 & 0 & 0 \end{pmatrix}$, has a zero row. The

corresponding row of the B-block has at least 1 non-zero entry - 1 is all it takes - so the equation has no solution. And that's the solution!

When that does not happen, there is more to do. The usual procedure is to backsolve. But now we don't start backsolving, unless there are only a few columns in B . We keep on working on the augmented matrix, using row operations, with the objective of transforming the A -block, as much as possible, into an identity matrix: we want to get 1's on or above the main diagonal, and 0's elsewhere, below the main diagonal *and above the 1's that are on the main diagonal*; we want each row to have at least one more leading 0 than the row preceding it, and we want all the all-zero rows to be together at the bottom. Such a matrix is in row echelon form. *We try to achieve this only in the A -block.* But we must do these row operations to the *whole* augmented matrix! **Example**

1: $\begin{pmatrix} 1 & 3 & 6 & 11 & 2 & 1 & 7 \\ 1 & 2 & 4 & 5 & 3 & 9 & 8 \\ 0 & 0 & 0 & -22 & 0 & 15 & -5 \end{pmatrix}$ is not in row-echelon form. To get row-echelon form, I'll first reduce

it, by subtracting row 1 from row 2: $\begin{pmatrix} 1 & 3 & 6 & 11 & 2 & 1 & 7 \\ 0 & -1 & -2 & -6 & 1 & 8 & 1 \\ 0 & 0 & 0 & -22 & 0 & 15 & -5 \end{pmatrix}$. Next, I'll combine $3[1]+[2]\rightarrow[1]$

and $-[2]\rightarrow[2]$: $\begin{pmatrix} 1 & 0 & 0 & -7 & 5 & 25 & 10 \\ 0 & 1 & 2 & 6 & -1 & -8 & -1 \\ 0 & 0 & 0 & -22 & 0 & 15 & -5 \end{pmatrix}$. Now the matrix is in row-echelon form. Notice the 2

in row 2, column 3. We can't make it into a 0 using row operations! That's because it is not above a leading 1. So it just stays there.

The row-echelon matrix you get *might* have an identity matrix in the A -block. Of course, this can only happen if A is square. But, as in the example, even if A is square, there might be some diagonal entries that are 0, and there might be some non-zero stuff in a row, that comes after the leading 1 in the same row, stuff you can't get rid of using *row* operations. But...

*If you **can** get an identity matrix in the A -block, the desired answer, X , is the matrix in the B -block, the columns originally occupied by B !*

Example 2: Solve $\begin{pmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{pmatrix} X = \begin{pmatrix} 11 & 2 & 1 & 7 \\ 5 & 3 & 9 & 8 \\ 6 & 0 & 10 & 4 \end{pmatrix}$, **and avoid fractions.**

Here is the augmented matrix:

$$\begin{pmatrix} 1 & 3 & 6 & 11 & 2 & 1 & 7 \\ 2 & 5 & 8 & 5 & 3 & 9 & 8 \\ 4 & 7 & 9 & 6 & 0 & 10 & 4 \end{pmatrix}.$$

Gauss elimination to a reduced matrix yields

$$\begin{pmatrix} 1 & 0 & -1 & 7 & -4 & -7 & -5 \\ 0 & 1 & 4 & 17 & 1 & -7 & 6 \\ 0 & 0 & 5 & 47 & -3 & -29 & 6 \end{pmatrix}.$$

I did these steps - probably not what everyone would do, just motivated by a desire to minimize arithmetic: $-2 [2] + [3] \rightarrow [3]$; $[3] + [1] \rightarrow [1]$; $-2 [1] + [2] \rightarrow [2]$; $[3] + [2] \rightarrow [2]$; $[2] + [3] \rightarrow [3]$; $2 [3] + [2] \rightarrow [2]$; $[2] \leftrightarrow [3]$; $- [2] \rightarrow [2]$; $- [3] \rightarrow [3]$. The reduced matrix is not yet in row echelon form. It looks as though it will be possible to do row operations, and make the first 3 columns of this reduced, augmented matrix into a 3×3 identity matrix. I want *to avoid fractions*, so I'll *multiply each row by what it takes to make each diagonal entry equal to the least common multiple of all the diagonal entries in the reduced matrix*. In this example, the least common multiple of all the diagonal entries in the reduced matrix is 5. That means, do $5[1] \rightarrow [1]$, and $5[2] \rightarrow [2]$. Then I continued, with the following steps, to reach a matrix in **modified(for fraction avoidance)** row-echelon form:

$-4 [3] + [2] \rightarrow [2]$; $[3] + [1] \rightarrow [1]$, finally getting

$$\begin{pmatrix} 5 & 0 & 0 & 82 & -23 & -64 & -19 \\ 0 & 5 & 0 & -103 & 17 & 81 & 6 \\ 0 & 0 & 5 & 47 & -3 & -29 & 6 \end{pmatrix}.$$

The next step is to divide each row by 5 to get the desired identity matrix in the first 3 columns, the ones originally occupied by $\begin{pmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{pmatrix}$. However, we won't do that. We'll just pull out

what's left over, beyond that nice 5-times-identity matrix, and write it down as our answer, multiplied by $(1/5)$: $X = (1/5) \begin{pmatrix} 82 & -23 & -64 & -19 \\ -103 & 17 & 81 & 6 \\ 47 & -3 & -29 & 6 \end{pmatrix}$. The last step is to check the answer: substitute

X into the equation, pull the $(1/5)$ out to the left, perform the matrix multiplication, divide each entry in the product by 5, and check that the result is $\begin{pmatrix} 11 & 2 & 1 & 7 \\ 5 & 3 & 9 & 8 \\ 6 & 0 & 10 & 4 \end{pmatrix}$. Please check it now!

Therefore, if A is a square matrix, and you can reduce A 's columns of the augmented matrix to an identity matrix, the problem is completely solved, and it has the unique solution X , comprised of the columns originally occupied by B , and this is the end of the procedure.

This does not always happen, unfortunately, so there may *still* be more to do. In case A is not square, there will surely be more to do! Before going on to the next steps to do, let's look at some examples. The first one will be "nice" - we'll get an identity matrix.

Example 3: Find the inverse, A^{-1} , of $A = \begin{pmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{pmatrix}$. This really is two matrix equations. We have to find a matrix X such that $AX = I = XA$.

IMPORTANT: Here's how to find inverses and determinants!!

First of all, A has to be a square matrix. If it isn't, you can say, "the inverse does not exist." If A is square, then, as usual, build an augmented matrix, starting with the columns of A , in order, then "augment" it by attaching an $n \times n$ identity matrix to play the role of the columns of B . Why? Because the equation we want to solve is $AX = I$, (we also want the equation $XA = I$ to be true - more on that presently). The determinant of a square matrix is a number that we can get as a by-product of the Gaussian elimination we do to find inverses. All I'll do now is to say how to find it, not what it means or why it's important. In our example, this is the augmented matrix:

$$\begin{pmatrix} 1 & 3 & 6 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 4 & 7 & 9 & 0 & 0 & 1 \end{pmatrix}.$$

We'll go thru essentially the same steps as in Example 2: $-4 [1] + [3] \rightarrow [3]$; $-2 [1] + [2] \rightarrow [2]$; $-5 [2] + [3] \rightarrow [3]$; $-1 [2] \rightarrow [2]$. This time we get the reduced matrix

$$\begin{pmatrix} 1 & 3 & 6 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 & -1 & 0 \\ 0 & 0 & 5 & 6 & -5 & 1 \end{pmatrix}.$$

How to find the determinant from the steps leading to the reduced matrix

We did two of the three kinds of row operations: adding a multiple of one row to another row (done 3 times), and multiplying a row by a non-zero constant. We did not interchange any pairs

of rows. To find the determinant, we need to keep track of (i) the numbers we multiply a row by - in this example we multiplied by -1 - and (ii) the number of times we interchanged rows - in this example we interchanged rows 0 times. When we have achieved zeroes below the main diagonal, we then do five things:

- (i) multiply together all the diagonal entries of the reduced matrix. If the result is 0, STOP: the determinant is 0; don't do any more steps. But, if the product isn't 0, - in our example it's 5,
- (ii) multiply together all the numbers we multiplied a row by - -1 in this example -
- (iii) find out whether the number of row changes was even or odd - in this example, 0 is even -
- (iv) multiply the product from (ii) by -1 if the number of row changes was odd - so do nothing in this example - the product stays -1 -
- (v) **divide** the product from (i) by the product from (iv); the result is the determinant of A. In the example, we get $5/(-1) = -5 = \text{determinant of A}$.

Now continue, with $5 [2] \rightarrow [2]$; $5 [1] \rightarrow [1]$; $-4 [3] + [2] \rightarrow [2]$; $-6 [3] + [1] \rightarrow [1]$;

$-3 [2] + [1] \rightarrow [1]$, to get the following matrix, in "fraction-avoiding" row-echelon form:

$$\begin{pmatrix} 5 & 0 & 0 & 11 & -15 & 6 \\ 0 & 5 & 0 & -14 & 15 & -4 \\ 0 & 0 & 5 & 6 & -5 & 1 \end{pmatrix}.$$

Then

$$A^{-1} = (1/5) \begin{pmatrix} 11 & -15 & 6 \\ -14 & 15 & -4 \\ 6 & -5 & 1 \end{pmatrix}.$$

The process has succeeded - we got an identity matrix (times 5) in the A-block. If we can't get one, it's because we wind up with a zero row in the A-block. And *because we started with an identity matrix as B*, it turns out that there MUST be a non-zero entry in each row and column of the I-block (we put I in place of B, so we call that part the I-block). So, the process for finding an inverse (if it exists) is to make an augmented matrix with n rows and $2n$ columns, that has the columns of A as its first n columns, and has the n columns of an identity matrix as its second n columns. Then "clean out" the columns of the A-block, using row operations, to get an identity

matrix, if possible. If not, you conclude that the inverse does not exist. If so, the inverse is what's in the columns where the identity matrix started out.

Now, what about the equation $XA = I$? It really does have to be checked, but that checking is done for us by a theorem that says, if T is a linear transformation on a finite dimensional space, meaning it maps the space into itself, then T has an inverse if and only if it has a right-inverse, namely, if there is a mapping S such that $TS = I$. It's also OK to work with left inverses. **BUT THIS IS ONLY TRUE IN THE CONTEXT OF FINITE DIMENSIONAL SPACES!**

Example 4: This example shows a way to handle a matrix equation of the given form $AX = B$ when you have to "do more work on it." It also introduces 5 important terms, homogeneous equation, particular solution, general solution, and, most important, basis and null space.

$$\text{Solve } \begin{pmatrix} 1 & 3 & 4 & 2 & 5 \\ 3 & 2 & 4 & 5 & 1 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix} X = \begin{pmatrix} 2 & 5 & 3 & 2 \\ 4 & 5 & 1 & 5 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

This might be how you'd combine 4 systems with the same coefficients, but with 4 different vectors making up the constants of each system. There are 2 different approaches we can take. I'll go thru the most straightforward one first; it works for *any* equation $AX = B$. The second approach is potentially easier to use, but more difficult to understand.

First approach

Suppose you're willing to do 4 backsolvings on this to get 4 answers, if any answers exist at all. There is a way to cut down on the work a bit. First do Gauss elimination on the augmented matrix

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

I get the reduced (*not* row-echelon) matrix $\begin{pmatrix} 1 & 3 & 4 & 2 & 5 & 2 & 5 & 3 & 2 \\ 0 & 1 & 14 & 13 & 2 & 6 & 18 & 2 & 11 \\ 0 & 0 & 45 & 45 & 0 & 20 & 58 & 3 & 38 \end{pmatrix}.$

Now work with the first *six* columns. Say you backsolve and get the answer (in terms of unknowns x_1, x_2, x_3, x_4, x_5):

$$x_1 = (8/9) - s + t,$$

$$\begin{aligned}
 (5.1) \quad x_2 &= -(2/9) + s - 2t, \\
 x_3 &= (4/9) - s, \\
 x_4 &= s, \\
 x_5 &= t.
 \end{aligned}$$

This means x_4 and x_5 have been replaced by parameters s and t . Express these formulas as column vectors, as you have done before:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 8 \\ -2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now please, stop a moment, and think what the answers have to be, if you multiply each of the column vectors $h_1 = (-1, 1, -1, 1, 0)$ and $h_2 = (1, -2, 0, 0, 1)$ by A . If you didn't guess, that's OK - but then please do the multiplications now. The answer is, "the vector that, when added to itself, remains unchanged." The vectors h_1 and h_2 are solutions of the homogeneous equation, $Ax = 0$. On the other hand, the vector $P = (1/9)(8, -2, 4, 0, 0)$ is a solution (one of infinitely many) of

$$\begin{pmatrix} 1 & 3 & 4 & 2 & 5 \\ 3 & 2 & 4 & 5 & 1 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix} X = \begin{pmatrix} 2 \\ 4 \\ 4 \end{pmatrix}.$$

This P , or any other solution of the equation, is called a particular solution of the equation. For example, if we let $s = 2/9$, $t = 0$ in (5.1), we get another solution, namely $(2/3, 0, 2/9, 2/9, 0)$, and this one can also be called a particular solution. Since ALL solutions can be obtained by choosing appropriate values of the parameters s and t , we call

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 8 \\ -2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

the general solution. The part of the general solution that involves the parameters s and t can be called the general homogeneous solution. The 2 vectors $(-1, 1, -1, 1, 0)$ and $(1, -2, 0, 0, 1)$ are known as a basis for the null space of the matrix A .

Exercise: Show that, if x_1 and x_2 are particular solutions of $Ax = b$, then $A(x_1 - x_2) = 0$ (the vector), that is, $x_1 - x_2$ is a solution of the homogeneous equation.

Fine. Now, to finish this approach, you *set both of s and t equal to zero in the reduced equations*, and backsolve them, one at a time, with the vectors $(5, 5, 1)$, $(3, 1, 2)$, and $(2, 5, 3)$ taking the place of $(2, 4, 4)$. The reason is to get a particular solution in each case. Then the general solution is your particular solution, plus the same homogeneous solution! This is what you would do if you had 4 different systems to solve, where the only difference was in the constants of the system. But, if you need to solve the original *matrix* equation,

$$\begin{pmatrix} 1 & 3 & 4 & 2 & 5 \\ 3 & 2 & 4 & 5 & 1 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix} X = \begin{pmatrix} 2 & 5 & 3 & 2 \\ 4 & 5 & 1 & 5 \\ 4 & 1 & 2 & 3 \end{pmatrix},$$

there is one more little (ha!) step. You take each general solution formula for the separate columns of B, (these columns being, in this example, $(2, 4, 4)$, $(5, 5, 1)$, $(3, 1, 2)$, and $(2, 5, 3)$), and replace s and t by s_1 and t_1 in the first general solution formula

(getting $(1/9)\begin{pmatrix} 8 \\ -2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ in this example), replace s and t by s_2 and t_2 in

the second general solution formula, and so on. Then you put them all together in an $n \times p$ matrix.

In this example, it would be a 5×4 matrix that could be written as a sum

$$X = P + (-1, 1, -1, 1, 0)(s_1 \ s_2 \ s_3 \ s_4) + (1, -2, 0, 0, 1)(t_1 \ t_2 \ t_3 \ t_4),$$

where the first matrix, P, would have, as its j^{th} column, a particular solution of the "simple" equation $Ax = j^{\text{th}}$ column of $\begin{pmatrix} 2 & 5 & 3 & 2 \\ 4 & 5 & 1 & 5 \\ 4 & 1 & 2 & 3 \end{pmatrix}$. Please write these products out now! You'll first need to do this

Exercise: Find the last 3 columns of P. The first column can be $(2/3, 0, 2/9, 2/9, 0)$.

This example is typical of what you do in situations like this. Notice the peculiar form of the matrices involving the parameters. The products do make sense! They are like that because the homogeneous solutions can be shared from column to column of B. The particular solutions, though, need to be figured out individually. This process works, and, if adopted, completes the description of how to solve matrix equations of the form $AX = B$. Here it is, condensed:

Given $AX = B$, where A is $m \times n$, and B has columns b_1, \dots, b_p . Do Gauss elimination on the augmented matrix $(A \mid B)$, and get a reduced matrix. Set up reduced equations, using only the first $n+1$ columns of the reduced matrix. Solve them, introducing parameters t_1, \dots, t_k , if necessary ($k \geq n-m$, remember).

Express your answer in the form of a general solution, consisting of a particular solution plus a linear combination, involving the parameters t_1, \dots, t_k , of the basis vectors of the null space of A :

$$(x_1, \dots, x_n) = (P_{11}, \dots, P_{1n}) + t_1(Z_{11}, \dots, Z_{1n}) + \dots + t_k(Z_{k1}, \dots, Z_{kn}).$$

Recall that, because of the commas, these are actually column vectors. There are double subscripts on the entries of the particular solution because it involved b_1 's column of the augmented matrix. Now go back to the reduced equations, set all of the parameters equal to zero, and using the next column of the reduced matrix, instead of the one you just used to get your general solution, backsolve, and write down your particular solution as the next column of a matrix $P = (P_{ij})$ that will eventually have p columns, one for each of the columns of B . Repeat until all the columns of the reduced matrix past the n^{th} column have been used. Then write down your answer in the form

$$X = P + \begin{pmatrix} Z_{11} \\ \dots \\ Z_{1n} \end{pmatrix} (t_{11} \dots t_{1p}) + \dots + \begin{pmatrix} Z_{k1} \\ \dots \\ Z_{kn} \end{pmatrix} (t_{k1} \dots t_{kp}).$$

This completes the procedure for solving an equation of the form $AX = B$.

Second approach

$$\text{Solve } \begin{pmatrix} 1 & 3 & 4 & 2 & 5 \\ 3 & 2 & 4 & 5 & 1 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix} X = \begin{pmatrix} 2 & 5 & 3 & 2 \\ 4 & 5 & 1 & 5 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

We notice that B has more columns than A has rows. So doing Gauss elimination looks like a lot of tedious work, especially if we want to avoid fractions. So, we ignore B for the moment, and augment A with a 3×3 identity matrix. We do Gauss elimination on the augmented matrix

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

and get, using the same steps as before, the reduced matrix

$$\begin{pmatrix} 1 & 3 & 4 & 2 & 5 & 1 & 0 & 0 \\ 0 & 1 & 14 & 13 & 2 & 1 & 3 & -2 \\ 0 & 0 & 45 & 45 & 0 & 2 & 11 & -7 \end{pmatrix}.$$

We keep going, trying to get an identity matrix in the first 3 columns of the 3×5 A-block of the augmented matrix. It turns out that we can, and we get, avoiding fractions, the “row-echelon” matrix

$$\begin{pmatrix} 45 & 0 & 0 & 45 & -45 & -14 & 13 & 4 \\ 0 & 45 & 0 & -45 & 90 & 17 & -19 & 8 \\ 0 & 0 & 45 & 45 & 0 & 2 & 11 & -7 \end{pmatrix}$$

It isn't quite a row-echelon matrix, because it doesn't have 1's as the leading entries of the rows.

But it *does* have the same number as the leading entry in each row. So, if we divide the whole augmented matrix by 45, we'll get a matrix with a 3×3 identity matrix in its first 3 columns.

Let's express the first 5 columns of the augmented matrix in block form: $(I \ S)$. That's the A-

block, where A was when we started. Thus, $S = \begin{pmatrix} 1 & -1 \\ -1 & 2 \\ 1 & 0 \end{pmatrix}$. Let's call the last 3 columns of the

augmented matrix E. Thus, $E = (1/45) \begin{pmatrix} -14 & 13 & 4 \\ 17 & -19 & 8 \\ 2 & 11 & -7 \end{pmatrix}$. That's where the 3×3 identity matrix was

when we started. Let's express the unknown matrix, X, a 5×4 matrix, in suitable block form too:

$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, where X_1 is 3×4 , and X_2 is 2×4 . Now we can write down the answer! Let X_2

be an arbitrary 2×4 matrix. It will play the role of the parameters, so let's set $X_2 = T = (t_{ij})$. Then

$X = \begin{pmatrix} X_1 \\ T \end{pmatrix}$, where $X_1 = EB - ST$:

$$X_1 = (1/45) \begin{pmatrix} -14 & 13 & 4 \\ 17 & -19 & 8 \\ 2 & 11 & -7 \end{pmatrix} \begin{pmatrix} 2 & 5 & 3 & 2 \\ 4 & 5 & 1 & 5 \\ 4 & 1 & 2 & 3 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ -1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \end{pmatrix}.$$

Start with an augmented matrix $(A \ I)$. If you can get an identity matrix in the first m columns of the A-block, the desired answer, X, is $X = \begin{pmatrix} EB - ST \\ T \end{pmatrix}$,

where E, $m \times m$, and S, $m \times (n-m)$, come from the final form, $(I \ S \ E)$, of the augmented matrix, after we get the identity matrix in the first m columns of the A-block, and T is an $(n-m) \times p$ parameter matrix, where p is the number of columns of B. This is the end of the procedure, in this case.

Instead of next giving an example to illustrate what happens if you can't get an identity matrix in the first m columns of the A -block, I'll state the condensed version now, and *then* give an example. This part of the second approach is mainly for reference!

But what if we can't get an identity matrix in the first m columns of the A -block?

This happens when, and only when, there are some rows in the reduced matrix that start off with n zeroes. The reduced matrix now has the form $(R \ E)$, where R consists of the first n columns of the reduced matrix, and E occupies the last m columns of the reduced matrix. The matrix R has some all-zero rows now.

There are 1 or 2 steps that remain. Form the matrix product EB . If, for some i , row i of R is zero, and row i of EB is NOT zero, then this is the last step: $AX = E$ has no solutions! What do we do if EB has a zero row wherever R does? The answer is simple, but might seem weird! Remember that R is reduced as much as it can be using row operations. We will now do column operations on the "non-zero part" of R ! Cross out any rows of R that are all zero.

Suppose the new matrix has r rows. The number r is very important in the theory of linear algebra. It is called the (row) rank of A . Call the resulting matrix R^r . Then write down an $n \times n$ identity matrix under the matrix R^r . Do column operations on the matrix $\begin{pmatrix} R^r \\ I \end{pmatrix}$, trying to get an $r \times r$ identity matrix into the first r columns of your big matrix, and zeroes everywhere else in the first r rows of your big matrix. This can be done, with modifications about 1's if you're avoiding fractions. You wind up with a matrix that looks like this, in block form:

$$\begin{pmatrix} I^r & 0 \\ C_1 & C_2 \end{pmatrix}.$$

Here, I^r is an $r \times r$ identity matrix, C_1 is an $n \times r$ matrix, and C_2 is an $n \times (n-r)$ matrix. If you're trying to avoid fractions, you'll have sI^r instead of I^r .

Remember, the last $m-r$ rows of EB are zero rows. Let $(EB)^r$ denote the matrix

consisting of the first r rows of EB . Let T be an $(n-r) \times p$ matrix (of parameters): $T = (t_{ij})$. Write down the answer, X , given by

$$X = (C_1 \ C_2) \begin{pmatrix} (EB)^r \\ T \end{pmatrix}.$$

Those trying to avoid fractions, will have to multiply the right-hand side by $1/s$.

This completes the second approach to solving equations of the form $AX = B$.

By the way, the columns of C_2 constitute a basis for the null space of A .

Example 5: Solve $Ax = b$, where

$$A = \begin{pmatrix} 2 & 4 & 6 & 3 & 45 \\ 2 & 4 & 5 & 2 & 36 \\ 1 & 2 & 3 & 2 & 25 \\ 1 & 2 & 4 & 2 & 29 \end{pmatrix}, \text{ and } b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}.$$

In this example, we are being asked to find for which b 's there is a solution, and for which not, and to find a formula, in terms of the entries of b , and parameters, perhaps, for *all* solutions in case solutions do exist - we know that infinitely many exist, if any do, because there are "more unknowns than equations." First, we set up the augmented matrix

$$(A \ I) = \begin{pmatrix} 2 & 4 & 6 & 3 & 45 & 1 & 0 & 0 & 0 \\ 2 & 4 & 5 & 2 & 36 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 2 & 25 & 0 & 0 & 1 & 0 \\ 1 & 2 & 4 & 2 & 29 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and then use Gauss elimination to find the reduced matrix

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

Then we set up the "augmented-under" matrix to follow, and carry out *column* operations on it:

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

becomes, after doing $-2 [1] + [2] \rightarrow [2]$; $[2] \leftrightarrow [3]$; $[3] \leftrightarrow [4]$; $-3 [1] + [5] \rightarrow [5]$;

$-4 [2] + [5] \rightarrow [5]$; $-5 [3] + [5] \rightarrow [5]$:

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

We can now write down the answer as

$$X = (C_1 \ C_2) \begin{pmatrix} (Eb)^r \\ T \end{pmatrix},$$

where C_1 is the first 3 columns of the 5×5 matrix above, C_2 is the last 2, E consists of the last 4 columns of the reduced matrix, on the previous page, and T is 2×1 . Of course, all this would not be done *unless* the 4th row (entry, that is) of Eb is 0. That is, we need

$$-2b_1 + b_2 + b_3 + b_4 = 0.$$

Exercise: Assuming $-2b_1 + b_2 + b_3 + b_4 = 0$, write out the formula for the general solution of $Ax = b$ in the last example.

Exercise: How can we solve matrix equations of the form $XA = B$?

Exercise: Solve $Ax = b$, where

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \text{ and } b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}.$$

More about determinants

In Example 3, we found the determinant of $A = \begin{pmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{pmatrix}$. We got -5 . We found the

determinant as a by-product of inverting A . You probably noticed that it *really* is a by-product of *reducing* A .

We *denote* the determinant of A in 2 standard ways: by $\det(A)$, usually used when we're referring to a matrix by name, and by $\begin{vmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{vmatrix}$, usually used when we're referring to a matrix

explicitly, by showing its entries, actually or schematically. So for our example we can write $\det(A)$

$$= -5 = \begin{vmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{vmatrix}.$$

What if all we want is the determinant? Is there some procedure we can use that doesn't demand setting up an augmented matrix? Yes! Here are some

Rules for operating on determinants

(D1) The determinant of a matrix A is defined when, and only when, A is square;

(D2) The determinant of a lower triangular matrix is the product of the entries on the main diagonal;

(D3) $\det(A^T) = \det(A)$ (so the preceding rule applies to upper triangular matrices too);

(D4) Interchanging two rows of a matrix just changes the sign of the determinant;

NOTE: This means the determinant of the new matrix - the one with interchanged rows - is the negative of the original one. The original determinant is still the same!

(D5) Multiplying a row by a non-zero scalar, c , multiplies the determinant by c ;

(D6) Adding a multiple of one row to *another* row does not change the determinant ;

(D7) If a row can be expressed as the sum of two row vectors, the determinant can be expressed as the sum of two determinants, each containing a different one of the summands in that row.

These are all the rules needed, in principle, but there are some handy derived rules that can be deduced from (D1) - (D7):

(D8) If two rows are proportional, the determinant is 0.

(D9) If two columns are proportional, the determinant is 0.

(D10) Interchanging two columns of a matrix just changes the sign of the determinant;

(D11) Multiplying a column by a non-zero scalar, c , multiplies the determinant by c ;

(D12) Adding a multiple of one column to *another* column does not change the determinant ;

(D13) If a column can be expressed as the sum of two column vectors, the determinant can be expressed as the sum of two determinants, each containing a different one of the summands in that column.

(D14) If a column, or a row, is zero, the determinant is 0.

There is another bunch of rules that is important. But first, let's use *these* to find $\begin{vmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{vmatrix}$

directly. Our objective is to reach a lower, or upper, triangular determinant. We can use row *and*

column operations! Here is one way to proceed:

$$\begin{vmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 4 & 7 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 6 \\ 2 & 5 & 8 \\ 0 & -3 & -7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 6 \\ 0 & -1 & -4 \\ 0 & -3 & -7 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 6 \\ 0 & 1 & 4 \\ 0 & -3 & -7 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 5 \end{vmatrix} = -5.$$

Steps for each $\Rightarrow -2[2]+[3]\rightarrow[3]$, by (D6) no change; $-2[1]+[2]\rightarrow[2]$, by (D6) no change;

$-[2]\rightarrow[2]$, by (D5) the original determinant is (-1) times the new one; $3[2]+[3]\rightarrow[3]$, by (D6)

no change; by (D2) the last determinant is 5, but is multiplied by -1 to give the final answer.

Example 6: Find $\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$.

$$\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix} = \frac{1}{6} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 3 & 6 & 3 \\ 0 & 0 & 1 & 2 \end{vmatrix} =$$

$$= \frac{1}{6} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 2 \end{vmatrix} = \frac{1}{24} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 4 & 8 \end{vmatrix} = \frac{1}{24} \begin{vmatrix} 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 5 \end{vmatrix} = 5.$$

Here, a lot of row multiplying was done inside, so to compensate, fractions had to be put outside.

After each fraction-avoiding multiply was done, a row was subtracted from the one just below it.

Example 7: Find $\begin{vmatrix} 1 & 3 & 6 \\ 1 & 2 & 4 \\ 1 & 5 & 10 \end{vmatrix}$. We notice that columns 2 and 3 are proportional, so the

determinant is 0, by (D9).

Example 8: Find $\begin{vmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 5 & 7 \end{vmatrix}$. Subtracting row 1 from rows 2 and 3 won't change the

determinant, so $\begin{vmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 5 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{vmatrix} = 0$, by (D8), because the new rows 2 and 3 are

proportional.

Example 9: Find $\begin{vmatrix} 1 & 3 & 5 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{vmatrix}$. If we interchange rows 2 and 3, or columns 2 and 3, we get

an upper triangular matrix, and we change the sign, so the determinant is -6 . Please do it!

Example 10: Find $\begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 1 & 0 & 7 \end{vmatrix}$. The last row can be expressed as $(1 \ 0 \ 0) + (0 \ 0 \ 7)$. So, by

$$\begin{aligned} \text{(D7), } \begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 1 & 0 & 7 \end{vmatrix} &= \begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{vmatrix} + 14 = \\ &= -\begin{vmatrix} 5 & 3 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{vmatrix} + 14 = -10 + 14 = 4. \end{aligned}$$

Rules that allow reduction in the size of a determinant

You might have noticed one way to reduce the size of a determinant. Here is an example to illustrate it.

Example 11: Find $\begin{vmatrix} 8 & a & b \\ 0 & 5 & 2 \\ 0 & 3 & 7 \end{vmatrix}$. This matrix is already partly reduced. So we know that, as we

continue to reduce to an upper triangular determinant, we'll just be copying the first row over and

over, and, in the end, we will just be using the 8 that starts the row. So we can say that

$$\begin{vmatrix} 8 & a & b \\ 0 & 5 & 2 \\ 0 & 3 & 7 \end{vmatrix} = 8 \begin{vmatrix} 5 & 2 \\ 3 & 7 \end{vmatrix} = 8 \begin{vmatrix} 1 & 2 \\ -11 & 7 \end{vmatrix} = 8 \begin{vmatrix} 1 & 2 \\ 0 & 29 \end{vmatrix} = 232.$$

Before we state the next rule, let's look at another example.

Example 12: Find $\begin{vmatrix} 7 & 0 & 5 \\ 2 & 0 & 3 \\ b & 8 & a \end{vmatrix}$. This matrix has a column with all but one of its entries 0. But

it's not the first column, so we will simply put it there, but it will take 2 steps. First, we

interchange rows 1 and 3, change the sign, then switch columns 1 and 2, and change the sign:

$$\begin{vmatrix} 7 & 0 & 5 \\ 2 & 0 & 3 \\ b & 8 & a \end{vmatrix} = -\begin{vmatrix} b & 8 & a \\ 2 & 0 & 3 \\ 7 & 0 & 5 \end{vmatrix} = \begin{vmatrix} 8 & b & a \\ 0 & 2 & 3 \\ 0 & 7 & 5 \end{vmatrix} = 8 \begin{vmatrix} 2 & 3 \\ 7 & 5 \end{vmatrix} = 8 \begin{vmatrix} 2 & 0 \\ 7 & -11/2 \end{vmatrix} = -88.$$

Rule for reducing size when a column has only one non-zero entry:

This rule demands that you figure out a possible sign change! First, the easy part:

Write down the determinant you get by crossing out the row and the column that

contain the single non-zero entry in your column that has only one non-zero entry,

then multiply by that non-zero entry. Here's how to find the sign change: Suppose

that non-zero entry is in row i , col j . Multiply by $(-1)^{i+j}$.

Let's redo Example 12 this way. Start with $\begin{vmatrix} 7 & 0 & 5 \\ 2 & 0 & 3 \\ b & 8 & a \end{vmatrix}$. The single non-zero entry in col 2 is also in row 3. So we do the first part, and get $8 \begin{vmatrix} 7 & 5 \\ 2 & 3 \end{vmatrix}$. Now we multiply by $(-1)^{3+2} = -1$, and get $\begin{vmatrix} 7 & 0 & 5 \\ 2 & 0 & 3 \\ b & 8 & a \end{vmatrix} = -8 \begin{vmatrix} 7 & 5 \\ 2 & 3 \end{vmatrix}$. Please compare this with the first way we did Example 12. It's not exactly the same, but the answer is the same. Please check that it is!

Rule for reducing size when a row has only one non-zero entry:

This rule demands that you figure out a possible sign change! First, the easy part: Write down the determinant you get by crossing out the row and the column that contain the single non-zero entry in your row that has only one non-zero entry, then multiply by that non-zero entry. Here's how to find the sign change: Suppose that non-zero entry is in row i , col j . Multiply by $(-1)^{i+j}$.

The rules just discussed are a special case of the next one. So "many" just means "more than one."

Rule for reducing size when a row has "many" non-zero entries:

To make this rule easy to write down, let's develop some notation for it. When A is an $n \times n$ matrix, let's let A^{ij} stand for the $(n-1) \times (n-1)$ matrix that is obtained when you delete row i and col j from A . Let's let a_{ij} stand for the entry of A that is in row i and col j . Here is the rule, expressed as an equation:

$$\text{For each } i, 1 \leq i \leq n, \det(A) = \sum_{k=1}^n (-1)^{i+k} a_{ik} \det(A^{ik}).$$

This is called "expanding a determinant along the i^{th} row." The most commonly used row, when there is no nice pattern to exploit, is the first row.

Rule for reducing size when a column has "many" non-zero entries:

$$\text{For each } j, 1 \leq j \leq n, \det(A) = \sum_{k=1}^n (-1)^{k+j} a_{kj} \det(A^{kj}).$$

This is called "expanding a determinant along the i^{th} column." The most commonly used column, when there is no nice pattern to exploit, is the first column.

Example 13: Find $\begin{vmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix}$. Let's expand along the first column. To do this, we need A^{k1} ,

for $k = 1, 2, 3, 4$. Here they are:

$$A^{11} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}, \quad A^{21} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}, \quad A^{31} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{vmatrix}, \quad A^{41} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix}.$$

We also need a_{k1} , for $k = 1, 2, 3, 4$. Here they are: $a_{11} = 2, a_{21} = 1, a_{31} = 0, a_{41} = 0$. We have to

$$\text{compute } \sum_{k=1}^4 (-1)^{k+j} a_{k1} \det(A^{k1}) = 2 \det(A^{11}) - 1 \det(A^{21}) + 0 \det(A^{31}) - 0 \det(A^{41}) =$$

$$2 \det(A^{11}) + \det(A^{21}) = 2 \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} - \begin{vmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}. \quad \text{We don't need the last 2 determinants, because}$$

they got multiplied by 0. Such a shame! They were easy to calculate! Look at the second one:

the first row has only one non-zero entry, in row 1, col 1, so we can reduce the size. We get:

$$\det(A) = 2 \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} - \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}. \quad \text{Let's expand the first one along the first column. Please carry out}$$

$$\text{the steps! Here is what I get: } \det(A) = 2(2 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} - \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix}) - \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}. \quad \text{Simplify: } \det(A) = 3 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} \\ - 2 \begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} = 3 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} - 2 \cdot 2 = 3 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} - 4 = 3 \begin{vmatrix} 2 & 1 \\ 0 & 3/2 \end{vmatrix} - 4 = 9 - 4 = 5.$$

When we have a large determinant, we use this method to reduce the size, but we have to compute more determinants, so we try to expand along a row or column that has a lot of zeroes on it. If there are no nice-that-way rows or columns, we might do some row or column operations first to get some zeroes. We have a lot of flexibility. Flexibility requires attention. If you need a method that always works, use the row-reduction method. It is guaranteed to work in about $n^3/3$ multiplication steps.

Formulas for 2×2 and 3×3 determinants

We could use size reduction to find the formula for an arbitrary 2×2 determinant; let's use Gauss elimination instead.

Example 13: Find $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$. Case 1: $a_{11} \neq 0$. Then

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21}. \quad \text{This is actually the answer in both cases.}$$

Case 2: $a_{11} = 0$. Then

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = - \begin{vmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{vmatrix} = -a_{21} a_{12} = 0 a_{22} - a_{21} a_{12} = a_{11} a_{22} - a_{21} a_{12}.$$

Example 14: Find $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$. Now let's use size reduction, this time along the first row.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \text{ Now we'll use the}$$

formula for a 2×2 , and we get

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} (a_{22} a_{33} - a_{23} a_{32}) - a_{12} (a_{21} a_{33} - a_{23} a_{31}) + a_{13} (a_{21} a_{32} - a_{22} a_{31}) =$$

$$a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31}.$$

There is a pattern in the subscripts! In each product, the first subscripts are always 1, 2, 3. When the pattern of second subscripts is in 1-2-3 cyclic order, namely $_1_2_3, _2_3_1, _3_1_2$, the product has a + sign. When the pattern of second subscripts is not in 1-2-3 cyclic order, namely $_1_3_2, _2_1_3, _3_2_1$, the product has a - sign. The pattern for the 2×2 is similar. Please check; do you agree? For $n \times n$ determinants, the pattern is more complex, but the idea is similar.

There are 2 terms in a 2×2 , 6 in a 3×3 , and $n!$ in an $n \times n$ determinant. The reason is, when we set up the first subscripts in order, $1_2_3_ \dots _n_$, there are n choices for the first blank, then only $n-1$ for the second one, and so on, one less each time, until finally at the last blank there is only one unused number, so really no *choice*, but we count it as 1, and take the product, getting $n!$ possible arrangements of second subscripts. To find out what sign goes with each arrangement, we switch them, 2 at a time, until we get them back in natural order. If the number of switches is odd, we multiply the product by -1 ; otherwise, if the number of switches is even, we multiply the product by $+1$. But the methods that have been described allow us to avoid doing all that.

Exercises: Find these determinants.

$$\begin{vmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{vmatrix} \quad \begin{vmatrix} 3 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 3 \end{vmatrix} \quad \begin{vmatrix} 2 & 1 & 0 & -1 \\ 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & 1 \\ -1 & 2 & 1 & 0 \end{vmatrix} \quad \begin{vmatrix} 2 & 3 & 1 & -1 \\ 4 & 5 & 7 & 9 \\ 4 & -1 & 0 & 1 \\ 5 & 6 & 3 & 7 \end{vmatrix}$$

$$\begin{vmatrix} 1+a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & 1+a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & 1+a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & 1+a_4b_4 \end{vmatrix} = \det(I + ab^t), \quad 3 \times 3.$$

$$\begin{vmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{vmatrix}, \quad 6 \times 6 \text{ version of this determinant}; \quad 7 \times 7 \text{ version of this determinant.}$$