Chapter 2

Solving Linear Equations

2.1 Vectors and Linear Equations

The central problem of linear algebra is to solve a system of equations. Those equations are linear, which means that the unknowns are only multiplied by numbers—we never see x times y. Our first linear system is certainly not big. But you will see how far it leads:

$$\begin{array}{ll} \textbf{Two equations} & x - 2y = 1 \\ \textbf{Two unknowns} & 3x + 2y = 11 \end{array} \qquad (1)$$

We begin *a row at a time*. The first equation $\boldsymbol{x - 2y = 1}$ produces a straight line in the xy plane. The point $x = 1, y = 0$ is on the line because it solves that equation. The point $x = 3, y = 1$ is also on the line because $3 - 2 = 1$. If we choose $x = 101$ we find $y = 50$.

The slope of this particular line is $\frac{1}{2}$, because y increases by 1 when x changes by 2. But slopes are important in calculus and this is linear algebra!

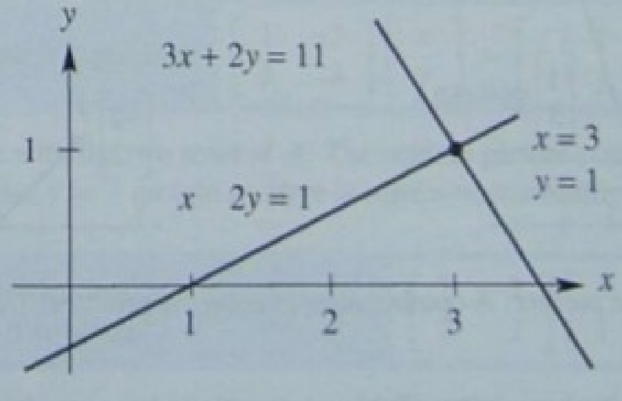

Figure 2.1: *Row picture*: The point (3, 1) where the lines meet is the solution.

Figure 2.1 shows that line $x - 2y = 1$. The second line in this "row picture" comes from the second equation $3x + 2y = 11$. You can't miss the intersection point where the two lines meet. *The point $x = 3$, $y = 1$ lies on both lines.* That point solves both equations at once. This is the solution to our system of linear equations.

> **ROWS** ***The row picture shows two lines meeting at a single point (the solution).***

Turn now to the column picture. I want to recognize the same linear system as a "vector equation". Instead of numbers we need to see *vectors*. If you separate the original system into its columns instead of its rows, you get a vector equation:

Combination equals b $$x\begin{bmatrix}1\\3\end{bmatrix}+y\begin{bmatrix}-2\\2\end{bmatrix}=\begin{bmatrix}1\\11\end{bmatrix}=b. \quad (2)$$

This has two column vectors on the left side. The problem is *to find the combination of those vectors that equals the vector on the right.* We are multiplying the first column by x and the second column by y, and adding. With the right choices $x = 3$ and $y = 1$ (the same numbers as before), this produces $3(\textit{column 1}) + 1(\textit{column 2}) = \boldsymbol{b}$.

> **COLUMNS** ***The column picture combines the column vectors on the left side to produce the vector b on the right side.***

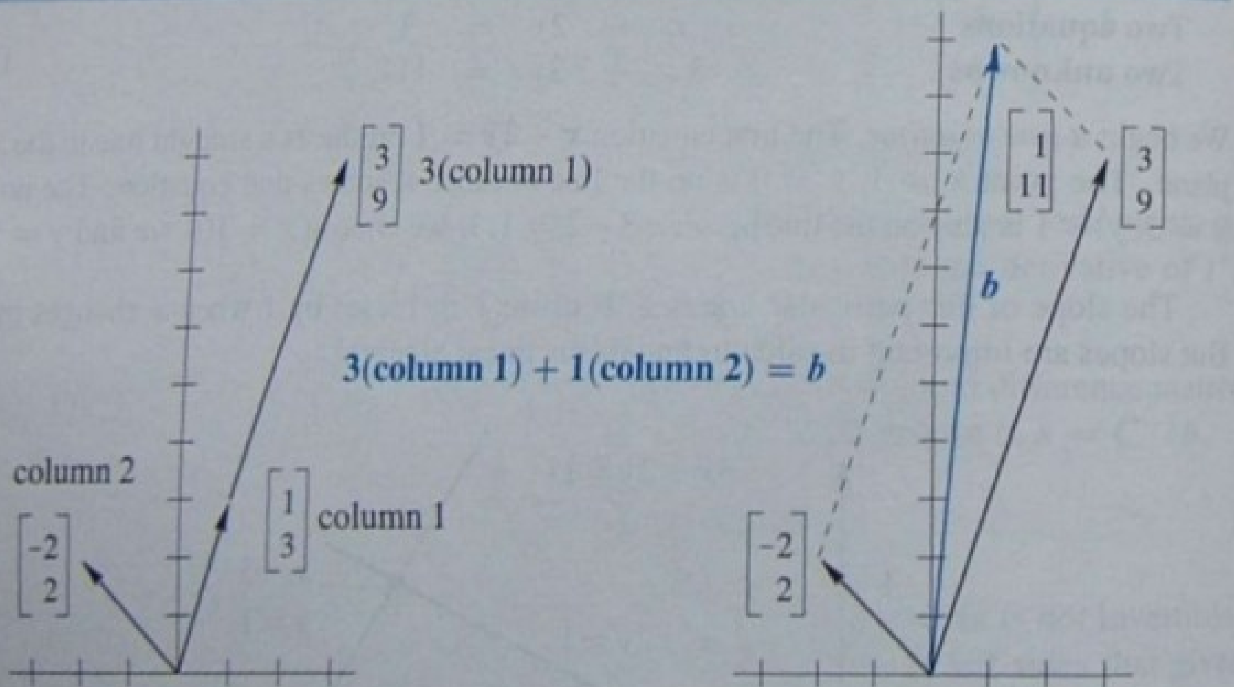

Figure 2.2: *Column picture*: A combination of columns produces the right side (1,11).

Figure 2.2 is the "column picture" of two equations in two unknowns. The first part shows the two separate columns, and that first column multiplied by 3. This multiplication by a *scalar* (a number) is one of the two basic operations in linear algebra:

Scalar multiplication $$3\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}3\\9\end{bmatrix}.$$

If the components of a vector v are v_1 and v_2, then cv has components cv_1 and cv_2.

The other basic operation is *vector addition*. We add the first components and the second components separately. The vector sum is $(1, 11)$ as desired:

$$\textbf{Vector addition} \qquad \begin{bmatrix} 3 \\ 9 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.$$

The right side of Figure 2.2 shows this addition. The sum along the diagonal is the vector $b = (1, 11)$ on the right side of the linear equations.

To repeat: The left side of the vector equation is a ***linear combination*** of the columns. The problem is to find the right coefficients $x = 3$ and $y = 1$. We are combining scalar multiplication and vector addition into one step. That step is crucially important, because it contains both of the basic operations:

$$\textbf{Linear combination} \qquad 3\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.$$

Of course the solution $x = 3, y = 1$ is the same as in the row picture. I don't know which picture you prefer! I suspect that the two intersecting lines are more familiar at first. You may like the row picture better, but only for one day. My own preference is to combine column vectors. It is a lot easier to see a combination of four vectors in four-dimensional space, than to visualize how four hyperplanes might possibly meet at a point. (*Even one hyperplane is hard enough. . .*)

The ***coefficient matrix*** on the left side of the equations is the 2 by 2 matrix A:

$$\textbf{Coefficient matrix} \qquad A = \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}.$$

This is very typical of linear algebra, to look at a matrix by rows and by columns. Its rows give the row picture and its columns give the column picture. Same numbers, different pictures, same equations. We write those equations as a matrix problem $A\boldsymbol{x} = \boldsymbol{b}$:

$$\textbf{Matrix equation} \qquad \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.$$

The row picture deals with the two rows of A. The column picture combines the columns. The numbers $x = 3$ and $y = 1$ go into $\boldsymbol{x}$. Here is matrix-vector multiplication:

$$\begin{array}{l}\textbf{Dot products with rows} \\ \textbf{Combination of columns}\end{array} \qquad A\boldsymbol{x} = \boldsymbol{b} \quad \text{is} \quad \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.$$

Looking ahead This chapter is going to solve n equations in n unknowns (for any n). I am not going at top speed, because smaller systems allow examples and pictures and a complete understanding. You are free to go faster, as long as **matrix multiplication and inversion** become clear. Those two ideas will be the keys to invertible matrices.

I can list four steps to understanding elimination using matrices.

1. Elimination goes from A to a triangular U by a sequence of matrix steps E_{ij}.

2. The inverse matrices E_{ij}^{-1} in reverse order bring U back to the original A.

3. In matrix language that reverse order is $A = LU =$ (lower triangle) (upper triangle).

4. Elimination succeeds if A is invertible. (It may need row exchanges.).

The most-used algorithm in computational science takes those steps (MATLAB calls it lu). But linear algebra goes beyond square invertible matrices! For m by n matrices, $Ax = 0$ may have many solutions. Those solutions will go into a **vector space**. The **rank** of A leads to the **dimension** of that vector space.

All this comes in Chapter 3, and I don't want to hurry. But I must get there.

Three Equations in Three Unknowns

The three unknowns are x, y, z. We have three linear equations:

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

We look for numbers x, y, z that solve all three equations at once. Those desired numbers might or might not exist. For this system, they do exist. When the number of unknowns matches the number of equations, there is *usually* one solution. Before solving the problem, we visualize it both ways:

ROW ***The row picture shows three planes meeting at a single point.***

COLUMN ***The column picture combines three columns to produce*** $(6, 4, 2)$.

In the row picture, each equation produces a *plane* in three-dimensional space. The first plane in Figure 2.3 comes from the first equation $x + 2y + 3z = 6$. That plane crosses the x and y and z axes at the points $(6, 0, 0)$ and $(0, 3, 0)$ and $(0, 0, 2)$. Those three points solve the equation and they determine the whole plane.

The vector $(x, y, z) = (0, 0, 0)$ does not solve $x + 2y + 3z = 6$. Therefore that plane does not contain the origin. The plane $x + 2y + 3z = 0$ does pass through the origin, and it is parallel to $x + 2y + 3z = 6$. When the right side increases to 6, the parallel plane moves away from the origin.

The second plane is given by the second equation $2x + 5y + 2z = 4$. *It intersects the first plane in a line* $\boldsymbol{L}$. The usual result of two equations in three unknowns is a line $\boldsymbol{L}$ of solutions. (Not if the equations were $x + 2y + 3z = 6$ and $x + 2y + 3z = 0$.)

The third equation gives a third plane. It cuts the line $\boldsymbol{L}$ at a single point. That point lies on all three planes and it solves all three equations. It is harder to draw this triple intersection point than to imagine it. The three planes meet at the solution (which we haven't found yet). **The column form will now show immediately why** $z = 2$.

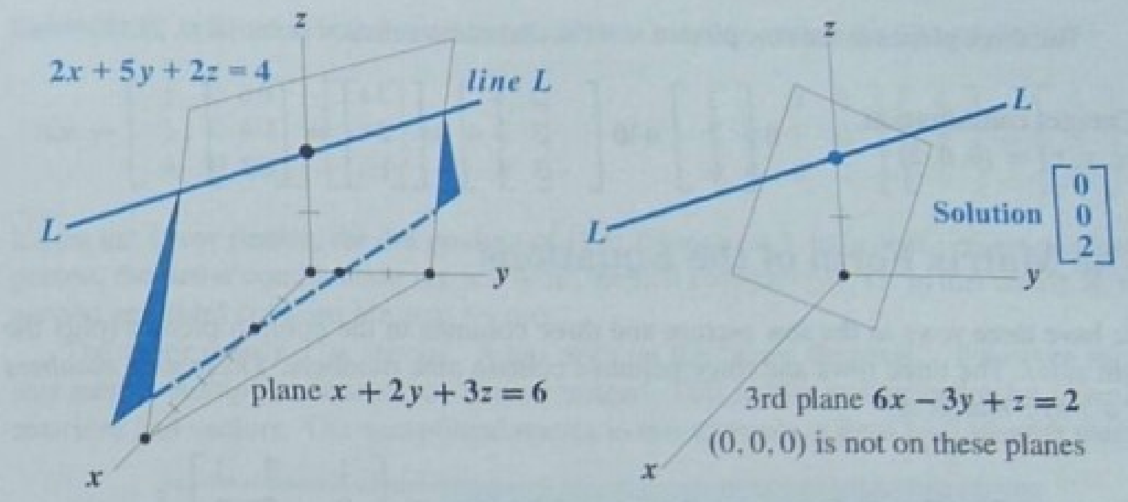

Figure 2.3: *Row picture*: Two planes meet at a line, three planes at a point.

The column picture starts with the vector form of the equations $Ax = b$:

Combine columns $$x\begin{bmatrix}1\\2\\6\end{bmatrix}+y\begin{bmatrix}2\\5\\-3\end{bmatrix}+z\begin{bmatrix}3\\2\\1\end{bmatrix}=\begin{bmatrix}6\\4\\2\end{bmatrix}. \qquad (4)$$

The unknowns are the coefficients x, y, z. We want to multiply the three column vectors by the correct numbers x, y, z to produce $\boldsymbol{b} = (6, 4, 2)$.

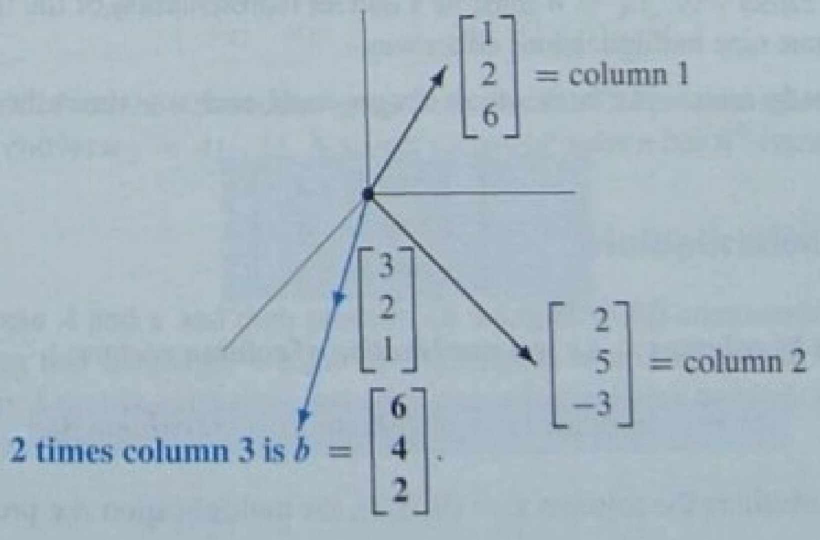

Figure 2.4: *Column picture*: $(x, y, z) = (0, 0, 2)$ because $2(3, 2, 1) = (6, 4, 2) = \boldsymbol{b}$.

Figure 2.4 shows this column picture. Linear combinations of those columns can produce any vector $\boldsymbol{b}$! The combination that produces $\boldsymbol{b} = (6, 4, 2)$ is just 2 times the third column. *The coefficients we need are* $x = 0$, $y = 0$, *and* $z = 2$.

The three planes in the row picture meet at that same solution point $(0,0,2)$:

Correct combination $(x,y,z)=(\mathbf{0},\mathbf{0},\mathbf{2})$
$$\mathbf{0}\begin{bmatrix}1\\2\\6\end{bmatrix}+\mathbf{0}\begin{bmatrix}2\\5\\-3\end{bmatrix}+\mathbf{2}\begin{bmatrix}3\\2\\1\end{bmatrix}=\begin{bmatrix}6\\4\\2\end{bmatrix}.$$

The Matrix Form of the Equations

We have three rows in the row picture and three columns in the column picture (plus the right side). The three rows and three columns contain nine numbers. *These nine numbers fill a 3 by 3 matrix A*:

The "coefficient matrix" in $Ax=b$ is $$A=\begin{bmatrix}1&2&3\\2&5&2\\6&-3&1\end{bmatrix}.$$

The capital letter A stands for all nine coefficients (in this square array). The letter $\boldsymbol{b}$ denotes the column vector with components $6,4,2$. The unknown $\boldsymbol{x}$ is also a column vector, with components x,y,z. (We use boldface because it is a vector, $\boldsymbol{x}$ because it is unknown.) By rows the equations were (3), by columns they were (4), and by matrices they are (5):

Matrix equation $A\boldsymbol{x}=\boldsymbol{b}$
$$\begin{bmatrix}1&2&3\\2&5&2\\6&-3&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\\4\\2\end{bmatrix}. \qquad (5)$$

Basic question: What does it mean to "multiply A times $\boldsymbol{x}$"? We can multiply by rows or by columns. Either way, $A\boldsymbol{x}=\boldsymbol{b}$ must be a correct representation of the three equations. You do the same nine multiplications either way.

Multiplication by rows $\quad A\boldsymbol{x}$ comes from *dot products*, each row times the column $\boldsymbol{x}$:

$$A\boldsymbol{x}=\begin{bmatrix}(\,\textbf{\textit{row 1}}\,)\cdot\boldsymbol{x}\\(\,\textbf{\textit{row 2}}\,)\cdot\boldsymbol{x}\\(\,\textbf{\textit{row 3}}\,)\cdot\boldsymbol{x}\end{bmatrix}. \qquad (6)$$

Multiplication by columns $\quad A\boldsymbol{x}$ is a ***combination of column vectors***:

$$A\boldsymbol{x}=x\,(\textbf{\textit{column 1}})+y\,(\textbf{\textit{column 2}})+z\,(\textbf{\textit{column 3}}). \qquad (7)$$

When we substitute the solution $\boldsymbol{x}=(0,0,2)$, the multiplication $A\boldsymbol{x}$ produces $\boldsymbol{b}$:

$$\begin{bmatrix}1&2&3\\2&5&2\\6&-3&1\end{bmatrix}\begin{bmatrix}0\\0\\2\end{bmatrix}=\textbf{\textit{2 times column 3}}=\begin{bmatrix}6\\4\\2\end{bmatrix}.$$

The dot product from the first row is $(1,2,3)\cdot(0,0,2)=6$. The other rows give dot products 4 and 2. ***This book sees $A\boldsymbol{x}$ as a combination of the columns of A.***

Example 1 Here are 3 by 3 matrices A and I = identity, with three 1's and six 0's:

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} \qquad Ix = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

If you are a row person, the dot product of $(1, 0, 0)$ with $(4, 5, 6)$ is 4. If you are a column person, the linear combination Ax is 4 times the first column $(1, 1, 1)$. In that matrix A, the second and third columns are zero vectors.

The other matrix I is special. It has ones on the "main diagonal". *Whatever vector this matrix multiplies, that vector is not changed.* This is like multiplication by 1, but for matrices and vectors. The exceptional matrix in this example is the 3 by 3 ***identity matrix***:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{always yields the multiplication} \quad Ix = x\,.$$

Matrix Notation

The first row of a 2 by 2 matrix contains a_{11} and a_{12}. The second row contains a_{21} and a_{22}. The first index gives the row number, so that a_{ij} is an entry in row i. The second index j gives the column number. But those subscripts are not very convenient on a keyboard! Instead of a_{ij} we type $A(i, j)$. ***The entry*** $a_{57} = A(5, 7)$ ***would be in row 5, column 7.***

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} A(1,1) & A(1,2) \\ A(2,1) & A(2,2) \end{bmatrix}.$$

For an m by n matrix, the row index i goes from 1 to m. The column index j stops at n. There are mn entries $a_{ij} = A(i, j)$. A square matrix of order n has n^2 entries.

Multiplication in MATLAB

I want to express A and x and their product Ax using MATLAB commands. This is a first step in learning that language. I begin by defining the matrix A and the vector x. This vector is a 3 by 1 matrix, with three rows and one column. Enter matrices a row at a time, and use a semicolon to signal the end of a row:

$$A = [1 \;\; 2 \;\; 3; \;\; 2 \;\; 5 \;\; 2; \;\; 6 \; -3 \;\; 1]$$
$$x = [0; 0; 2]$$

Here are three ways to multiply Ax in MATLAB. In reality, $A * x$ is the good way to do it. MATLAB is a high level language, and it works with matrices:

Matrix multiplication $b = A * x$

2.2 The Idea of Elimination

This chapter explains a systematic way to solve linear equations. The method is called ***"elimination"***, and you can see it immediately in our 2 by 2 example. Before elimination, x and y appear in both equations. After elimination, the first unknown x has disappeared from the second equation $8y = 8$:

$$\textbf{Before}\quad \begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned} \qquad \textbf{After}\quad \begin{aligned} x - 2y &= 1 \\ 8y &= 8 \end{aligned} \qquad \begin{aligned} &(\textit{multiply equation } 1 \textit{ by } 3) \\ &(\textit{subtract to eliminate } 3x) \end{aligned}$$

The new equation $8y = 8$ instantly gives $y = 1$. Substituting $y = 1$ back into the first equation leaves $x - 2 = 1$. Therefore $x = 3$ and the solution $(x, y) = (3, 1)$ is complete.

Elimination produces an ***upper triangular system***—this is the goal. The nonzero coefficients $1, -2, 8$ form a triangle. That system is solved from the bottom upwards—first $y = 1$ and then $x = 3$. This quick process is called ***back substitution***. It is used for upper triangular systems of any size, after elimination gives a triangle.

Important point: The original equations have the same solution $x = 3$ and $y = 1$. Figure 2.5 shows each system as a pair of lines, intersecting at the solution point $(3, 1)$. After elimination, the lines still meet at the same point. Every step worked with correct equations.

How did we get from the first pair of lines to the second pair? We subtracted 3 times the first equation from the second equation. The step that eliminates x from equation 2 is the fundamental operation in this chapter. We use it so often that we look at it closely:

To eliminate x: Subtract a multiple of equation 1 from equation 2.

Three times $x - 2y = 1$ gives $3x - 6y = 3$. When this is subtracted from $3x + 2y = 11$, the right side becomes 8. The main point is that $3x$ cancels $3x$. What remains on the left side is $2y - (-6y)$ or $8y$, and x is eliminated. **The system became triangular.**

Ask yourself how that multiplier $\ell = 3$ was found. The first equation contains $1x$. ***So the first pivot was* 1** (the coefficient of x). The second equation contains $3x$, **so the multiplier was 3**. Then subtraction $3x - 3x$ produced the zero and the triangle.

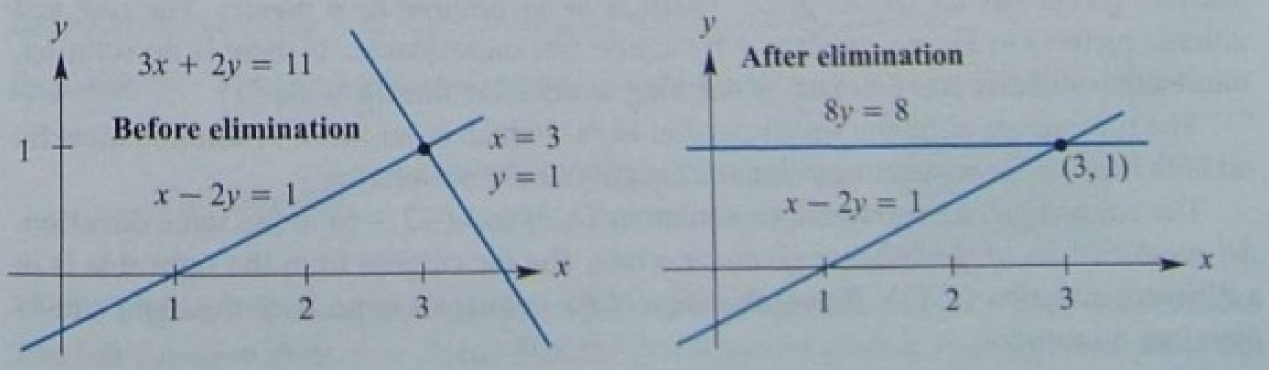

Figure 2.5: Eliminating x makes the second line horizontal. Then $8y = 8$ gives $y = 1$.

You will see the multiplier rule if I change the first equation to $4x - 8y = 4$. (Same straight line but the first pivot becomes 4.) The correct multiplier is now $\ell = \frac{3}{4}$. *To find the multiplier, divide the coefficient "3" to be eliminated by the pivot "4"*:

$$\begin{array}{lll} 4x - 8y = 4 & \textbf{Multiply equation 1 by } \frac{3}{4} & 4x - 8y = 4 \\ 3x + 2y = 11 & \textbf{Subtract from equation 2} & \quad\;\; 8y = 8. \end{array}$$

The final system is triangular and the last equation still gives $y = 1$. Back substitution produces $4x - 8 = 4$ and $4x = 12$ and $x = 3$. We changed the numbers but not the lines or the solution. ***Divide by the pivot to find that multiplier*** $\ell = \frac{3}{4}$:

Pivot	=	***first nonzero in the row that does the elimination***
Multiplier	=	***(entry to eliminate) divided by (pivot)*** $= \frac{3}{4}$.

The new second equation starts with the second pivot, which is 8. We would use it to eliminate y from the third equation if there were one. *To solve n equations we want n pivots.* ***The pivots are on the diagonal of the triangle after elimination.***

You could have solved those equations for x and y without reading this book. It is an extremely humble problem, but we stay with it a little longer. Even for a 2 by 2 system, elimination might break down. By understanding the possible breakdown (when we can't find a full set of pivots), you will understand the whole process of elimination.

Breakdown of Elimination

Normally, elimination produces the pivots that take us to the solution. But failure is possible. At some point, the method might ask us to *divide by zero*. We can't do it. The process has to stop. There might be a way to adjust and continue—or failure may be unavoidable.

Example 1 fails with ***no solution to*** $0y = 8$. Example 2 fails with ***too many solutions to*** $0y = 0$. Example 3 succeeds by exchanging the equations.

Example 1 ***Permanent failure with no solution.*** Elimination makes this clear:

$$\begin{array}{lll} x - 2y = 1 & \text{Subtract 3 times} & x - 2y = 1 \\ 3x - 6y = 11 & \text{eqn. 1 from eqn. 2} & \quad\;\; \mathbf{0}y = 8. \end{array}$$

There is *no* solution to $0y = 8$. Normally we divide the right side 8 by the second pivot, but *this system has no second pivot.* ***(Zero is never allowed as a pivot!)*** The row and column pictures in Figure 2.6 show why failure was unavoidable. If there is no solution, elimination will discover that fact by reaching an equation like $0y = 8$.

The row picture of failure shows parallel lines—which never meet. A solution must lie on both lines. With no meeting point, the equations have no solution.

The column picture shows the two columns $(1, 3)$ and $(-2, -6)$ in the same direction. *All combinations of the columns lie along a line.* But the column from the right side is in a different direction $(1, 11)$. No combination of the columns can produce this right side—therefore no solution.

When we change the right side to $(1, 3)$, failure shows as a whole line of solution points. Instead of no solution, next comes Example 2 with infinitely many.

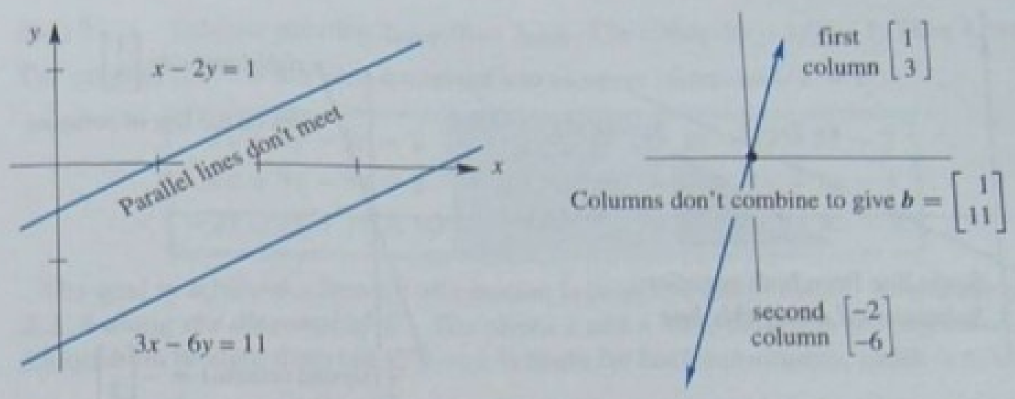

Figure 2.6: Row picture and column picture for Example 1: ***no solution***.

Example 2 ***Failure with infinitely many solutions. Change b = (1, 11) to (1, 3).***

$x - 2y = 1$	Subtract 3 times	$x - 2y = 1$	Still only
$3x - 6y = 3$	eqn. 1 from eqn. 2	$\mathbf{0}y = \mathbf{0}.$	**one pivot.**

Every y satisfies $0y = 0$. There is really only one equation $x - 2y = 1$. The unknown y is ***"free"***. After y is freely chosen, x is determined as $x = 1 + 2y$.

In the row picture, the parallel lines have become the same line. Every point on that line satisfies both equations. We have a whole line of solutions in Figure 2.7.

In the column picture, $\boldsymbol{b} = (1, 3)$ is now the same as column 1. So we can choose $x = 1$ and $y = 0$. We can also choose $x = 0$ and $y = -\frac{1}{2}$; column 2 times $-\frac{1}{2}$ equals $\boldsymbol{b}$. Every (x, y) that solves the row problem also solves the column problem.

Failure For n equations we do not get n pivots

Elimination leads to an equation $\mathbf{0} \neq \mathbf{0}$ (no solution) or $\mathbf{0} = \mathbf{0}$ (many solutions)

Success comes with n pivots. But we may have to exchange the n equations.

Elimination can go wrong in a third way—but this time it can be fixed. *Suppose the first pivot position contains zero.* We refuse to allow zero as a pivot. When the first equation has no term involving x, we can exchange it with an equation below:

Example 3 ***Temporary failure (zero in pivot). A row exchange produces two pivots:***

Permutation	$0x + 2y = 4$	Exchange the	$3x - 2y = 5$
	$3x - 2y = 5$	two equations	$2y = 4.$

The new system is already triangular. This small example is ready for back substitution. The last equation gives $y = 2$, and then the first equation gives $x = 3$. The row picture is normal (two intersecting lines). The column picture is also normal (column vectors not in the same direction). The pivots 3 and 2 are normal—but a ***row exchange*** was required.

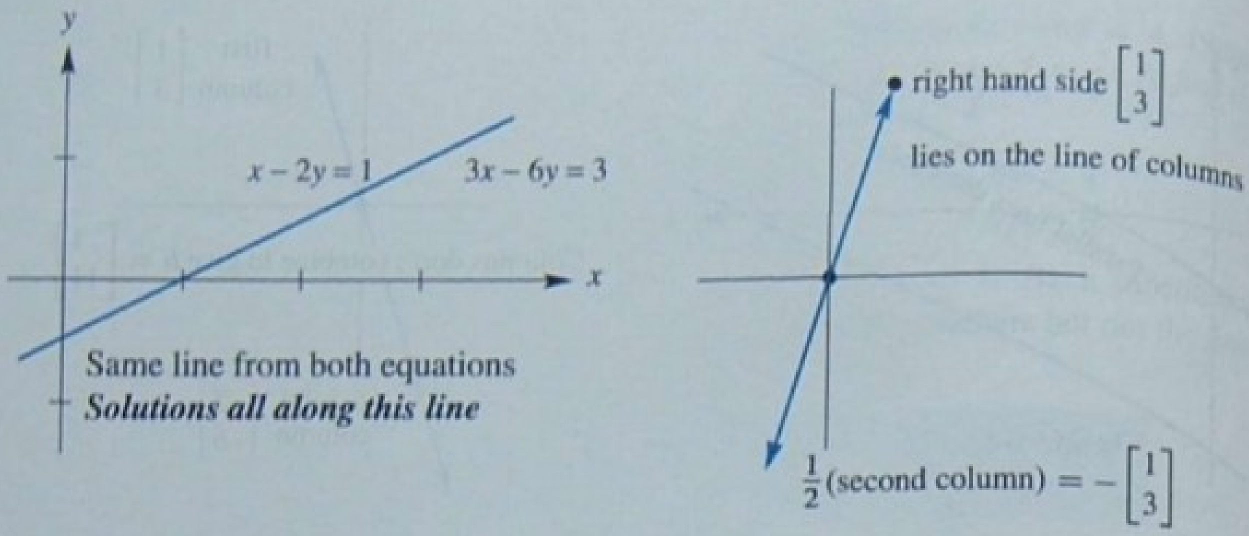

Figure 2.7: Row and column pictures for Example 2: *infinitely many solutions.*

Examples 1 and 2 are ***singular***—there is no second pivot. Example 3 is ***nonsingular***—there is a full set of pivots and exactly one solution. Singular equations have no solution or infinitely many solutions. Pivots must be nonzero because we have to divide by them.

Three Equations in Three Unknowns

To understand Gaussian elimination, you have to go beyond 2 by 2 systems. Three by three is enough to see the pattern. For now the matrices are square—an equal number of rows and columns. Here is a 3 by 3 system, specially constructed so that all steps lead to whole numbers and not fractions:

$$\begin{aligned} \mathbf{2}x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 10 \end{aligned} \tag{1}$$

What are the steps? The first pivot is the boldface **2** (upper left). Below that pivot we want to eliminate the 4. *The first multiplier is the ratio* $4/2 = 2$. Multiply the pivot equation by $\ell_{21} = 2$ and subtract. Subtraction removes the $4x$ from the second equation:

Step 1 Subtract 2 times equation 1 from equation 2. This leaves $y + z = 4$.

We also eliminate $-2x$ from equation 3—still using the first pivot. The quick way is to add equation 1 to equation 3. Then $2x$ cancels $-2x$. We do exactly that, but the rule in this book is to *subtract rather than add*. The systematic pattern has multiplier $\ell_{31} = -2/2 = -1$. Subtracting -1 times an equation is the same as adding:

Step 2 Subtract -1 times equation 1 from equation 3. This leaves $y + 5z = 12$.

The two new equations involve only y and z. The second pivot (in boldface) is 1:

x is eliminated
$$\begin{aligned} \mathbf{1}y + 1z &= 4 \\ 1y + 5z &= 12 \end{aligned}$$

We have reached a 2 by 2 system. The final step eliminates y to make it 1 by 1:

Step 3 Subtract equation 2_{new} from 3_{new}. The multiplier is $1/1 = 1$. Then $4z = 8$.

The original $A\boldsymbol{x} = \boldsymbol{b}$ has been converted into an upper triangular $U\boldsymbol{x} = \boldsymbol{c}$:

$$\begin{array}{r} 2x + 4y - 2z = 2 \\ 4x + 9y - 3z = 8 \\ -2x - 3y + 7z = 10 \end{array} \qquad \begin{array}{c} A\boldsymbol{x} = \boldsymbol{b} \\ \text{has become} \\ U\boldsymbol{x} = \boldsymbol{c} \end{array} \qquad \begin{array}{r} 2x + 4y - 2z = 2 \\ 1y + 1z = 4 \\ 4z = 8. \end{array} \tag{2}$$

The goal is achieved—forward elimination is complete from A to U. ***Notice the pivots* 2, 1, 4 *along the diagonal of*** U. The pivots 1 and 4 were hidden in the original system. Elimination brought them out. $U\boldsymbol{x} = \boldsymbol{c}$ is ready for ***back substitution***, which is quick:

$$(4z = 8 \quad \text{gives} \quad z = 2) \qquad (y + z = 4 \quad \text{gives} \quad y = 2) \qquad (\text{equation 1 gives} \quad x = -1)$$

The solution is $(x, y, z) = (-1, 2, 2)$. The row picture has three planes from three equations. All the planes go through this solution. The original planes are sloping, but the last plane $4z = 8$ after elimination is horizontal.

The column picture shows a combination $A\boldsymbol{x}$ of column vectors producing the right side $\boldsymbol{b}$. The coefficients in that combination are $-1, 2, 2$ (the solution):

$$A\boldsymbol{x} = (-1)\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 9 \\ -3 \end{bmatrix} + 2\begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix} \quad \text{equals} \quad \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} = \boldsymbol{b}. \tag{3}$$

The numbers x, y, z multiply columns $1, 2, 3$ in $A\boldsymbol{x} = \boldsymbol{b}$ and also in the triangular $U\boldsymbol{x} = \boldsymbol{c}$.

For a 4 by 4 problem, or an n by n problem, elimination proceeds the same way. Here is the whole idea, column by column from A to U, when elimination succeeds.

Column 1. ***Use the first equation to create zeros below the first pivot.***

Column 2. ***Use the new equation 2 to create zeros below the second pivot.***

Columns 3 to *n*. ***Keep going to find all*** n ***pivots and the triangular*** U.

$$\text{After column 2 we have} \quad \begin{bmatrix} \boldsymbol{x} & x & x & x \\ 0 & \boldsymbol{x} & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix}. \quad \text{We want} \quad \begin{bmatrix} \boldsymbol{x} & x & x & x \\ & \boldsymbol{x} & x & x \\ & & \boldsymbol{x} & x \\ & & & \boldsymbol{x} \end{bmatrix}. \tag{4}$$

The result of forward elimination is an upper triangular system. It is nonsingular if there is a full set of n pivots (never zero!). *Question*: Which x on the left could be changed to boldface $\boldsymbol{x}$ because the pivot is known? Here is a final example to show the original $A\boldsymbol{x} = \boldsymbol{b}$, the triangular system $U\boldsymbol{x} = \boldsymbol{c}$, and the solution (x, y, z) from back substitution:

$$\begin{array}{rl} x + y + z = 6 & \\ x + 2y + 2z = 9 & \textbf{Forward} \\ x + 2y + 3z = 10 & \textbf{Forward} \end{array} \qquad \begin{array}{r} x + y + z = 6 \\ y + z = 3 \\ z = 1 \end{array} \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad \begin{array}{l} \textbf{Back} \\ \textbf{Back} \\ \end{array}$$

All multipliers are 1. All pivots are 1. All planes meet at the solution $(3, 2, 1)$. The columns of A combine with $3, 2, 1$ to give $\boldsymbol{b} = (6, 9, 10)$. The triangle shows $U\boldsymbol{x} = \boldsymbol{c} = (6, 3, 1)$.

■ REVIEW OF THE KEY IDEAS ■

1. A linear system ($Ax = b$) becomes upper triangular ($Ux = c$) after elimination.
2. We subtract ℓ_{ij} times equation j from equation i, to make the (i, j) entry zero.
3. The multiplier is $\ell_{ij} = \dfrac{\text{entry to eliminate in row } i}{\text{pivot in row } j}$. Pivots can not be zero!
4. A zero in the pivot position can be repaired if there is a nonzero below it.
5. The upper triangular system is solved by back substitution (starting at the bottom).
6. When breakdown is permanent, the system has no solution or infinitely many.

■ WORKED EXAMPLES ■

2.2 A When elimination is applied to this matrix A, what are the first and second pivots? What is the multiplier ℓ_{21} in the first step (ℓ_{21} times row 1 is *subtracted* from row 2)?

A has a ***first difference*** in row 1 and a ***second difference*** $-1, 2, -1$ in row 2.

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix} \longrightarrow U = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

What entry in the 2, 2 position (instead of 2) would force an exchange of rows 2 and 3? Why is the lower left multiplier $\ell_{31} = 0$, subtracting zero times row 1 from row 3?
If you change the corner entry from $a_{33} = 2$ to $a_{33} = 1$, why does elimination fail?

Solution The first pivot is 1. The multiplier ℓ_{21} is $-1/1 = -1$. When -1 times row 1 is subtracted (so row 1 is added to row 2), the second pivot is revealed as 1.

If we reduce the middle entry "2" to "1", that would force a row exchange. (Zero will appear in the second pivot position.) The multiplier ℓ_{31} is zero because $a_{31} = 0$. A zero at the start of a row needs no elimination. This A is a "*band matrix*".

The last pivot is 1. So if the original corner entry a_{33} is reduced by 1 (to $a_{33} = 1$), elimination would produce 0. **No third pivot, elimination fails.**

2.2 B Suppose A is already a ***triangular matrix*** (upper triangular or lower triangular). *Where do you see its pivots?* When does $Ax = b$ have exactly one solution for every b?

Solution The pivots of a triangular matrix are already set along the main diagonal. *Elimination succeeds when all those numbers are nonzero.* Use ***back*** substitution when A is upper triangular, go ***forward*** when A is lower triangular.

2.2 C Use elimination to reach upper triangular matrices U. Solve by back substitution or explain why this is impossible. What are the pivots (never zero)? Exchange equations when necessary. The only difference is the $-x$ in the last equation.

$$\begin{array}{rll} \textbf{Success} & x+y+z=7 & x+y+z=7 \\ \textbf{then} & x+y-z=5 & x+y-z=5 \\ \textbf{Failure} & x-y+z=3 & -x-y+z=3 \end{array}$$

Solution For the first system, subtract equation 1 from equations 2 and 3 (the multipliers are $\ell_{21}=1$ and $\ell_{31}=1$). The 2, 2 entry becomes zero, so exchange equations:

$$\textbf{Success} \quad \begin{aligned} x+y+z &= 7 \\ \mathbf{0}y-2z &= -2 \\ -2y+0z &= -4 \end{aligned} \quad \text{exchanges into} \quad \begin{aligned} x+y+z &= 7 \\ \mathbf{-2}y+0z &= -4 \\ -2z &= -2 \end{aligned}$$

Then back substitution gives $z=1$ and $y=2$ and $x=4$. The pivots are $1, -2, -2$.

For the second system, subtract equation 1 from equation 2 as before. Add equation 1 to equation 3. This leaves zero in the 2, 2 entry *and also below*:

$$\textbf{Failure} \quad \begin{aligned} x+y+z &= 7 \\ \mathbf{0}y-2z &= -2 \\ \mathbf{0}y+2z &= 10 \end{aligned} \quad \begin{array}{l} \text{There is } \textit{no pivot} \text{ in column 2 (it was } - \text{ column 1)} \\ \text{A further elimination step gives } 0z=8 \\ \text{The three planes don't meet} \end{array}$$

Plane 1 meets plane 2 in a line. Plane 1 meets plane 3 in a parallel line. *No solution.*

If we change the "3" in the original third equation to "-5" then elimination would lead to $0=0$. There are infinitely many solutions! *The three planes now meet along a whole line.*

Changing 3 to -5 moved the third plane to meet the other two. The second equation gives $z=1$. Then the first equation leaves $x+y=6$. **No pivot in column 2 makes** y **free** (it can have any value). Then $x=6-y$.

Problem Set 2.2

Problems 1–10 are about elimination on 2 by 2 systems.

1 What multiple ℓ_{21} of equation 1 should be subtracted from equation 2?

$$\begin{aligned} 2x+3y &= 1 \\ 10x+9y &= 11. \end{aligned}$$

After this elimination step, write down the upper triangular system and circle the two pivots. The numbers 1 and 11 have no influence on those pivots.

2 Solve the triangular system of Problem 1 by back substitution, y before x. Verify that x times $(2,10)$ plus y times $(3,9)$ equals $(1,11)$. If the right side changes to $(4,44)$, what is the new solution?

Problems 11–20 study elimination on 3 by 3 systems (and possible failure).

11 (Recommended) A system of linear equations can't have exactly two solutions. *Why*?

(a) If (x, y, z) and (X, Y, Z) are two solutions, what is another solution?

(b) If 25 planes meet at two points, where else do they meet?

12 Reduce this system to upper triangular form by two row operations:

$$\begin{aligned} 2x + 3y + z &= 8 \\ 4x + 7y + 5z &= 20 \\ -2y + 2z &= 0. \end{aligned}$$

Circle the pivots. Solve by back substitution for z, y, x.

13 Apply elimination (circle the pivots) and back substitution to solve

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

List the three row operations: Subtract ____ times row ____ from row ____.

14 Which number d forces a row exchange, and what is the triangular system (not singular) for that d? Which d makes this system singular (no third pivot)?

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

15 Which number b leads later to a row exchange? Which b leads to a missing pivot? In that singular case find a nonzero solution x, y, z.

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

16 (a) Construct a 3 by 3 system that needs two row exchanges to reach a triangular form and a solution.

(b) Construct a 3 by 3 system that needs a row exchange to keep going, but breaks down later.

17 If rows 1 and 2 are the same, how far can you get with elimination (allowing row exchange)? If columns 1 and 2 are the same, which pivot is missing?

Equal rows	$2x - y + z = 0$	$2x + 2y + z = 0$	**Equal columns**
	$2x - y + z = 0$	$4x + 4y + z = 0$	
	$4x + y + z = 2$	$6x + 6y + z = 2.$	

2.3 Elimination Using Matrices

We now combine two ideas—elimination and matrices. The goal is to express all the steps of elimination (and the final result) in the clearest possible way. In a 3 by 3 example, elimination could be described in words. For larger systems, a long list of steps would be hopeless. You will see how to subtract a multiple of row j from row i—*using a matrix* E.

The 3 by 3 example in the previous section has the beautifully short form $A\boldsymbol{x} = \boldsymbol{b}$:

$$\begin{aligned} 2x_1 + 4x_2 - 2x_3 &= 2 \\ 4x_1 + 9x_2 - 3x_3 &= 8 \\ -2x_1 - 3x_2 + 7x_3 &= 10 \end{aligned} \quad \text{is the same as} \quad \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix}. \quad (1)$$

The nine numbers on the left go into the matrix A. That matrix not only sits beside $\boldsymbol{x}$, it *multiplies* $\boldsymbol{x}$. The rule for "A times $\boldsymbol{x}$" is exactly chosen to yield the three equations.

Review of A times x. A matrix times a vector gives a vector. The matrix is square when the number of equations (three) matches the number of unknowns (three). Our matrix is 3 by 3. A general square matrix is n by n. Then the vector $\boldsymbol{x}$ is in n-dimensional space.

$$\textbf{\textit{The unknown in }}\mathbf{R}^3\textbf{\textit{ is}} \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \textbf{\textit{and the solution is}} \quad \boldsymbol{x} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Key point: $A\boldsymbol{x} = \boldsymbol{b}$ represents the row form and also the column form of the equations.

$$\textbf{Column form} \qquad A\boldsymbol{x} = (-1)\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 9 \\ -3 \end{bmatrix} + 2\begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} = \boldsymbol{b}.$$

This rule for $A\boldsymbol{x}$ is used so often that we express it once more for emphasis.

Ax is a combination of the columns of A. Components of $\boldsymbol{x}$ multiply those columns:

$$A\boldsymbol{x} = x_1 \text{ times (column 1)} + \cdots + x_n \text{ times (column } n).$$

When we compute the components of $A\boldsymbol{x}$, we use the row form of matrix multiplication. The ith component is a dot product with row i of A, which is $[a_{i1} \ \ a_{i2} \ \ \ldots \ \ a_{in}]$. The short formula for that dot product with $\boldsymbol{x}$ uses "sigma notation".

Components of $A\boldsymbol{x}$ are dot products with rows of A.

The ith component of $A\boldsymbol{x}$ is $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$. This is $\displaystyle\sum_{j=1}^{n} a_{ij}x_j$.

The sigma symbol $\sum$ is an instruction to add.[1] Start with $j = 1$ and stop with $j = n$. Start the sum with $a_{i1}x_1$ and stop with $a_{in}x_n$. That produces (row i) $\cdot\, \boldsymbol{x}$.

[1] Einstein shortened this even more by omitting the $\sum$. The repeated j in $a_{ij}x_j$ automatically meant addition. He also wrote the sum as $a_i^j x_j$. Not being Einstein, we include the $\sum$.

One point to repeat about matrix notation: The entry in row 1, column 1 (the top left corner) is a_{11}. The entry in row 1, column 3 is a_{13}. The entry in row 3, column 1 is a_{31}. (Row number comes before column number.) The word "entry" for a matrix corresponds to "component" for a vector. General rule: $a_{ij} = A(i, j)$ ***is in row** i, **column** j.*

Example 1 This matrix has $a_{ij} = 2i + j$. Then $a_{11} = 3$. Also $a_{12} = 4$ and $a_{21} = 5$. Here is Ax with numbers and letters:

$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \cdot 2 + 4 \cdot 1 \\ 5 \cdot 2 + 6 \cdot 1 \end{bmatrix} \qquad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}.$$

The first component of Ax is $6 + 4 = 10$. ***A row times a column gives a dot product.***

The Matrix Form of One Elimination Step

$Ax = b$ is a convenient form for the original equation. What about the elimination steps? The first step in this example subtracts 2 times the first equation from the second equation. On the right side, 2 times the first component of b is subtracted from the second component:

$$\textbf{First step} \qquad b = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} \quad \text{changes to} \quad b_{\text{new}} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix}.$$

We want to do that subtraction with a matrix! The same result $b_{\text{new}} = Eb$ is achieved when we multiply an "elimination matrix" E times b. It subtracts $2b_1$ from b_2:

$$\textit{The elimination matrix is} \qquad E = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Multiplication by E subtracts 2 times row 1 from row 2. Rows 1 and 3 stay the same:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 10 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 \end{bmatrix}$$

The first and third rows of E are rows from the identity matrix I. The new second component is the number 4 that appeared after the elimination step. This is $b_2 - 2b_1$.

It is easy to describe the "elementary matrices" or "elimination matrices" like this E. Start with the identity matrix I. *Change one of its zeros to the multiplier* $-\ell$:

The ***identity matrix*** has 1's on the diagonal and otherwise 0's. Then $Ib = b$ for all b. The ***elementary matrix or elimination matrix*** E_{ij} that subtracts a multiple ℓ of row j from row i has the extra nonzero entry $-\ell$ in the i, j position (still diagonal 1's).

Example 2 The matrix E_{31} has $-\ell$ in the 3, 1 position:

$$\textbf{Identity} \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \textbf{Elimination} \quad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell & 0 & 1 \end{bmatrix}.$$

When you multiply I times $\boldsymbol{b}$, you get $\boldsymbol{b}$. But E_{31} subtracts ℓ times the first component from the third component. With $\ell = 4$ this example gives $9 - 4 = 5$:

$$I\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} \quad \text{and} \quad E\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \\ \mathbf{9} \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ \mathbf{5} \end{bmatrix}.$$

What about the left side of $A\boldsymbol{x} = \boldsymbol{b}$? Both sides are multiplied by E_{31}. ***The purpose of E_{31} is to produce a zero in the (3, 1) position of the matrix.***

The notation fits this purpose. Start with A. Apply E's to produce zeros below the pivots (the first E is E_{21}). End with a triangular U. We now look in detail at those steps.

First a small point. The vector $\boldsymbol{x}$ stays the same. The solution is not changed by elimination. (That may be more than a small point.) It is the coefficient matrix that is changed. When we start with $A\boldsymbol{x} = \boldsymbol{b}$ and multiply by E, the result is $EA\boldsymbol{x} = E\boldsymbol{b}$. The new matrix EA is the result of *multiplying E times A*.

Confession The *elimination matrices* E_{ij} are great examples, but you won't see them later. They show how a matrix acts on rows. By taking several elimination steps, we will see how to *multiply matrices* (and the order of the E's becomes important). ***Products and inverses*** are especially clear for E's. It is those two ideas that the book will now use.

Matrix Multiplication

The big question is: ***How do we multiply two matrices?*** When the first matrix is E, we already know what to expect for EA. This particular E subtracts 2 times row 1 from row 2 of this matrix A and any matrix. The multiplier is $\ell = 2$:

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 \\ \mathbf{0} & \mathbf{1} & \mathbf{1} \\ -2 & -3 & 7 \end{bmatrix} \quad (\textit{with the zero}). \tag{2}$$

This step does not change rows 1 and 3 of A. Those rows are unchanged in EA—only row 2 is different. *Twice the first row has been subtracted from the second row.* Matrix multiplication agrees with elimination—and the new system of equations is $EA\boldsymbol{x} = E\boldsymbol{b}$.

$EA\boldsymbol{x}$ is simple but it involves a subtle idea. Start with $A\boldsymbol{x} = \boldsymbol{b}$. Multiplying both sides by E gives $E(A\boldsymbol{x}) = E\boldsymbol{b}$. With matrix multiplication, this is also $(EA)\boldsymbol{x} = E\boldsymbol{b}$. **The first was E times $A\boldsymbol{x}$, the second is EA times $\boldsymbol{x}$. They are the same.** Parentheses are not needed. We just write $EA\boldsymbol{x}$.

That rule extends to a matrix C with several column vectors like $C = [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \boldsymbol{c}_3]$. When multiplying EAC, you can do AC first or EA first. This is the point of an "associative law" like $3 \times (4 \times 5) = (3 \times 4) \times 5$. Multiply 3 times 20, or multiply 12 times 5. Both answers are 60. That law seems so clear that it is hard to imagine it could be false.

The "commutative law" $3 \times 4 = 4 \times 3$ looks even more obvious. But EA is usually different from AE. When E multiplies on the right, it acts on the ***columns*** of A.

Associative law is true	$A(BC) = (AB)C$
Commutative law is false	Often $AB \neq BA$

There is another requirement on matrix multiplication. Suppose B has only one column (this column is $\boldsymbol{b}$). The matrix-matrix law for EB should agree with the matrix-vector law for $E\boldsymbol{b}$. Even more, we should be able to *multiply matrices EB a column at a time*:

If B has several columns $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$, ***then the columns of EB are*** $E\boldsymbol{b}_1, E\boldsymbol{b}_2, E\boldsymbol{b}_3$.

Matrix multiplication $\quad AB = A\,[\boldsymbol{b}_1\ \boldsymbol{b}_2\ \boldsymbol{b}_3] = [A\boldsymbol{b}_1\ A\boldsymbol{b}_2\ A\boldsymbol{b}_3].$ (3)

This holds true for the matrix multiplication in (2). If you multiply column 3 of A by E, you correctly get column 3 of EA:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix} \qquad E(\text{column } j \text{ of } A) = \text{column } j \text{ of } EA.$$

This requirement deals with columns, while elimination is applied to rows. **The next section describes each entry of every product** AB. The beauty of matrix multiplication is that all three approaches (*rows, columns, whole matrices*) come out right.

The Matrix P_{ij} for a Row Exchange

To subtract row j from row i we use E_{ij}. To exchange or "permute" those rows we use another matrix P_{ij} (a **permutation matrix**). A row exchange is needed when zero is in the pivot position. Lower down, that pivot column may contain a nonzero. By exchanging the two rows, we have a pivot and elimination goes forward.

What matrix P_{23} exchanges row 2 with row 3? We can find it by exchanging rows of the identity matrix I:

$$\textbf{Permutation matrix} \qquad P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

This is a ***row exchange matrix***. Multiplying by P_{23} exchanges components 2 and 3 of any column vector. Therefore it also exchanges rows 2 and 3 of any matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{3} \\ \mathbf{5} \end{bmatrix} = \begin{bmatrix} 1 \\ \mathbf{5} \\ \mathbf{3} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{3} \\ 0 & 6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 1 \\ 0 & 6 & 5 \\ \mathbf{0} & \mathbf{0} & \mathbf{3} \end{bmatrix}.$$

On the right, P_{23} is doing what it was created for. With zero in the second pivot position and "6" below it, the exchange puts 6 into the pivot.

Matrices *act*. They don't just sit there. We will soon meet other permutation matrices, which can change the order of several rows. Rows 1, 2, 3 can be moved to 3, 1, 2. Our P_{23} is one particular permutation matrix—it exchanges rows 2 and 3.

Row Exchange Matrix P_{ij} is the identity matrix with rows i and j reversed. When this **"permutation matrix"** P_{ij} multiplies a matrix, it exchanges rows i and j.

$$\textit{To exchange equations 1 and 3 multiply by} \quad P_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Usually row exchanges are not required. The odds are good that elimination uses only the E_{ij}. But the P_{ij} are ready if needed, to move a pivot up to the diagonal.

The Augmented Matrix

This book eventually goes far beyond elimination. Matrices have all kinds of practical applications, in which they are multiplied. Our best starting point was a square E times a square A, because we met this in elimination—and we know what answer to expect for EA. The next step is to allow a *rectangular matrix*. It still comes from our original equations, but now it includes the right side $\boldsymbol{b}$.

Key idea: Elimination does the same row operations to A and to $\boldsymbol{b}$. ***We can include b as an extra column and follow it through elimination.*** The matrix A is enlarged or "augmented" by the extra column $\boldsymbol{b}$:

$$\textbf{\textit{Augmented matrix}} \qquad \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & \mathbf{2} \\ 4 & 9 & -3 & \mathbf{8} \\ -2 & -3 & 7 & \mathbf{10} \end{bmatrix}.$$

Elimination acts on whole rows of this matrix. The left side and right side are both multiplied by E, to subtract 2 times equation 1 from equation 2. With $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$ those steps happen together:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 & \mathbf{2} \\ 4 & 9 & -3 & \mathbf{8} \\ -2 & -3 & 7 & \mathbf{10} \end{bmatrix} = \begin{bmatrix} 2 & 4 & -2 & \mathbf{2} \\ 0 & 1 & 1 & \mathbf{4} \\ -2 & -3 & 7 & \mathbf{10} \end{bmatrix}.$$

The new second row contains 0, 1, 1, 4. The new second equation is $x_2 + x_3 = 4$. Matrix multiplication works by rows and at the same time by columns:

ROWS Each row of E acts on $[\,A \;\; \boldsymbol{b}\,]$ to give a row of $[\,EA \;\; E\boldsymbol{b}\,]$.

COLUMNS E acts on each column of $[\,A \;\; \boldsymbol{b}\,]$ to give a column of $[\,EA \;\; E\boldsymbol{b}\,]$.

Notice again that word "acts." This is essential. Matrices do something! The matrix A acts on $\boldsymbol{x}$ to produce $\boldsymbol{b}$. The matrix E operates on A to give EA. The whole process of elimination is a sequence of row operations, alias matrix multiplications. A goes to $E_{21}A$ which goes to $E_{31}E_{21}A$. Finally $E_{32}E_{31}E_{21}A$ is a triangular matrix.

The right side is included in the augmented matrix. The end result is a triangular system of equations. We stop for exercises on multiplication by E, before writing down the rules for all matrix multiplications (including block multiplication).

■ REVIEW OF THE KEY IDEAS ■

1. $Ax = x_1$ times column $1 + \cdots + x_n$ times column n. And $(Ax)_i = \sum_{j=1}^{n} a_{ij}x_j$.
2. Identity matrix $= I$, elimination matrix $= E_{ij}$ using ℓ_{ij}, exchange matrix $= P_{ij}$.
3. Multiplying $Ax = b$ by E_{21} subtracts a multiple ℓ_{21} of equation 1 from equation 2. The number $-\ell_{21}$ is the $(2,1)$ entry of the elimination matrix E_{21}.
4. For the augmented matrix $\begin{bmatrix} A & b \end{bmatrix}$, that elimination step gives $\begin{bmatrix} E_{21}A & E_{21}b \end{bmatrix}$.
5. When A multiplies any matrix B, it multiplies each column of B separately.

■ WORKED EXAMPLES ■

2.3 A What 3 by 3 matrix E_{21} subtracts 4 times row 1 from row 2? What matrix P_{32} exchanges row 2 and row 3? If you multiply A on the *right* instead of the left, describe the results AE_{21} and AP_{32}.

Solution By doing those operations on the identity matrix I, we find

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad P_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Multiplying by E_{21} on the right side will subtract 4 times **column 2** from **column 1**. Multiplying by P_{32} on the right will exchange **columns 2** and **3**.

2.3 B Write down the augmented matrix $[A \quad b]$ with an extra column:

$$\begin{aligned} x + 2y + 2z &= 1 \\ 4x + 8y + 9z &= 3 \\ 3y + 2z &= 1 \end{aligned}$$

Apply E_{21} and then P_{32} to reach a triangular system. Solve by back substitution. What combined matrix $P_{32}E_{21}$ will do both steps at once?

Solution E_{21} removes the 4 in column 1. But zero appears in column 2:

$$[A \;\; b] = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 4 & 8 & 9 & 3 \\ 0 & 3 & 2 & 1 \end{bmatrix} \quad \text{and} \quad E_{21}[A \;\; b] = \begin{bmatrix} 1 & 2 & 2 & 1 \\ \mathbf{0} & \mathbf{0} & 1 & -1 \\ 0 & 3 & 2 & 1 \end{bmatrix}$$

Now P_{32} exchanges rows 2 and 3. Back substitution produces z then y and x.

$$P_{32}\, E_{21}[A \;\; b] = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

For the matrix $P_{32}\, E_{21}$ that does both steps at once, *apply* P_{32} *to* E_{21}.

One matrix
Both steps

$$P_{32}\, E_{21} = \text{ exchange the rows of } E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -4 & 1 & 0 \end{bmatrix}.$$

2.3 C Multiply these matrices in two ways. First, rows of A times columns of B. Second, ***columns of A times rows of B***. That unusual way produces two matrices that add to AB. How many separate ordinary multiplications are needed?

Both ways

$$AB = \begin{bmatrix} 3 & 4 \\ 1 & 5 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{10} & \mathbf{16} \\ \mathbf{7} & \mathbf{9} \\ \mathbf{4} & \mathbf{8} \end{bmatrix}$$

Solution Rows of A times columns of B are dot products of vectors:

$$(\text{row } 1) \cdot (\text{column } 1) = [3 \;\; 4] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{10} \qquad \text{is the } (1,1) \text{ entry of } AB$$

$$(\text{row } 2) \cdot (\text{column } 1) = [1 \;\; 5] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{7} \qquad \text{is the } (2,1) \text{ entry of } AB$$

We need 6 dot products, 2 multiplications each, 12 in all $(3 \cdot 2 \cdot 2)$. The same AB comes from *columns of A times rows of B*. A column times a row is a matrix.

$$AB = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} [2 \;\; 4] + \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} [1 \;\; 1] = \begin{bmatrix} \mathbf{6} & \mathbf{12} \\ \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \end{bmatrix} + \begin{bmatrix} \mathbf{4} & \mathbf{4} \\ \mathbf{5} & \mathbf{5} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Problem Set 2.3

Problems 1–15 are about elimination matrices.

1 Write down the 3 by 3 matrices that produce these elimination steps:

(a) E_{21} subtracts 5 times row 1 from row 2.

(b) E_{32} subtracts -7 times row 2 from row 3.

(c) P exchanges rows 1 and 2, then rows 2 and 3.

2 In Problem 1, applying E_{21} and then E_{32} to $\boldsymbol{b} = (1, 0, 0)$ gives $E_{32}E_{21}\boldsymbol{b} =$ ____. Applying E_{32} before E_{21} gives $E_{21}E_{32}\boldsymbol{b} =$ ____. When E_{32} comes first, row ____ feels no effect from row ____.

3 Which three matrices E_{21}, E_{31}, E_{32} put A into triangular form U?

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 4 & 6 & 1 \\ -2 & 2 & 0 \end{bmatrix} \quad \text{and} \quad E_{32}E_{31}E_{21}A = U.$$

Multiply those E's to get one matrix M that does elimination: $MA = U$.

4 Include $\boldsymbol{b} = (1, 0, 0)$ as a fourth column in Problem 3 to produce $[\,A \;\; \boldsymbol{b}\,]$. Carry out the elimination steps on this augmented matrix to solve $A\boldsymbol{x} = \boldsymbol{b}$.

5 Suppose $a_{33} = 7$ and the third pivot is 5. If you change a_{33} to 11, the third pivot is ____. If you change a_{33} to ____, there is no third pivot.

6 If every column of A is a multiple of $(1, 1, 1)$, then $A\boldsymbol{x}$ is always a multiple of $(1, 1, 1)$. Do a 3 by 3 example. How many pivots are produced by elimination?

7 Suppose E subtracts 7 times row 1 from row 3.

(a) To *invert* that step you should ____ 7 times row ____ to row ____.

(b) What "inverse matrix" E^{-1} takes that reverse step (so $E^{-1}E = I$)?

(c) If the reverse step is applied first (and then E) show that $EE^{-1} = I$.

8 The ***determinant*** of $M = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is $\det M = ad - bc$. Subtract ℓ times row 1 from row 2 to produce a new M^*. Show that $\det M^* = \det M$ for every ℓ. When $\ell = c/a$, *the product of pivots equals the determinant*: $(a)(d - \ell b)$ equals $ad - bc$.

9 (a) E_{21} subtracts row 1 from row 2 and then P_{23} exchanges rows 2 and 3. What matrix $M = P_{23}E_{21}$ does both steps at once?

(b) P_{23} exchanges rows 2 and 3 and then E_{31} subtracts row 1 from row 3. What matrix $M = E_{31}P_{23}$ does both steps at once? Explain why the M's are the same but the E's are different.

10 (a) What 3 by 3 matrix E_{13} will add row 3 to row 1?

(b) What matrix adds row 1 to row 3 and *at the same time* row 3 to row 1?

(c) What matrix adds row 1 to row 3 and *then* adds row 3 to row 1?

11 Create a matrix that has $a_{11} = a_{22} = a_{33} = 1$ but elimination produces two negative pivots without row exchanges. (The first pivot is 1.)

12 Multiply these matrices:

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

13 Explain these facts. If the third column of B is all zero, the third column of EB is all zero (for any E). If the third *row* of B is all zero, the third row of EB might *not* be zero.

14 This 4 by 4 matrix will need elimination matrices E_{21} and E_{32} and E_{43}. What are those matrices?

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

15 Write down the 3 by 3 matrix that has $a_{ij} = 2i - 3j$. This matrix has $a_{32} = 0$, but elimination still needs E_{32} to produce a zero in the 3, 2 position. Which previous step destroys the original zero and what is E_{32}?

Problems 16–23 are about creating and multiplying matrices.

16 Write these ancient problems in a 2 by 2 matrix form $Ax = b$ and solve them:

(a) X is twice as old as Y and their ages add to 33.

(b) $(x, y) = (2, 5)$ and $(3, 7)$ lie on the line $y = mx + c$. Find m and c.

17 The parabola $y = a + bx + cx^2$ goes through the points $(x, y) = (1, 4)$ and $(2, 8)$ and $(3, 14)$. Find and solve a matrix equation for the unknowns (a, b, c).

18 Multiply these matrices in the orders EF and FE:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \qquad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix}.$$

Also compute $E^2 = EE$ and $F^3 = FFF$. You can guess F^{100}.

2.4 Rules for Matrix Operations

I will start with basic facts. A matrix is a rectangular array of numbers or "entries". When A has m rows and n columns, it is an "m by n" matrix. Matrices can be added if their shapes are the same. They can be multiplied by any constant c. Here are examples of $A + B$ and $2A$, for 3 by 2 matrices:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 4 & 4 \\ 9 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 7 & 8 \\ 9 & 9 \end{bmatrix} \quad \text{and} \quad 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \\ 0 & 0 \end{bmatrix}.$$

Matrices are added exactly as vectors are—one entry at a time. We could even regard a column vector as a matrix with only one column (so $n = 1$). The matrix $-A$ comes from multiplication by $c = -1$ (reversing all the signs). Adding A to $-A$ leaves the *zero matrix*, with all entries zero. All this is only common sense.

The entry in row i and column j is called a_{ij} or $A(i, j)$. The n entries along the first row are $a_{11}, a_{12}, \ldots, a_{1n}$. The lower left entry in the matrix is a_{m1} and the lower right is a_{mn}. The row number i goes from 1 to m. The column number j goes from 1 to n.

Matrix addition is easy. The serious question is ***matrix multiplication.*** When can we multiply A times B, and what is the product AB? We cannot multiply when A and B are 3 by 2. They don't pass the following test:

To multiply AB: ***If A has n columns, B must have n rows.***

When A is 3 by 2, the matrix B can be 2 by 1 (a vector) or 2 by 2 (square) or 2 by 20. ***Every column of B is multiplied by A.*** I will begin matrix multiplication the *dot product way*, and then return to this *column way*: A times columns of B. The most important rule is that ***AB times C equals A times BC.*** A Challenge Problem will prove this.

Suppose A is m by n and B is n by p. We can multiply. The product AB is m by p.

$$(m \times n)(n \times p) = (m \times p) \qquad \begin{bmatrix} m \textbf{ rows} \\ n \textit{ columns} \end{bmatrix} \begin{bmatrix} n \textit{ rows} \\ p \textbf{ columns} \end{bmatrix} = \begin{bmatrix} m \textbf{ rows} \\ p \textbf{ columns} \end{bmatrix}.$$

A row times a column is an extreme case. Then 1 by n multiplies n by 1. The result is 1 by 1. That single number is the "dot product".

In every case AB is filled with dot products. For the top corner, the $(1, 1)$ entry of AB is (row 1 of A) $\cdot$ (column 1 of B). To multiply matrices, take the dot product of ***each row of A with each column of B.***

The entry in row i and column j of AB is (row i of A) $\cdot$ (column j of B) .

Figure 2.8 picks out the second row ($i = 2$) of a 4 by 5 matrix A. It picks out the third column ($j = 3$) of a 5 by 6 matrix B. Their dot product goes into row 2 and column 3 of AB. The matrix AB has *as many rows as A* (4 rows), and *as many columns as B*.

$$\begin{bmatrix} * \\ a_{i1}\ a_{i2}\ \cdots\ a_{i5} \\ * \\ * \end{bmatrix} \begin{bmatrix} * & * & b_{1j} & * & * & * \\ & & b_{2j} & & & \\ & & \vdots & & & \\ & & b_{5j} & & & \end{bmatrix} = \begin{bmatrix} & & * & & & \\ * & * & (AB)_{ij} & * & * & * \\ & & * & & & \\ & & * & & & \end{bmatrix}$$

A is 4 by 5 $\quad$ B is 5 by 6 $\quad$ AB is 4 by 6

Figure 2.8: Here $i = 2$ and $j = 3$. Then $(AB)_{23}$ is (row 2) $\cdot$ (column 3) $= \Sigma a_{2k} b_{k3}$.

Example 1 Square matrices can be multiplied if and only if they have the same size:

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

The first dot product is $1 \cdot 2 + 1 \cdot 3 = 5$. Three more dot products give 6, 1, and 0. Each dot product requires two multiplications—thus eight in all.

If A and B are n by n, so is AB. It contains n^2 dot products, row of A times column of B. Each dot product needs n multiplications, so ***the computation of AB uses n^3 separate multiplications***. For $n = 100$ we multiply a million times. For $n = 2$ we have $n^3 = 8$.

Mathematicians thought until recently that AB absolutely needed $2^3 = 8$ multiplications. Then somebody found a way to do it with 7 (and extra additions). By breaking n by n matrices into 2 by 2 blocks, this idea also reduced the count for large matrices. Instead of n^3 it went below $n^{2.8}$, and the exponent keeps falling.[1] The best at this moment is $n^{2.376}$. But the algorithm is so awkward that scientific computing is done the regular way: n^2 dot products in AB, and n multiplications for each one.

Example 2 Suppose A is a row vector (1 by 3) and B is a column vector (3 by 1). Then AB is 1 by 1 (only one entry, the dot product). On the other hand B times A (***a column times a row***) is a full 3 by 3 matrix. This multiplication is allowed!

Column times row
$(n \times 1)(1 \times n) = (n \times n)$

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}.$$

A row times a column is an "*inner*" product—that is another name for dot product. A column times a row is an "*outer*" product. These are extreme cases of matrix multiplication.

Rows and Columns of AB

In the big picture, A multiplies each column of B. The result is a column of AB. In that column, we are combining the columns of A. ***Each column of AB is a combination of***

[1] Maybe 2.376 will drop to 2. No other number looks special, but no change for 10 years.

the columns of A. That is the column picture of matrix multiplication:

$$\textbf{Matrix } A \textbf{ times column of } B \qquad A\begin{bmatrix} \boldsymbol{b}_1 \cdots \boldsymbol{b}_p \end{bmatrix} = \begin{bmatrix} A\boldsymbol{b}_1 \cdots A\boldsymbol{b}_p \end{bmatrix}.$$

The row picture is reversed. Each row of A multiplies the whole matrix B. The result is a row of AB. It is a combination of the rows of B:

$$\textbf{Row times matrix} \qquad \begin{bmatrix} \text{row } i \text{ of } A \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} \text{row } i \text{ of } AB \end{bmatrix}.$$

We see row operations in elimination (E times A). We see columns in A times $\boldsymbol{x}$. The "row-column picture" has the dot products of rows with columns. Believe it or not, ***there is also a column-row picture.*** Not everybody knows that columns $1, \ldots, n$ of A multiply rows $1, \ldots, n$ of B and add up to the same answer AB. Worked Example **2.3 C** had numbers for $n = 2$. ***Example 3 will show how to multiply AB using columns times rows.***

The Laws for Matrix Operations

May I put on record six laws that matrices do obey, while emphasizing an equation they don't obey? The matrices can be square or rectangular, and the laws involving $A + B$ are all simple and all obeyed. Here are three addition laws:

$$\begin{aligned} A + B &= B + A && \text{(commutative law)} \\ c(A + B) &= cA + cB && \text{(distributive law)} \\ A + (B + C) &= (A + B) + C && \text{(associative law).} \end{aligned}$$

Three more laws hold for multiplication, but $AB = BA$ is not one of them:

$$\begin{aligned} AB &\neq BA && \text{(the commutative "law" is } \textit{usually broken}) \\ C(A + B) &= CA + CB && \text{(distributive law from the left)} \\ (A + B)C &= AC + BC && \text{(distributive law from the right)} \\ A(BC) &= (AB)C && \text{(associative law for } ABC) \ (\textbf{parentheses not needed}). \end{aligned}$$

When A and B are not square, AB is a different size from BA. These matrices can't be equal—even if both multiplications are allowed. For square matrices, almost any example shows that AB is different from BA:

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{1} \end{bmatrix} \quad \text{but} \quad BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{bmatrix}.$$

It is true that $AI = IA$. All square matrices commute with I and also with cI. Only these matrices cI commute with all other matrices.

The law $A(B + C) = AB + AC$ is proved a column at a time. Start with $A(\boldsymbol{b} + \boldsymbol{c}) = A\boldsymbol{b} + A\boldsymbol{c}$ for the first column. That is the key to everything—***linearity***. Say no more.

The law $A(BC) = (AB)C$ means that you can multiply BC first or else AB first. The direct proof is sort of awkward (Problem 37) but this law is extremely useful. We highlighted it above; it is the key to the way we multiply matrices.

Look at the special case when $A = B = C =$ square matrix. Then (A *times* A^2) is equal to (A^2 *times* A). The product in either order is A^3. The matrix powers A^p follow the same rules as numbers:

$$A^p = AAA\cdots A \text{ (}p\text{ factors)} \qquad (A^p)(A^q) = A^{p+q} \qquad (A^p)^q = A^{pq}.$$

Those are the ordinary laws for exponents. A^3 times A^4 is A^7 (seven factors). A^3 to the fourth power is A^{12} (twelve A's). When p and q are zero or negative these rules still hold, provided A has a "-1 power"—which is the *inverse matrix* A^{-1}. Then $A^0 = I$ is the identity matrix (no factors).

For a number, a^{-1} is $1/a$. For a matrix, the inverse is written A^{-1}. (It is *never* I/A, except this is allowed in MATLAB.) Every number has an inverse except $a = 0$. To decide when A has an inverse is a central problem in linear algebra. Section 2.5 will start on the answer. This section is a Bill of Rights for matrices, to say when A and B can be multiplied and how.

Block Matrices and Block Multiplication

We have to say one more thing about matrices. They can be cut into ***blocks*** (which are smaller matrices). This often happens naturally. Here is a 4 by 6 matrix broken into blocks of size 2 by 2—in this example each block is just I:

4 by 6 matrix
2 by 2 blocks

$$A = \left[\begin{array}{cc|cc|cc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right] = \begin{bmatrix} I & I & I \\ I & I & I \end{bmatrix}.$$

If B is also 4 by 6 and the block sizes match, you can add $A + B$ *a block at a time*.

We have seen block matrices before. The right side vector $\boldsymbol{b}$ was placed next to A in the "augmented matrix". Then $[\,A \;\; \boldsymbol{b}\,]$ has two blocks of different sizes. Multiplying by an elimination matrix gave $[\,EA \;\; E\boldsymbol{b}\,]$. No problem to multiply blocks times blocks, when their shapes permit.

Block multiplication If the cuts between columns of A match the cuts between rows of B, then block multiplication of AB is allowed:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & \cdots \\ B_{21} & \cdots \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & \cdots \\ A_{21}B_{11} + A_{22}B_{21} & \cdots \end{bmatrix}. \qquad (1)$$

This equation is the same as if the blocks were numbers (which are 1 by 1 blocks). We are careful to keep A's in front of B's, because BA can be different.

Main point When matrices split into blocks, it is often simpler to see how they act. The block matrix of I's above is much clearer than the original 4 by 6 matrix A.

Example 3 **(Important special case)** Let the blocks of A be its n columns. Let the blocks of B be its n rows. Then block multiplication AB adds up ***columns times rows***:

Columns times rows
$$\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & b_1 & - \\ & \vdots & \\ - & b_n & - \end{bmatrix} = \begin{bmatrix} a_1 b_1 + \cdots + a_n b_n \end{bmatrix}. \quad (2)$$

This is another way to multiply matrices. Compare it with the usual rows times columns. Row 1 of A times column 1 of B gave the $(1,1)$ entry in AB. Now *column* 1 of A times *row* 1 of B gives a full matrix—not just a single number. Look at this example:

$$\begin{bmatrix} 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Column 1 times row 1 + Column 2 times row 2
$$= \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 5 & 0 \end{bmatrix}. \quad (3)$$

We stop there so you can see columns multiplying rows. If a 2 by 1 matrix (a column) multiplies a 1 by 2 matrix (a row), the result is 2 by 2. That is what we found. Dot products are *inner* products and these are *outer* products. In the top left corner the answer is $3 + 4 = 7$. This agrees with the row-column dot product of $(1,4)$ with $(3,1)$.

Summary The usual way, rows times columns, gives four dot products (8 multiplications). The new way, columns times rows, gives two full matrices (the same 8 multiplications). The 8 multiplications, and the 4 additions, are just executed in a different order.

Example 4 **(Elimination by blocks)** Suppose the first column of A contains $1, 3, 4$. To change 3 and 4 to 0 and 0, multiply the pivot row by 3 and 4 and subtract. Those row operations are really multiplications by elimination matrices E_{21} and E_{31}:

One at a time
$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}.$$

The "block idea" is to do both eliminations with one matrix E. That matrix clears out the whole first column of A below the pivot $a = 1$:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \text{ multiplies } \begin{bmatrix} \mathbf{1} & x & x \\ \mathbf{3} & x & x \\ \mathbf{4} & x & x \end{bmatrix} \text{ to give } EA = \begin{bmatrix} \mathbf{1} & x & x \\ \mathbf{0} & x & x \\ \mathbf{0} & x & x \end{bmatrix}.$$

Using inverses from 2.5, a block matrix E can do elimination on a whole (block) column of A. Suppose A has four blocks A, B, C, D. Watch how E multiplies A by blocks:

Block elimination
$$\left[\begin{array}{c|c} I & \mathbf{0} \\ \hline -CA^{-1} & I \end{array}\right] \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline \mathbf{0} & \boldsymbol{D - CA^{-1}B} \end{array}\right]. \quad (4)$$

Elimination multiplies the first row $[\,A \;\; B\,]$ by CA^{-1} (previously c/a). It subtracts from C to get a zero block in the first column. It subtracts from D to get $S = D - CA^{-1}B$.

The 3-step paths are counted by A^3; we look at paths to node 2:

$$A^3 = \begin{bmatrix} 3 & \mathbf{2} \\ 2 & \mathbf{1} \end{bmatrix} \quad \text{counts the paths with three steps} \quad \begin{bmatrix} \cdots & \text{1 to 1 to 1 to 2, 1 to 2 to 1 to 2} \\ \cdots & \text{2 to 1 to 1 to 2} \end{bmatrix}$$

These A^k contain the Fibonacci numbers $0, 1, 1, 2, 3, 5, 8, 13, \ldots$ coming in Section 6.2. Multiplying A by A^k involves Fibonacci's rule $F_{k+2} = F_{k+1} + F_k$ (as in $13 = 8 + 5$):

$$(A)(A^k) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{bmatrix} = \begin{bmatrix} F_{k+2} & F_{k+1} \\ F_{k+1} & F_k \end{bmatrix} = A^{k+1}.$$

There are 13 six-step paths from node 1 to node 1, but I can't find them all.

A^k *also counts words.* A path like 1 to 1 to 2 to 1 corresponds to the word **aaba**. The letter **b** can't repeat because there is no edge from 2 to 2. The i, j entry of A^k counts the words of length $k + 1$ that start with the ith letter and end with the jth.

Problem Set 2.4

Problems 1–16 are about the laws of matrix multiplication.

1 A is 3 by 5, B is 5 by 3, C is 5 by 1, and D is 3 by 1. *All entries are* 1. Which of these matrix operations are allowed, and what are the results?

$$BA \qquad AB \qquad ABD \qquad DBA \qquad A(B + C).$$

2 What rows or columns or matrices do you multiply to find

(a) the third column of AB?

(b) the first row of AB?

(c) the entry in row 3, column 4 of AB?

(d) the entry in row 1, column 1 of CDE?

3 Add AB to AC and compare with $A(B + C)$:

$$A = \begin{bmatrix} 1 & 5 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}.$$

4 In Problem 3, multiply A times BC. Then multiply AB times C.

5 Compute A^2 and A^3. Make a prediction for A^5 and A^n:

$$A = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}.$$

6 Show that $(A + B)^2$ is different from $A^2 + 2AB + B^2$, when

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}.$$

Write down the correct rule for $(A + B)(A + B) = A^2 + \_\_\_\_ + B^2$.

7 True or false. Give a specific example when false:

(a) If columns 1 and 3 of B are the same, so are columns 1 and 3 of AB.

(b) If rows 1 and 3 of B are the same, so are rows 1 and 3 of AB.

(c) If rows 1 and 3 of A are the same, so are rows 1 and 3 of ABC.

(d) $(AB)^2 = A^2B^2$.

8 How is each row of DA and EA related to the rows of A, when

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \quad \text{and} \quad E = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}?$$

How is each column of AD and AE related to the columns of A?

9 Row 1 of A is added to row 2. This gives EA below. Then column 1 of EA is added to column 2 to produce $(EA)F$:

$$EA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix}$$

$$\text{and} \quad (EA)F = (EA) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & a+b \\ a+c & a+c+b+d \end{bmatrix}.$$

(a) Do those steps in the opposite order. First add column 1 of A to column 2 by AF, then add row 1 of AF to row 2 by $E(AF)$.

(b) Compare with $(EA)F$. What law is obeyed by matrix multiplication?

10 Row 1 of A is again added to row 2 to produce EA. Then F adds row 2 of EA to row 1. The result is $F(EA)$:

$$F(EA) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix} = \begin{bmatrix} 2a+c & 2b+d \\ a+c & b+d \end{bmatrix}.$$

(a) Do those steps in the opposite order: first add row 2 to row 1 by FA, then add row 1 of FA to row 2.

(b) What law is or is not obeyed by matrix multiplication?

11 (3 by 3 matrices) Choose the only B so that for every matrix A

(a) $BA = 4A$
(b) $BA = 4B$
(c) BA has rows 1 and 3 of A reversed and row 2 unchanged
(d) All rows of BA are the same as row 1 of A.

12 Suppose $AB = BA$ and $AC = CA$ for these two particular matrices B and C:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ commutes with } B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Prove that $a = d$ and $b = c = 0$. Then A is a multiple of I. The only matrices that commute with B and C and all other 2 by 2 matrices are A = multiple of I.

13 Which of the following matrices are guaranteed to equal $(A - B)^2$: $A^2 - B^2$, $(B - A)^2$, $A^2 - 2AB + B^2$, $A(A - B) - B(A - B)$, $A^2 - AB - BA + B^2$?

14 True or false:

(a) If A^2 is defined then A is necessarily square.
(b) If AB and BA are defined then A and B are square.
(c) If AB and BA are defined then AB and BA are square.
(d) If $AB = B$ then $A = I$.

15 If A is m by n, how many separate multiplications are involved when

(a) A multiplies a vector $\boldsymbol{x}$ with n components?
(b) A multiplies an n by p matrix B?
(c) A multiplies itself to produce A^2? Here $m = n$.

16 For $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 4 \\ 1 & 0 & 6 \end{bmatrix}$, compute these answers *and nothing more*:

(a) column 2 of AB
(b) row 2 of AB
(c) row 2 of $AA = A^2$
(d) row 2 of $AAA = A^3$.

Problems 17–19 use a_{ij} for the entry in row i, column j of A.

17 Write down the 3 by 3 matrix A whose entries are

(a) a_{ij} = minimum of i and j
(b) $a_{ij} = (-1)^{i+j}$
(c) $a_{ij} = i/j$.

18 What words would you use to describe each of these classes of matrices? Give a 3 by 3 example in each class. Which matrix belongs to all four classes?

(a) $a_{ij} = 0$ if $i \neq j$

(b) $a_{ij} = 0$ if $i < j$

(c) $a_{ij} = a_{ji}$

(d) $a_{ij} = a_{1j}$.

19 The entries of A are a_{ij}. Assuming that zeros don't appear, what is

(a) the first pivot?

(b) the multiplier ℓ_{31} of row 1 to be subtracted from row 3?

(c) the new entry that replaces a_{32} after that subtraction?

(d) the second pivot?

Problems 20–24 involve powers of A.

20 Compute A^2, A^3, A^4 and also Av, A^2v, A^3v, A^4v for

$$A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}.$$

21 Find all the powers $A^2, A^3, \ldots$ and $AB, (AB)^2, \ldots$ for

$$A = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

22 By trial and error find real nonzero 2 by 2 matrices such that

$$A^2 = -I \qquad BC = 0 \qquad DE = -ED \text{ (not allowing } DE = 0\text{)}.$$

23 (a) Find a nonzero matrix A for which $A^2 = 0$.

(b) Find a matrix that has $A^2 \neq 0$ but $A^3 = 0$.

24 By experiment with $n = 2$ and $n = 3$ predict A^n for these matrices:

$$A_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad A_3 = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}.$$

2.5 Inverse Matrices

Suppose A is a square matrix. We look for an "***inverse matrix***" A^{-1} of the same size, such that A^{-1} *times* A *equals* I. Whatever A does, A^{-1} undoes. Their product is the identity matrix—which does nothing to a vector, so $A^{-1}Ax = x$. *But* A^{-1} *might not exist.*

What a matrix mostly does is to multiply a vector x. Multiplying $Ax = b$ by A^{-1} gives $A^{-1}Ax = A^{-1}b$. ***This is*** $x = A^{-1}b$. The product $A^{-1}A$ is like multiplying by a number and then dividing by that number. A number has an inverse if it is not zero—matrices are more complicated and more interesting. The matrix A^{-1} is called "A inverse."

DEFINITION The matrix A is ***invertible*** if there exists a matrix A^{-1} such that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I. \tag{1}$$

Not all matrices have inverses. This is the first question we ask about a square matrix: Is A invertible? We don't mean that we immediately calculate A^{-1}. In most problems we never compute it! Here are six "notes" about A^{-1}.

Note 1 ***The inverse exists if and only if elimination produces n pivots*** (row exchanges are allowed). Elimination solves $Ax = b$ without explicitly using the matrix A^{-1}.

Note 2 The matrix A cannot have two different inverses. Suppose $BA = I$ and also $AC = I$. Then $B = C$, according to this "proof by parentheses":

$$B(AC) = (BA)C \quad \text{gives} \quad BI = IC \quad \text{or} \quad B = C. \tag{2}$$

This shows that a *left-inverse* B (multiplying from the left) and a *right-inverse* C (multiplying A from the right to give $AC = I$) must be the *same matrix*.

Note 3 If A is invertible, the one and only solution to $Ax = b$ is $x = A^{-1}b$:

Multiply $Ax = b$ ***by*** A^{-1}. ***Then*** $x = A^{-1}Ax = A^{-1}b$.

Note 4 (Important) ***Suppose there is a nonzero vector x such that*** $Ax = \mathbf{0}$. *Then A cannot have an inverse.* No matrix can bring $\mathbf{0}$ back to x.

If A is invertible, then $Ax = \mathbf{0}$ can only have the zero solution $x = A^{-1}\mathbf{0} = \mathbf{0}$.

Note 5 A 2 by 2 matrix is invertible if and only if $ad - bc$ is not zero:

$$\textbf{2 by 2 Inverse:} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{3}$$

This number $ad - bc$ is the *determinant* of A. A matrix is invertible if its determinant is no[t] zero (Chapter 5). The test for n pivots is usually decided before the determinant appears.

Note 6 A diagonal matrix has an inverse provided no diagonal entries are zero:

$$\text{If} \quad A = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \quad \text{then} \quad A^{-1} = \begin{bmatrix} 1/d_1 & & \\ & \ddots & \\ & & 1/d_n \end{bmatrix}.$$

Example 1 The 2 by 2 matrix $A = \left[\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix}\right]$ is not invertible. It fails the test in Note 5, because $ad - bc$ equals $2 - 2 = 0$. It fails the test in Note 3, because $Ax = \mathbf{0}$ when $x = (2, -1)$. It fails to have two pivots as required by Note 1.

Elimination turns the second row of this matrix A into a zero row.

The Inverse of a Product AB

For two nonzero numbers a and b, the sum $a + b$ might or might not be invertible. The numbers $a = 3$ and $b = -3$ have inverses $\frac{1}{3}$ and $-\frac{1}{3}$. Their sum $a + b = 0$ has no inverse. But the product $ab = -9$ does have an inverse, which is $\frac{1}{3}$ times $-\frac{1}{3}$.

For two matrices A and B, the situation is similar. It is hard to say much about the invertibility of $A + B$. But the *product* AB has an inverse, if and only if the two factors A and B are separately invertible (and the same size). The important point is that A^{-1} and B^{-1} come in ***reverse order***:

If A and B are invertible then so is AB. The inverse of a product AB is

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{4}$$

To see why the order is reversed, multiply AB times $B^{-1}A^{-1}$. Inside that is $BB^{-1} = I$:

$$\textbf{Inverse of } AB \quad (AB)(B^{-1}A^{-1}) = AIA^{-1} = AA^{-1} = I.$$

We moved parentheses to multiply BB^{-1} first. Similarly $B^{-1}A^{-1}$ times AB equals I. This illustrates a basic rule of mathematics: Inverses come in reverse order. It is also common sense: If you put on socks and then shoes, the first to be taken off are the _____. The same reverse order applies to three or more matrices:

$$\textbf{Reverse order} \quad (ABC)^{-1} = C^{-1}B^{-1}A^{-1}. \tag{5}$$

Example 2 ***Inverse of an elimination matrix.*** If E subtracts 5 times row 1 from row 2, then E^{-1} *adds* 5 times row 1 to row 2:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Multiply EE^{-1} to get the identity matrix I. Also multiply $E^{-1}E$ to get I. We are adding and subtracting the same 5 times row 1. Whether we add and then subtract (this is EE^{-1}) or subtract and then add (this is $E^{-1}E$), we are back at the start.

For square matrices, an inverse on one side is automatically an inverse on the other side. If $AB = I$ then automatically $BA = I$. In that case B is A^{-1}. This is very useful to know but we are not ready to prove it.

Example 3 Suppose F subtracts 4 times row 2 from row 3, and F^{-1} adds it back:

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \quad \text{and} \quad F^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}.$$

Now multiply F by the matrix E in Example 2 to find FE. Also multiply E^{-1} times F^{-1} to find $(FE)^{-1}$. Notice the orders FE and $E^{-1}F^{-1}$!

$$FE = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 20 & -4 & 1 \end{bmatrix} \quad \text{is inverted by} \quad E^{-1}F^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}. \tag{6}$$

The result is beautiful and correct. The product FE contains "20" but its inverse doesn't. E subtracts 5 times row 1 from row 2. Then F subtracts 4 times the *new* row 2 (changed by row 1) from row 3. ***In this order*** FE***, row 3 feels an effect from row 1.***

In the order $E^{-1}F^{-1}$, that effect does not happen. First F^{-1} adds 4 times row 2 to row 3. After that, E^{-1} adds 5 times row 1 to row 2. There is no 20, because row 3 doesn't change again. ***In this order*** $E^{-1}F^{-1}$***, row 3 feels no effect from row 1.***

> In elimination order F follows E. In reverse order E^{-1} follows F^{-1}.
> $E^{-1}F^{-1}$ ***is quick. The multipliers 5, 4 fall into place below the diagonal of 1's.***

This special multiplication $E^{-1}F^{-1}$ and $E^{-1}F^{-1}G^{-1}$ will be useful in the next section. We will explain it again, more completely. In this section our job is A^{-1}, and we expect some serious work to compute it. Here is a way to organize that computation.

Calculating A^{-1} by Gauss-Jordan Elimination

I hinted that A^{-1} might not be explicitly needed. The equation $Ax = b$ is solved by $x = A^{-1}b$. But it is not necessary or efficient to compute A^{-1} and multiply it times b. *Elimination goes directly to* x. Elimination is also the way to calculate A^{-1}, as we now show. The Gauss-Jordan idea is to solve $AA^{-1} = I$, *finding each column of* A^{-1}.

A multiplies the first column of A^{-1} (call that x_1) to give the first column of I (call that e_1). This is our equation $Ax_1 = e_1 = (1,0,0)$. There will be two more equations. Each of the columns x_1, x_2, x_3 of A^{-1} is multiplied by A to produce a column of I:

$$\textbf{3 columns of } A^{-1} \qquad AA^{-1} = A\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = I. \tag{7}$$

To invert a 3 by 3 matrix A, we have to solve three systems of equations: $Ax_1 = e_1$ and $Ax_2 = e_2 = (0,1,0)$ and $Ax_3 = e_3 = (0,0,1)$. Gauss-Jordan finds A^{-1} this way.

The *Gauss-Jordan method* computes A^{-1} by solving *all n equations together*. Usually the "augmented matrix" $[A \ \ b]$ has one extra column b. Now we have three right sides e_1, e_2, e_3 (when A is 3 by 3). They are the columns of I, so the augmented matrix is really the block matrix $[A \ \ I]$. I take this chance to invert my favorite matrix K, with 2's on the main diagonal and -1's next to the 2's:

$$[K \ e_1 \ e_2 \ e_3] = \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{bmatrix} \quad \textbf{Start Gauss-Jordan on } K$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \frac{3}{2} & \mathbf{-1} & \frac{1}{2} & \mathbf{1} & \mathbf{0} \\ 0 & -1 & 2 & 0 & 0 & 1 \end{bmatrix} \quad (\tfrac{1}{2}\ \textit{row}\ 1 + \textit{row}\ 2)$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & \mathbf{1} \end{bmatrix} \quad (\tfrac{2}{3}\ \textit{row}\ 2 + \textit{row}\ 3)$$

We are halfway to K^{-1}. The matrix in the first three columns is U (upper triangular). The pivots $2, \frac{3}{2}, \frac{4}{3}$ are on its diagonal. Gauss would finish by back substitution. The contribution of Jordan is *to continue with elimination*! He goes all the way to the ***"reduced echelon form"***. Rows are added to rows above them, to produce ***zeros above the pivots***:

$$\begin{pmatrix} \text{Zero above} \\ \text{third pivot} \end{pmatrix} \to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \frac{3}{2} & \mathbf{0} & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{bmatrix} \quad (\tfrac{3}{4}\ \textit{row}\ 3 + \textit{row}\ 2)$$

$$\begin{pmatrix} \text{Zero above} \\ \text{second pivot} \end{pmatrix} \to \begin{bmatrix} \mathbf{2} & \mathbf{0} & \mathbf{0} & \frac{3}{2} & \mathbf{1} & \frac{1}{2} \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{bmatrix} \quad (\tfrac{2}{3}\ \textit{row}\ 2 + \textit{row}\ 1)$$

The last Gauss-Jordan step is to divide each row by its pivot. The new pivots are 1. We have reached I in the first half of the matrix, because K is invertible. ***The three columns of K^{-1} are in the second half of $[I \ \ K^{-1}]$:***

$$\begin{matrix} \text{(divide by 2)} \\ \text{(divide by } \frac{3}{2}) \\ \text{(divide by } \frac{4}{3}) \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix} = [I \ \ x_1 \ \ x_2 \ \ x_3] = [I \ \ K^{-1}].$$

Starting from the 3 by 6 matrix $[K \ \ I]$, we ended with $[I \ \ K^{-1}]$. Here is the whole Gauss-Jordan process on one line for any invertible matrix A:

Gauss-Jordan ***Multiply $[A \ \ I]$ by A^{-1} to get $[I \ \ A^{-1}]$.***

The elimination steps create the inverse matrix while changing A to I. For large matrices, we probably don't want A^{-1} at all. But for small matrices, it can be very worthwhile to know the inverse. We add three observations about this particular K^{-1} because it is an important example. We introduce the words *symmetric*, *tridiagonal*, and *determinant*:

1. K is ***symmetric*** across its main diagonal. So is K^{-1}.

2. K is ***tridiagonal*** (only three nonzero diagonals). But K^{-1} is a dense matrix with no zeros. That is another reason we don't often compute inverse matrices. The inverse of a band matrix is generally a dense matrix.

3. The *product of pivots* is $2(\frac{3}{2})(\frac{4}{3}) = 4$. This number 4 is the ***determinant*** of K.

$$K^{-1} \textbf{\textit{ involves division by the determinant}} \qquad K^{-1} = \frac{1}{4}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}. \tag{8}$$

This is why an invertible matrix cannot have a zero determinant.

Example 4 Find A^{-1} by Gauss-Jordan elimination starting from $A = \left[\begin{smallmatrix} 2 & 3 \\ 4 & 7 \end{smallmatrix}\right]$. There are two row operations and then a division to put 1's in the pivots:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{3} & 1 & 0 \\ \mathbf{4} & \mathbf{7} & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 1 & 0 \\ \mathbf{0} & \mathbf{1} & -2 & 1 \end{bmatrix} \quad (\text{this is } \begin{bmatrix} U & L^{-1} \end{bmatrix})$$

$$\to \begin{bmatrix} \mathbf{2} & \mathbf{0} & 7 & -3 \\ 0 & 1 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{7}{2} & -\frac{3}{2} \\ 0 & 1 & -2 & 1 \end{bmatrix} \quad (\text{this is } \begin{bmatrix} I & A^{-1} \end{bmatrix}).$$

That A^{-1} involves division by the determinant $ad - bc = 2 \cdot 7 - 3 \cdot 4 = 2$. The code for $X =$ inverse(A) can use **rref**, the "row reduced echelon form" from Chapter 3:

```
I = eye (n);                 % Define the n by n identity matrix
R = rref ([A I]);            % Eliminate on the augmented matrix [A I]
X = R(:, n + 1 : n + n)      % Pick A^-1 from the last n columns of R
```

A must be invertible, or elimination cannot reduce it to I (in the left half of R).

Gauss-Jordan shows why A^{-1} is expensive. We must solve n equations for its n columns.

To solve $Ax = b$ without A^{-1}, we deal with *one* column b to find one column x.

In defense of A^{-1}, we want to say that its cost is not n times the cost of one system $Ax = b$. Surprisingly, the cost for n columns is only multiplied by 3. This saving is because the n equations $Ax_i = e_i$ all involve the same matrix A. Working with the right sides is relatively cheap, because elimination only has to be done once on A.

The complete A^{-1} needs n^3 elimination steps, where a single x needs $n^3/3$. The next section calculates these costs.

L goes to I by a product of elimination matrices $E_{32}E_{31}E_{21}$. So that product is L^{-1}. All pivots are 1's (a full set). L^{-1} is lower triangular, with the strange entry "11".

That 11 does not appear to spoil 3, 4, 5 in the good order $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = L$.

■ REVIEW OF THE KEY IDEAS ■

1. The inverse matrix gives $AA^{-1} = I$ and $A^{-1}A = I$.
2. A is invertible if and only if it has n pivots (row exchanges allowed).
3. If $A\boldsymbol{x} = \boldsymbol{0}$ for a nonzero vector $\boldsymbol{x}$, then A has no inverse.
4. The inverse of AB is the reverse product $B^{-1}A^{-1}$. And $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
5. The Gauss-Jordan method solves $AA^{-1} = I$ to find the n columns of A^{-1}. The augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$ is row-reduced to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$.

■ WORKED EXAMPLES ■

2.5 A The inverse of a triangular **difference matrix** A is a triangular **sum matrix** S:

$$\begin{bmatrix} A & I \end{bmatrix} = \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right] = \begin{bmatrix} I & A^{-1} \end{bmatrix} = \begin{bmatrix} I & \textit{sum matrix} \end{bmatrix}.$$

If I change a_{13} to -1, then all rows of A add to zero. The equation $A\boldsymbol{x} = \boldsymbol{0}$ will now have the nonzero solution $\boldsymbol{x} = (1, 1, 1)$. A clear signal: ***This new A can't be inverted.***

Problem Set 2.5

1 Find the inverses (directly or from the 2 by 2 formula) of A, B, C:

$$A = \begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 3 & 4 \\ 5 & 7 \end{bmatrix}.$$

2 For these "permutation matrices" find P^{-1} by trial and error (with 1's and 0's):

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

3 Solve for the first column (x, y) and second column (t, z) of A^{-1}:

$$\begin{bmatrix} 10 & 20 \\ 20 & 50 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 10 & 20 \\ 20 & 50 \end{bmatrix}\begin{bmatrix} t \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

4 Show that $\left[\begin{smallmatrix} 1 & 2 \\ 3 & 6 \end{smallmatrix}\right]$ is not invertible by trying to solve $AA^{-1} = I$ for column 1 of A^{-1}:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \left(\begin{matrix} \textit{For a different } A, \text{ could column 1 of } A^{-1} \\ \text{be possible to find but not column 2?} \end{matrix}\right)$$

5 Find an upper triangular U (not diagonal) with $U^2 = I$ which gives $U = U^{-1}$.

6 (a) If A is invertible and $AB = AC$, prove quickly that $B = C$.

(b) If $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, find two different matrices such that $AB = AC$.

7 (Important) If A has row 1 + row 2 = row 3, show that A is not invertible:

(a) Explain why $A\boldsymbol{x} = (1, 0, 0)$ cannot have a solution.

(b) Which right sides (b_1, b_2, b_3) might allow a solution to $A\boldsymbol{x} = \boldsymbol{b}$?

(c) What happens to row 3 in elimination?

8 If A has column 1 + column 2 = column 3, show that A is not invertible:

(a) Find a nonzero solution $\boldsymbol{x}$ to $A\boldsymbol{x} = \boldsymbol{0}$. The matrix is 3 by 3.

(b) Elimination keeps column 1 + column 2 = column 3. Explain why there is no third pivot.

9 Suppose A is invertible and you exchange its first two rows to reach B. Is the new matrix B invertible and how would you find B^{-1} from A^{-1}?

10 Find the inverses (in any legal way) of

$$A = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ 0 & 4 & 0 & 0 \\ 5 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 2 & 0 & 0 \\ 4 & 3 & 0 & 0 \\ 0 & 0 & 6 & 5 \\ 0 & 0 & 7 & 6 \end{bmatrix}.$$

2.6 Elimination = Factorization: $A = LU$

Students often say that mathematics courses are too theoretical. Well, not this section. It is almost purely practical. The goal is to describe Gaussian elimination in the most useful way. Many key ideas of linear algebra, when you look at them closely, are really *factorizations* of a matrix. The original matrix A becomes the product of two or three special matrices. The first factorization—also the most important in practice—comes now from elimination. ***The factors L and U are triangular matrices. The factorization that comes from elimination is $A = LU$.***

We already know U, the upper triangular matrix with the pivots on its diagonal. The elimination steps take A to U. We will show how reversing those steps (taking U back to A) is achieved by a lower triangular L. ***The entries of L are exactly the multipliers*** ℓ_{ij}—which multiplied the pivot row j when it was subtracted from row i.

Start with a 2 by 2 example. The matrix A contains 2, 1, 6, 8. The number to eliminate is 6. ***Subtract 3 times row*** 1 ***from row*** 2. That step is E_{21} in the forward direction with multiplier $\ell_{21} = 3$. The return step from U to A is $L = E_{21}^{-1}$ (an addition using $+3$):

$$\textbf{\textit{Forward from } } A \textbf{\textit{ to }} U: \quad E_{21}A = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 5 \end{bmatrix} = U$$

$$\textbf{\textit{Back from }} U \textbf{\textit{ to }} A: \quad E_{21}^{-1}U = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 6 & 8 \end{bmatrix} = A.$$

The second line is our factorization $LU = A$. Instead of E_{21}^{-1} we write L. Move now to larger matrices with many E's. ***Then L will include all their inverses.***

Each step from A to U multiplies by a matrix E_{ij} to produce zero in the (i, j) position. To keep this clear, we stay with the most frequent case—***when no row exchanges are involved.*** If A is 3 by 3, we multiply by E_{21} and E_{31} and E_{32}. The multipliers ℓ_{ij} produce zeros in the $(2, 1)$ and $(3, 1)$ and $(3, 2)$ positions—all below the diagonal. Elimination ends with the upper triangular U.

Now move those E's onto the other side, *where their inverses multiply* U:

$$(E_{32}E_{31}E_{21})A = U \quad \textbf{becomes} \quad A = (E_{21}^{-1}E_{31}^{-1}E_{32}^{-1})U \quad \textbf{which is} \quad A = LU. \tag{1}$$

The inverses go in opposite order, as they must. That product of three inverses is L. ***We have reached $A = LU$.*** Now we stop to understand it.

Explanation and Examples

First point: Every inverse matrix E^{-1} is *lower triangular*. Its off-diagonal entry is ℓ_{ij}, to undo the subtraction produced by $-\ell_{ij}$. The main diagonals of E and E^{-1} contain 1's. Our example above had $\ell_{21} = 3$ and $E = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$ and $L = E^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$.

Second point: Equation (1) shows a lower triangular matrix (the product of the E_{ij}) multiplying A. It also shows all the E_{ij}^{-1} multiplying U to bring back A. ***This lower triangular product of inverses is L.***

One reason for working with the inverses is that we want to factor A, not U. The "inverse form" gives $A = LU$. Another reason is that we get something extra, almost more than we deserve. This is the third point, showing that L is exactly right.

Third point: Each multiplier ℓ_{ij} goes directly into its i, j position—*unchanged*—in the product of inverses which is L. Usually matrix multiplication will mix up all the numbers. Here that doesn't happen. The order is right for the inverse matrices, to keep the ℓ's unchanged. The reason is given below in equation (3).

Since each E^{-1} has 1's down its diagonal, the final good point is that L does too.

($A = LU$) *This is elimination without row exchanges.* The upper triangular U has the pivots on its diagonal. The lower triangular L has all 1's on its diagonal. ***The multipliers ℓ_{ij} are below the diagonal of L.***

Example 1 Elimination subtracts $\frac{1}{2}$ times row 1 from row 2. The last step subtracts $\frac{2}{3}$ times row 2 from row 3. The lower triangular L has $\ell_{21} = \frac{1}{2}$ and $\ell_{32} = \frac{2}{3}$. Multiplying LU produces A:

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & \frac{4}{3} \end{bmatrix} = LU.$$

The (3, 1) multiplier is zero because the (3, 1) entry in A is zero. No operation needed.

Example 2 Change the top left entry from 2 to 1. The pivots all become 1. The multipliers are all 1. That pattern continues when A is 4 by 4:

$$\textbf{Special pattern} \quad A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 0 & 1 & 1 & \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ & 1 & 1 & 0 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}.$$

These LU examples are showing something extra, which is very important in practice. Assume no row exchanges. When can we predict *zeros* in L and U?

When a row of A starts with zeros, so does that row of L.

When a column of A starts with zeros, so does that column of U.

If a row starts with zero, we don't need an elimination step. L has a zero, which saves computer time. Similarly, zeros at the *start* of a column survive into U. But please realize: Zeros in the *middle* of a matrix are likely to be filled in, while elimination sweeps forward. We now explain why L has the multipliers ℓ_{ij} in position, with no mix-up.

The key reason why A equals LU: Ask yourself about the pivot rows that are subtracted from lower rows. Are they the original rows of A? *No*, elimination probably changed them. Are they rows of U? *Yes*, the pivot rows never change again. When computing the third row of U, we subtract multiples of earlier rows of U (*not rows of A!*):

$$\text{Row 3 of } U = (\text{Row 3 of } A) - \ell_{31}(\text{Row 1 of } U) - \ell_{32}(\text{Row 2 of } U). \tag{2}$$

Rewrite this equation to see that the row $[\ell_{31} \quad \ell_{32} \quad 1]$ is multiplying U:

$$(\text{Row 3 of } A) = \ell_{31}(\text{Row 1 of } U) + \ell_{32}(\text{Row 2 of } U) + 1(\text{Row 3 of } U). \tag{3}$$

***This is exactly row* 3 *of* $A = LU$.** That row of L holds $\ell_{31}, \ell_{32}, 1$. All rows look like this, whatever the size of A. With no row exchanges, we have $A = LU$.

Better balance The LU factorization is "unsymmetric" because U has the pivots on its diagonal where L has 1's. This is easy to change. ***Divide U by a diagonal matrix D that contains the pivots.*** That leaves a new matrix with 1's on the diagonal:

$$\text{Split } U \text{ into } \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & u_{12}/d_1 & u_{13}/d_1 & \cdot \\ & 1 & u_{23}/d_2 & \cdot \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}.$$

It is convenient (but a little confusing) to keep the same letter U for this new upper triangular matrix. It has 1's on the diagonal (like L). Instead of the normal LU, the new form has D in the middle: ***Lower triangular L times diagonal D times upper triangular U.***

The triangular factorization can be written $A = LU$ *or* $A = LDU$.

Whenever you see LDU, it is understood that U has 1's on the diagonal. *Each row is divided by its first nonzero entry—the pivot.* Then L and U are treated evenly in LDU:

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 8 \\ 0 & 5 \end{bmatrix} \quad \text{splits further into} \quad \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & \\ & 5 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}. \tag{4}$$

The pivots 2 and 5 went into D. Dividing the rows by 2 and 5 left the rows $[1 \quad 4]$ and $[0 \quad 1]$ in the new U with diagonal ones. The multiplier 3 is still in L.

My own lectures sometimes stop at this point. The next paragraphs show how elimination codes are organized, and how long they take. If MATLAB (or any software) is available, you can measure the computing time by just counting the seconds.

One Square System = Two Triangular Systems

The matrix L contains our memory of Gaussian elimination. It holds the numbers that multiplied the pivot rows, before subtracting them from lower rows. When do we need this record and how do we use it in solving $A\boldsymbol{x} = \boldsymbol{b}$?

We need L as soon as there is a *right side* $\boldsymbol{b}$. The factors L and U were completely decided by the left side (the matrix A). On the right side of $A\boldsymbol{x} = \boldsymbol{b}$, we use L^{-1} and then U^{-1}. That ***Solve*** step deals with two triangular matrices.

1 ***Factor*** (into L and U, by elimination on the left side matrix A)

2 ***Solve*** (forward elimination on $\boldsymbol{b}$ using L, then back substitution for $\boldsymbol{x}$ using U).

Earlier, we worked on A and $\boldsymbol{b}$ at the same time. No problem with that—just augment to $[\,A \;\; \boldsymbol{b}\,]$. But most computer codes keep the two sides separate. The memory of elimination is held in L and U, to process $\boldsymbol{b}$ whenever we want to. The User's Guide to LAPACK remarks that "This situation is so common and the savings are so important that no provision has been made for solving a single system with just one subroutine."

How does ***Solve*** work on $\boldsymbol{b}$? First, apply forward elimination to the right side (the multipliers are stored in L, use them now). This changes $\boldsymbol{b}$ to a new right side $\boldsymbol{c}$. *We are really solving* $L\boldsymbol{c} = \boldsymbol{b}$. Then back substitution solves $U\boldsymbol{x} = \boldsymbol{c}$ as always. The original system $A\boldsymbol{x} = \boldsymbol{b}$ is factored into ***two triangular systems***:

Forward and backward *Solve* $L\boldsymbol{c} = \boldsymbol{b}$ *and then solve* $U\boldsymbol{x} = \boldsymbol{c}$. (5)

To see that $\boldsymbol{x}$ is correct, multiply $U\boldsymbol{x} = \boldsymbol{c}$ by L. Then $LU\boldsymbol{x} = L\boldsymbol{c}$ is just $A\boldsymbol{x} = \boldsymbol{b}$.

To emphasize: There is *nothing new* about those steps. This is exactly what we have done all along. We were really solving the triangular system $L\boldsymbol{c} = \boldsymbol{b}$ as elimination went forward. Then back substitution produced $\boldsymbol{x}$. An example shows what we actually did.

Example 3 Forward elimination (downward) on $A\boldsymbol{x} = \boldsymbol{b}$ ends at $U\boldsymbol{x} = \boldsymbol{c}$:

$$A\boldsymbol{x} = \boldsymbol{b} \quad \begin{aligned} u + 2v &= 5 \\ 4u + 9v &= 21 \end{aligned} \quad \text{becomes} \quad \begin{aligned} u + 2v &= 5 \\ v &= 1 \end{aligned} \quad U\boldsymbol{x} = \boldsymbol{c}$$

The multiplier was 4, which is saved in L. The right side used it to change 21 to 1:

$L\boldsymbol{c} = \boldsymbol{b}$ The lower triangular system $\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{c} \end{bmatrix} = \begin{bmatrix} 5 \\ 21 \end{bmatrix}$ gave $\boldsymbol{c} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

$U\boldsymbol{x} = \boldsymbol{c}$ The upper triangular system $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ gives $\boldsymbol{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

L and U can go into the n^2 storage locations that originally held A (now forgettable).

The Cost of Elimination

A very practical question is cost—or computing time. We can solve 1000 equations on a PC. What if $n = 100{,}000$? (*Not if A is dense.*) Large systems come up all the time in scientific computing, where a three-dimensional problem can easily lead to a million unknowns. We can let the calculation run overnight, but we can't leave it for 100 years.

Problems 15-16 use L and U (without needing A) to solve $Ax = b$.

15 Solve the triangular system $Lc = b$ to find c. Then solve $Ux = c$ to find x:

$$L = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 2 \\ 11 \end{bmatrix}.$$

For safety multiply LU and solve $Ax = b$ as usual. Circle c when you see it.

16 Solve $Lc = b$ to find c. Then solve $Ux = c$ to find x. **What was** A?

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

17 (a) When you apply the usual elimination steps to L, what matrix do you reach?

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}.$$

(b) When you apply the same steps to I, what matrix do you get?

(c) When you apply the same steps to LU, what matrix do you get?

18 If $A = LDU$ and also $A = L_1 D_1 U_1$ with all factors invertible, then $L = L_1$ and $D = D_1$ and $U = U_1$. "*The three factors are unique.*"

Derive the equation $L_1^{-1} L D = D_1 U_1 U^{-1}$. Are the two sides triangular or diagonal? Deduce $L = L_1$ and $U = U_1$ (they all have diagonal 1's). Then $D = D_1$.

19 *Tridiagonal matrices* have zero entries except on the main diagonal and the two adjacent diagonals. Factor these into $A = LU$ and $A = LDL^{\mathrm{T}}$:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix}.$$

20 When T is tridiagonal, its L and U factors have only two nonzero diagonals. How would you take advantage of knowing the zeros in T, in a code for Gaussian elimination? Find L and U.

$$\textbf{Tridiagonal} \qquad T = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 4 \end{bmatrix}.$$

21 If A and B have nonzeros in the positions marked by x, which zeros (marked by 0) *stay zero* in their factors L and U?

$$A = \begin{bmatrix} x & x & x & x \\ x & x & x & 0 \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix} \qquad B = \begin{bmatrix} x & x & x & 0 \\ x & x & 0 & x \\ x & 0 & x & x \\ 0 & x & x & x \end{bmatrix}.$$

2.7 Transposes and Permutations

We need one more matrix, and fortunately it is much simpler than the inverse. It is the ***"transpose"*** of A, which is denoted by A^{T}. *The columns of A^{T} are the rows of A.*

When A is an m by n matrix, the transpose is n by m:

$$\textbf{Transpose} \qquad \text{If} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{then} \quad A^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 4 \end{bmatrix}.$$

You can write the rows of A into the columns of A^{T}. Or you can write the columns of A into the rows of A^{T}. The matrix "flips over" its main diagonal. The entry in row i, column j of A^{T} comes from row j, column i of the original A:

$$\textbf{Exchange rows and columns} \qquad (A^{\mathrm{T}})_{ij} = A_{ji}.$$

The transpose of a lower triangular matrix is upper triangular. (But the inverse is still lower triangular.) The transpose of A^{T} is A.

Note MATLAB's symbol for the transpose of A is A'. Typing [1 2 3] gives a row vector and the column vector is $v = [1 \ \ 2 \ \ 3]'$. To enter a matrix M with second column $w = [\ 4\ 5\ 6\ ]'$ you could define $M = [\ v\ w\ ]$. Quicker to enter by rows and then transpose the whole matrix: $M = [1 \ \ 2 \ \ 3; \ \ 4 \ \ 5 \ \ 6]'$.

The rules for transposes are very direct. We can transpose $A + B$ to get $(A + B)^{\mathrm{T}}$. Or we can transpose A and B separately, and then add $A^{\mathrm{T}} + B^{\mathrm{T}}$—with the same result. The serious questions are about the transpose of a product AB and an inverse A^{-1}:

$$\textbf{Sum} \qquad \text{The transpose of} \quad A + B \quad \text{is} \quad A^{\mathrm{T}} + B^{\mathrm{T}}. \tag{1}$$

$$\textbf{Product} \qquad \text{The transpose of} \quad AB \quad \text{is} \quad (AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}. \tag{2}$$

$$\textbf{Inverse} \qquad \text{The transpose of} \quad A^{-1} \quad \text{is} \quad (A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}. \tag{3}$$

Notice especially how $B^{\mathrm{T}}A^{\mathrm{T}}$ comes in reverse order. For inverses, this reverse order was quick to check: $B^{-1}A^{-1}$ times AB produces I. To understand $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$, start with $(Ax)^{\mathrm{T}} = x^{\mathrm{T}}A^{\mathrm{T}}$:

Ax combines the columns of A while $x^{\mathrm{T}}A^{\mathrm{T}}$ combines the rows of A^{T}.

It is the same combination of the same vectors! In A they are columns, in A^{T} they are rows. So the transpose of the column Ax is the row $x^{\mathrm{T}}A^{\mathrm{T}}$. That fits our formula $(Ax)^{\mathrm{T}} = x^{\mathrm{T}}A^{\mathrm{T}}$. Now we can prove the formula $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$, when B has several columns.

If $B = [x_1 \ \ x_2]$ has two columns, apply the same idea to each column. The columns of AB are Ax_1 and Ax_2. Their transposes are the rows of $B^{\mathrm{T}}A^{\mathrm{T}}$:

$$\text{Transposing } AB = \begin{bmatrix} Ax_1 & Ax_2 & \cdots \end{bmatrix} \text{ gives } \begin{bmatrix} x_1^{\mathrm{T}}A^{\mathrm{T}} \\ x_2^{\mathrm{T}}A^{\mathrm{T}} \\ \vdots \end{bmatrix} \text{ which is } B^{\mathrm{T}}A^{\mathrm{T}}. \tag{4}$$

The right answer B^TA^T comes out a row at a time. Here are numbers in $(AB)^T = B^TA^T$:

$$AB = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 9 & 1 \end{bmatrix} \quad \text{and} \quad B^TA^T = \begin{bmatrix} 5 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 9 \\ 0 & 1 \end{bmatrix}.$$

The reverse order rule extends to three or more factors: $(ABC)^T$ equals $C^TB^TA^T$.

If $A = LDU$ *then* $A^T = U^TD^TL^T$. *The pivot matrix has* $D = D^T$.

Now apply this product rule to both sides of $A^{-1}A = I$. On one side, I^T is I. We confirm the rule that $(A^{-1})^T$ *is the inverse of* A^T, because their product is I:

$$\textbf{Transpose of inverse} \qquad A^{-1}A = I \quad \text{is transposed to} \quad A^T(A^{-1})^T = I. \tag{5}$$

Similarly $AA^{-1} = I$ leads to $(A^{-1})^TA^T = I$. We can invert the transpose or we can transpose the inverse. Notice especially: A^T ***is invertible exactly when*** A ***is invertible.***

Example 1 The inverse of $A = \begin{bmatrix} 1 & 0 \\ 6 & 1 \end{bmatrix}$ is $A^{-1} = \begin{bmatrix} 1 & 0 \\ -6 & 1 \end{bmatrix}$. The transpose is $A^T = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$.

$$(A^{-1})^T \quad \textit{and} \quad (A^T)^{-1} \quad \textit{are both equal to} \quad \begin{bmatrix} 1 & -6 \\ 0 & 1 \end{bmatrix}.$$

The Meaning of Inner Products

We know the dot product (inner product) of $\boldsymbol{x}$ and $\boldsymbol{y}$. It is the sum of numbers x_iy_i. Now we have a better way to write $\boldsymbol{x} \cdot \boldsymbol{y}$, without using that unprofessional dot. Use matrix notation instead:

^{T} **is inside**	*The dot product or inner product is* $\boldsymbol{x}^T\boldsymbol{y}$	$(1 \times n)(n \times 1)$
^{T} **is outside**	*The rank one product or outer product is* $\boldsymbol{x}\boldsymbol{y}^T$	$(n \times 1)(1 \times n)$

$\boldsymbol{x}^T\boldsymbol{y}$ is a number, $\boldsymbol{x}\boldsymbol{y}^T$ is a matrix. Quantum mechanics would write those as $< x|y >$ (inner) and $|x >< y|$ (outer). I think the world is governed by linear algebra, but physics disguises it well. Here are examples where the inner product has meaning:

From mechanics	Work = (Movements) (Forces) = $\boldsymbol{x}^T\boldsymbol{f}$
From circuits	Heat loss = (Voltage drops) (Currents) = $\boldsymbol{e}^T\boldsymbol{y}$
From economics	Income = (Quantities) (Prices) = $\boldsymbol{q}^T\boldsymbol{p}$

We are really close to the heart of applied mathematics, and there is one more point to explain. It is the deeper connection between inner products and the transpose of A.

We defined A^T by flipping the matrix across its main diagonal. That's not mathematics. There is a better way to approach the transpose. A^T *is the matrix that makes these two inner products equal for every* $\boldsymbol{x}$ *and* $\boldsymbol{y}$:

$(Ax)^T y = x^T(A^T y)$ **Inner product of** Ax **with** y **= Inner product of** x **with** A^Ty

Example 2 Start with $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ $\quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ $\quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

On one side we have Ax multiplying y: $(x_2 - x_1)y_1 + (x_3 - x_2)y_2$

That is the same as $x_1(-y_1) + x_2(y_1 - y_2) + x_3(y_2)$. Now x is multiplying $A^{\mathrm{T}}y$.

$$A^{\mathrm{T}}y \text{ must be } \begin{bmatrix} -y_1 \\ y_1 - y_2 \\ y_2 \end{bmatrix} \text{ which produces } A^{\mathrm{T}} = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \text{ as expected.}$$

Example 3 Will you allow me a little calculus? It is extremely important or I wouldn't leave linear algebra. (This is really linear algebra for functions $x(t)$.) **The difference matrix changes to a derivative** $A = d/dt$. Its transpose will now come from $(dx/dt, y) = (x, -dy/dt)$.

The inner product changes from a finite sum of $x_k y_k$ to an integral of $x(t)y(t)$.

Inner product of functions $\quad x^{\mathrm{T}}y = (x, y) = \displaystyle\int_{-\infty}^{\infty} x(t)\, y(t)\, dt$ by definition

Transpose rule $(Ax)^{\mathrm{T}}y = x^{\mathrm{T}}(A^{\mathrm{T}}y)$ $\quad \displaystyle\int_{-\infty}^{\infty} \frac{dx}{dt} y(t)\, dt = \int_{-\infty}^{\infty} x(t)\left(-\frac{dy}{dt}\right) dt$ shows A^{T} $\quad$ (6)

I hope you recognize "***integration by parts***". The derivative moves from the first function $x(t)$ to the second function $y(t)$. During that move, a minus sign appears. This tells us that ***the "transpose" of the derivative is minus the derivative***.

The derivative is *anti-symmetric*: $A = d/dt$ ***and*** $A^{\mathrm{T}} = -d/dt$. Symmetric matrices have $A^{\mathrm{T}} = A$, anti-symmetric matrices have $A^{\mathrm{T}} = -A$. In some way, the 2 by 3 difference matrix above followed this pattern. The 3 by 2 matrix A^{T} was *minus* a difference matrix. It produced $y_1 - y_2$ in the middle component of $A^{\mathrm{T}}y$ instead of the difference $y_2 - y_1$.

Symmetric Matrices

For a *symmetric matrix*, transposing A to A^{T} produces no change. Then $A^{\mathrm{T}} = A$. Its (j, i) entry across the main diagonal equals its (i, j) entry. In my opinion, these are the most important matrices of all.

DEFINITION A ***symmetric matrix*** has $A^{\mathrm{T}} = A$. This means that $a_{ji} = a_{ij}$.

Symmetric matrices $\quad A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = A^{\mathrm{T}}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} = D^{\mathrm{T}}$.

The inverse of a symmetric matrix is also symmetric. The transpose of A^{-1} is $(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1} = A^{-1}$. That says A^{-1} is symmetric (when A is invertible):

Symmetric inverses $\quad A^{-1} = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$ and $D^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}$.

Now we produce symmetric matrices by ***multiplying any matrix R by*** R^{T}.

Choose any matrix R, probably rectangular. Multiply R^T times R. Then the product R^TR is automatically a square symmetric matrix:

$$\textit{The transpose of } R^TR \textit{ is } R^T(R^T)^T \textit{ which is } R^TR. \tag{7}$$

That is a quick proof of symmetry for R^TR. We could also look at the (i,j) entry of R^TR. It is the dot product of row i of R^T (column i of R) with column j of R. The (j,i) entry is the same dot product, column j with column i. So R^TR is symmetric.

The matrix RR^T is also symmetric. (The shapes of R and R^T allow multiplication.) But RR^T is a different matrix from R^TR. In our experience, most scientific problems that start with a rectangular matrix R end up with R^TR or RR^T or both. As in least squares.

Example 4 Multiply $R = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ and $R^T = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$ in both orders.

$$RR^T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \text{ and } R^TR = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \text{ are both symmetric matrices.}$$

The product R^TR is n by n. In the opposite order, RR^T is m by m. Both are symmetric, with positive diagonal (*why*?). But even if $m = n$, it is not very likely that $R^TR = RR^T$. Equality can happen, but it is abnormal.

Symmetric matrices in elimination $A^T = A$ makes elimination faster, because we can work with half the matrix (plus the diagonal). It is true that the upper triangular U is probably not symmetric. ***The symmetry is in the triple product*** $A = LDU$. Remember how the diagonal matrix D of pivots can be divided out, to leave 1's on the diagonal of both L and U:

$$\begin{bmatrix} 1 & 2 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \qquad LU \text{ misses the symmetry of } A$$

$$= \begin{bmatrix} \mathbf{1} & 0 \\ \mathbf{2} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 0 & \mathbf{1} \end{bmatrix} \qquad \begin{matrix} LDU \text{ captures the symmetry} \\ \textit{Now } U \textit{ is the transpose of } L. \end{matrix}$$

When A is symmetric, the usual form $A = LDU$ becomes $A = \boldsymbol{LDL^T}$. The final U (with 1's on the diagonal) is the transpose of L (also with 1's on the diagonal). The diagonal matrix D containing the pivots is symmetric by itself.

If $A = A^T$ is factored into LDU with no row exchanges, then U is exactly L^T.

$$\textit{The symmetric factorization of a symmetric matrix is } A = LDL^T.$$

Notice that the transpose of LDL^T is automatically $(L^T)^TD^TL^T$ which is LDL^T again. The work of elimination is cut in half, from $n^3/3$ multiplications to $n^3/6$. The storage is also cut essentially in half. We only keep L and D, not U which is just L^T.

Permutation Matrices

The transpose plays a special role for a *permutation matrix*. This matrix P has a single "1" in every row and every column. Then P^{T} is also a permutation matrix—maybe the same or maybe different. Any product $P_1 P_2$ is again a permutation matrix. We now create every P from the identity matrix, by reordering the rows of I.

The simplest permutation matrix is $P = I$ (*no exchanges*). The next simplest are the row exchanges P_{ij}. Those are constructed by exchanging two rows i and j of I. Other permutations reorder more rows. By doing all possible row exchanges to I, we get all possible permutation matrices:

DEFINITION ***A permutation matrix P has the rows of the identity I in any order.***

Example 5 There are six 3 by 3 permutation matrices. Here they are without the zeros:

$$I = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \quad P_{21} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} \quad P_{32}P_{21} = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$$

$$P_{31} = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \quad P_{32} = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} \quad P_{21}P_{32} = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}.$$

There are $n!$ permutation matrices of order n. The symbol $n!$ means "n factorial," the product of the numbers $(1)(2)\cdots(n)$. Thus $3! = (1)(2)(3)$ which is 6. There will be 24 permutation matrices of order $n = 4$. And 120 permutations of order 5.

There are only two permutation matrices of order 2, namely $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$.

Important: P^{-1} is also a permutation matrix. Among the six 3 by 3 P's displayed above, the four matrices on the left are their own inverses. The two matrices on the right are inverses of each other. In all cases, a single row exchange is its own inverse. If we repeat the exchange we are back to I. But for $P_{32}P_{21}$, the inverses go in opposite order as always. The inverse is $P_{21}P_{32}$.

More important: $\boldsymbol{P^{-1}}$ ***is always the same as*** $\boldsymbol{P^{\mathrm{T}}}$. The two matrices on the right are transposes—and inverses—of each other. When we multiply PP^{T}, the "1" in the first row of P hits the "1" in the first column of P^{T} (since the first row of P is the first column of P^{T}). It misses the ones in all the other columns. So $PP^{\mathrm{T}} = I$.

Another proof of $P^{\mathrm{T}} = P^{-1}$ looks at P as a product of row exchanges. Every row exchange is its own transpose and its own inverse. P^{T} and P^{-1} both come from the product of row exchanges *in reverse order*. So P^{T} and P^{-1} are the same.

Symmetric matrices led to $A = LDL^{\mathrm{T}}$. Now permutations lead to $PA = LU$.

Problem Set 2.7

Questions 1–7 are about the rules for transpose matrices.

1 Find A^T and A^{-1} and $(A^{-1})^T$ and $(A^T)^{-1}$ for

$$A = \begin{bmatrix} 1 & 0 \\ 9 & 3 \end{bmatrix} \quad \text{and also} \quad A = \begin{bmatrix} 1 & c \\ c & 0 \end{bmatrix}.$$

2 Verify that $(AB)^T$ equals $B^T A^T$ but those are different from $A^T B^T$:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \quad AB = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}.$$

In case $AB = BA$ (not generally true!) how do you prove that $B^T A^T = A^T B^T$?

3 (a) The matrix $((AB)^{-1})^T$ comes from $(A^{-1})^T$ and $(B^{-1})^T$. *In what order?*

(b) If U is upper triangular then $(U^{-1})^T$ is ____ triangular.

4 Show that $A^2 = 0$ is possible but $A^T A = 0$ is not possible (unless $A =$ zero matrix).

5 (a) The row vector x^T times A times the column y produces what number?

$$x^T A y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \underline{\quad}.$$

(b) This is the row $x^T A =$ ____ times the column $y = (0, 1, 0)$.

(c) This is the row $x^T = [0 \ \ 1]$ times the column $Ay =$ ____.

6 The transpose of a block matrix $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is $M^T =$ ____. Test an example. Under what conditions on A, B, C, D is the block matrix symmetric?

7 True or false:

(a) The block matrix $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ is automatically symmetric.

(b) If A and B are symmetric then their product AB is symmetric.

(c) If A is not symmetric then A^{-1} is not symmetric.

(d) When A, B, C are symmetric, the transpose of ABC is CBA.

Questions 8–15 are about permutation matrices.

8 Why are there $n!$ permutation matrices of order n?

9 If P_1 and P_2 are permutation matrices, so is $P_1 P_2$. This still has the rows of I in some order. Give examples with $P_1 P_2 \neq P_2 P_1$ and $P_3 P_4 = P_4 P_3$.

10 There are 12 "*even*" permutations of $(1, 2, 3, 4)$, with an *even number of exchanges*. Two of them are $(1, 2, 3, 4)$ with no exchanges and $(4, 3, 2, 1)$ with two exchanges. List the other ten. Instead of writing each 4 by 4 matrix, just order the numbers.

11 Which permutation makes PA upper triangular? Which permutations make $P_1 A P_2$ lower triangular? ***Multiplying A on the right by P_2 exchanges the*** ____ *of A.*