(b) If the third equation has zero as its second coefficient (it contains $0v$) then no multiple of equation 2 will be subtracted from equation 3.

(c) If the third equation contains $0u$ *and* $0v$, then no multiple of equation 1 or equation 2 will be subtracted from equation 3.

29. (Very optional) Normally the multiplication of two complex numbers

$$(a+ib)(c+id) = (ac-bd)+i(bc+ad)$$

involves the four separate multiplications ac, bd, be, ad. Ignoring i, can you compute $ac-bd$ and $bc+ad$ with only three multiplications? (You may do additions, such as forming $a+b$ before multiplying, without any penalty.)

30. Use elimination to solve

$$\begin{array}{rcrcrcr} u & + & v & + & w & = & 6 \\ u & + & 2v & + & 2w & = & 11 \\ 2u & + & 3v & - & 4w & = & 3 \end{array} \quad \text{and} \quad \begin{array}{rcrcrcr} u & + & v & + & w & = & 7 \\ u & + & 2v & + & 2w & = & 10 \\ 2u & + & 3v & - & 4w & = & 3. \end{array}$$

31. For which three numbers a will elimination fail to give three pivots?

$$\begin{aligned} ax+2y+3z &= b_1 \\ ax+ay+4z &= b_2 \\ ax+ay+az &= b_3. \end{aligned}$$

32. Find experimentally the average size (absolute value) of the first and second and third pivots for MATLAB's lu(rand(3,3)). The average of the first pivot from abs(A(1,1)) should be 0.5.

1.4 Matrix Notation and Matrix Multiplication

With our 3 by 3 example, we are able to write out all the equations in full. We can list the elimination steps, which subtract a multiple of one equation from another and reach a triangular matrix. For a large system, this way of keeping track of elimination would be hopeless; a much more concise record is needed.

We now introduce **matrix notation** to describe the original system, and **matrix multiplication** to describe the operations that make it simpler. Notice that three different types of quantities appear in our example:

$$\begin{array}{lrcrcrcr} \textbf{Nine coefficients} & 2u & + & v & + & w & = & 5 \\ \textbf{Three unknowns} & 4u & - & 6v & & & = & -2 \\ \textbf{Three right-hand sides} & -2u & + & 7v & + & 2w & = & 9 \end{array} \qquad (1)$$

On the right-hand side is the column vector b. On the left-hand side are the unknowns u, v, w. Also on the left-hand side are nine coefficients (one of which happens to be zero). It is natural to represent the three unknowns by a vector:

$$\text{The unknown is } x = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \text{The solution is } x = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

The nine coefficients fall into three rows and three columns, producing a 3 ***by*** 3 ***matrix***:

$$\textbf{Coefficient matrix} \qquad A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}.$$

A is a *square* matrix, because the number of equations equals the number of unknowns. If there are n equations in n unknowns, we have a square n by n matrix. More generally, we might have m equations and n unknowns. Then A is *rectangular*, with m rows and n columns. It will be an "m by n matrix."

Matrices are added to each other, or multiplied by numerical constants, exactly as vectors are—one entry at a time. In fact we may regard vectors as special cases of matrices; *they are matrices with only one column*. As with vectors, two matrices can be added only if they have the same shape:

$$\begin{array}{l} \textbf{Addition } A+B \\ \textbf{Multiplication } 2A \end{array} \qquad \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 1 \\ 1 & 6 \end{bmatrix} \qquad 2\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & 0 \\ 0 & 8 \end{bmatrix}.$$

Multiplication of a Matrix and a Vector

We want to rewrite the three equations with three unknowns u, v, w in the simplified matrix form $Ax = b$. Written out in full, matrix times vector equals vector:

$$\textbf{Matrix form } Ax = b \qquad \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}. \tag{2}$$

The right-hand side b is the column vector of "inhomogeneous terms." **The left-hand side is A times x.** This multiplication will be defined *exactly so as to reproduce the original system*. The first component of Ax comes from "multiplying" the first row of A into the column vector x:

$$\textbf{Row times column} \qquad \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u+v+w \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}. \tag{3}$$

The second component of the product Ax is $4u-6v+0w$, from the second row of A. The matrix equation $Ax=b$ is equivalent to the three simultaneous equations in equation (1).

Row times column is fundamental to all matrix multiplications. From two vectors it produces a single number. This number is called the ***inner product*** of the two vectors. In other words, the product of a 1 by n matrix (a *row vector*) and an n by 1 matrix (a *column vector*) is a 1 by 1 matrix:

$$\textbf{Inner product} \qquad \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2\cdot 1+1\cdot 1+1\cdot 2 \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}.$$

This confirms that the proposed solution $x=(1,1,2)$ does satisfy the first equation.

There are two ways to multiply a matrix A ***and a vector*** x. One way is *a row at a time*, Each row of A combines with x to give a component of Ax. There are three inner products when A has three rows:

$$Ax \textbf{ by rows} \qquad \begin{bmatrix} 1 & 1 & 6 \\ 3 & 0 & 1 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1\cdot 2+1\cdot 5+6\cdot 0 \\ 3\cdot 2+0\cdot 5+3\cdot 0 \\ 1\cdot 2+1\cdot 5+4\cdot 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \\ 7 \end{bmatrix}. \tag{4}$$

That is how Ax is usually explained, but the second way is equally important. In fact it is more important! It does the multiplication *a column at a time*. The product Ax is found all at once, as ***a combination of the three columns of*** A:

$$Ax \textbf{ by columns} \qquad 2\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \\ 7 \end{bmatrix}. \tag{5}$$

The answer is twice column 1 plus 5 times column 2. It corresponds to the "column picture" of linear equations. If the right-hand side b has components 7, 6, 7, then the solution has components 2, 5, 0. Of course the row picture agrees with that (and we eventually have to do the same multiplications).

The column rule will be used over and over, and we repeat it for emphasis:

> **1A** Every product Ax can be found using whole columns as in equation (5). Therefore Ax is ***a combination of the columns of*** A. The coefficients are the components of x.

To multiply A times x in n dimensions, we need a notation for the individual entries in A. *The entry in the ith row and jth column is always denoted by* a_{ij}. The first subscript gives the row number, and the second subscript indicates the column. (In equation (4), a_{21} is 3 and a_{13} is 6.) If A is an m by n matrix, then the index i goes from 1 to m—there are m rows—and the index j goes from 1 to n. Altogether the matrix has mn entries, and a_{mn} is in the lower right corner.

One subscript is enough for a vector. The jth component of x is denoted by x_j. (The multiplication above had $x_1 = 2$, $x_2 = 5$, $x_3 = 0$.) Normally x is written as a column vector—like an n by 1 matrix. But sometimes it is printed on a line, as in $x = (2,5,0)$. The parentheses and commas emphasize that it is not a 1 by 3 matrix. It is a column vector, and it is just temporarily lying down.

To describe the product Ax, we use the "*sigma*" symbol Σ for summation:

$$\textbf{Sigma notation} \qquad \text{The } i\text{th component of } Ax \text{ is } \sum_{j=1}^{n} a_{ij}x_j.$$

This sum takes us along the ith row of A. The column index j takes each value from 1 to n and we add up the results—the sum is $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$.

We see again that the length of the rows (the number of columns in A) must match the length of x. **An m by n matrix multiplies an n-dimensional vector** (and produces an m-dimensional vector). Summations are simpler than writing everything out in full, but matrix notation is better. (Einstein used "tensor notation," in which a repeated index automatically means summation. He wrote $a_{ij}x_j$ or even $a_i^j x_j$, without the Σ. Not being Einstein, we keep the Σ.)

The Matrix Form of One Elimination Step

So far we have a convenient shorthand $Ax = b$ for the original system of equations. What about the operations that are carried out during elimination? In our example, the first step subtracted 2 times the first equation from the second. On the right-hand side, 2 times the first component of b was subtracted from the second component. *The same result is achieved if we multiply b by this elementary matrix* (or *elimination matrix*):

$$\textbf{Elementary matrix} \qquad E = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This is verified just by obeying the rule for multiplying a matrix and a vector:

$$Eb = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = \begin{bmatrix} 5 \\ \mathbf{-12} \\ 9 \end{bmatrix}.$$

The components 5 and 9 stay the same (because of the 1, 0, 0 and 0, 0, 1 in the rows of E). The new second component -12 appeared after the first elimination step.

It is easy to describe the matrices like E, which carry out the separate elimination steps. We also notice the "identity matrix," which does nothing at all.

1B The ***identity matrix*** I, with 1s on the diagonal and 0s everywhere else, leaves every vector unchanged. The ***elementary matrix*** E_{ij} subtracts ℓ times

row j from row i. This E_{ij} includes $-\ell$ in row i, column j.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ has } Ib = b \qquad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell & 0 & 1 \end{bmatrix} \text{ has } E_{31}b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 - \ell b_1 \end{bmatrix}.$$

$Ib = b$ is the matrix analogue of multiplying by 1. A typical elimination step multiplies by E_{31}. The important question is: What happens to A on the left-hand side?

To maintain equality, we must apply the same operation to both sides of $Ax = b$. In other words, we must also multiply the vector Ax by the matrix E. Our original matrix E subtracts 2 times the first component from the second, After this step the new and simpler system (equivalent to the old) is just $E(Ax) = Eb$. It is simpler because of the zero that was created below the first pivot. It is equivalent because we can recover the original system (by adding 2 times the first equation back to the second). So the two systems have exactly the same solution x.

Matrix Multiplication

Now we come to the most important question: *How do we multiply two matrices*? There is a partial clue from Gaussian elimination: We know the original coefficient matrix A, we know the elimination matrix E, and we know the result EA after the elimination step. We hope and expect that

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ times } A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \text{ gives } \quad EA = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}.$$

Twice the first row of A has been subtracted from the second row. Matrix multiplication is consistent with the row operations of elimination. We can write the result either as $E(Ax) = Eb$, applying E to both sides of our equation, or as $(EA)x = Eb$. The matrix EA is constructed exactly so that these equations agree, and we don't need parentheses:

Matrix multiplication (EA times x) equals (E times Ax). We just write EAx.

This is the whole point of an "associative law" like $2 \times (3 \times 4) = (2 \times 3) \times 4$. The law seems so obvious that it is hard to imagine it could be false. But the same could be said of the "commutative law" $2 \times 3 = 3 \times 2$—and for matrices EA is not AE.

There is another requirement on matrix multiplication. We know how to multiply Ax, a matrix and a vector. The new definition should be consistent with that one. When a matrix B contains only a single column x, the matrix-matrix product AB should be identical with the matrix-vector product Ax. *More than that*: When B contains several

columns b_1, b_2, b_3, the columns of AB should be Ab_1, Ab_2, Ab_3!

$$\textbf{Multiplication by columns} \qquad AB = A\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} Ab_1 \\ Ab_2 \\ Ab_3 \end{bmatrix}.$$

Our first requirement had to do with rows, and this one is concerned with columns. A third approach is to describe each individual entry in AB and hope for the best. In fact, there is only one possible rule, and I am not sure who discovered it. It makes everything work. It does not allow us to multiply every pair of matrices. If they are square, they must have the same size. If they are rectangular, they must *not* have the same shape; ***the number of columns in*** A ***has to equal the number of rows in*** B. Then A can be multiplied into each column of B.

If A is m by n, and B is n by p, then multiplication is possible. *The product AB will be m by p.* We now find the entry in row i and column j of AB.

1C The i, j entry of AB is the inner product of the ith row of A and the jth column of B. In Figure 1.7, the 3, 2 entry of AB comes from row 3 and column 2:

$$(AB)_{32} = a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} + a_{34}b_{42}. \tag{6}$$

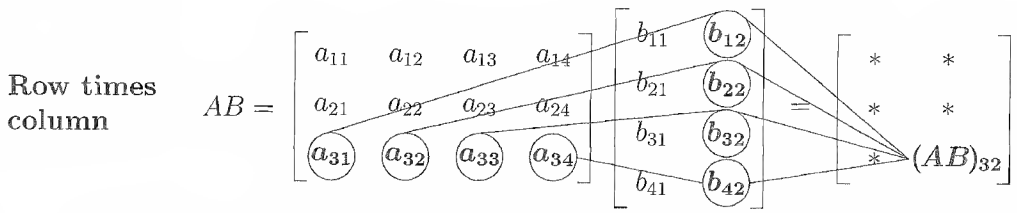

Figure 1.7: A 3 by 4 matrix A times a 4 by 2 matrix B is a 3 by 2 matrix AB.

Note. We write AB when the matrices have nothing special to do with elimination. Our earlier example was EA, because of the elementary matrix E. Later we have PA, or LU, or even LDU. The rule for matrix multiplication stays the same.

Example 1.

$$AB = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 5 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 17 & 1 & 0 \\ 4 & 8 & 0 \end{bmatrix}.$$

The entry 17 is $(2)(1)+(3)(5)$, the inner product of the first row of A and first column of B. The entry 8 is $(4)(2)+(0)(-1)$, from the second row and second column.

The third column is zero in B, so it is zero in AB. B consists of three columns side by side, and A multiplies each column separately. ***Every column of*** AB ***is a combination of the columns of*** A. Just as in a matrix-vector multiplication, the columns of A are multiplied by the entries in B.

Example 2.

$$\textbf{Row exchange matrix} \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 7 & 8 \\ 2 & 3 \end{bmatrix}.$$

Example 3. The 1s in the identity matrix I leave every matrix unchanged:

$$\textbf{Identity matrix} \qquad IA = A \quad \text{and} \quad BI = B.$$

Important: The multiplication AB can also be done *a row at a time*. In Example 1, the first row of AB uses the numbers 2 and 3 from the first row of A. Those numbers give $2[\text{row } 1] + 3[\text{row } 2] = [17 \;\; 1 \;\; 0]$. Exactly as in elimination, where all this started, each row of AB is a ***combination of the rows of*** B.

We summarize these three different ways to look at matrix multiplication.

1D

(i) Each entry of AB is the product of a ***row*** and a ***column***:

$$(AB)_{ij} = (\text{row } i \text{ of } A) \text{ times } (\text{column } j \text{ of } B)$$

(ii) Each column of AB is the product of a matrix and a column:

$$\text{column } j \text{ of } AB = A \text{ times } (\text{column } j \text{ of } B)$$

(iii) Each row of AB is the product of a row and a matrix:

$$\text{row } i \text{ of } AB = (\text{row } i \text{ of } A) \text{ times } B.$$

This leads hack to a key property of matrix multiplication. Suppose the shapes of three matrices A, B, C (possibly rectangular) permit them to be multiplied. The rows in A and B multiply the columns in B and C. Then the key property is this:

1E Matrix multiplication is associative: $(AB)C = A(BC)$. Just write ABC.

AB times C equals A times BC. If C happens to be just a vector (a matrix with only one column) this is the requirement $(EA)x = E(Ax)$ mentioned earlier. It is the whole basis for the laws of matrix multiplication. And if C has several columns, we have only to think of them placed side by side, and apply the same rule several times. Parentheses are not needed when we multiply several matrices.

There are two more properties to mention—one property that matrix multiplication has, and another which it *does not have*. The property that it does possess is:

1F Matrix operations are distributive:

$$A(B+C) = AB + AC \quad \text{and} \quad (B+C)D = BD + CD.$$

Of course the shapes of these matrices must match properly—B and C have the same shape, so they can be added, and A and D are the right size for premultiplication and postmultiplication. The proof of this law is too boring for words.

The property that fails to hold is a little more interesting:

1G Matrix multiplication is not commutative: Usually $FE \neq EF$.

Example 4. Suppose E subtracts twice the first equation from the second. Suppose F is the matrix for the next step, *to add row* 1 *to row* 3:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix}.$$

These two matrices do commute and the product does both steps at once:

$$EF = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix} = FE.$$

In either order, EF or FE, this changes rows 2 and 3 using row 1.

Example 5. Suppose E is the same but G *adds row* 2 *to row* 3. Now the order makes a difference. When we apply E and then G, the second row is altered *before* it affects the third. If E comes *after* G, then the third equation feels no effect from the first. You will see a zero in the $(3,1)$ entry of EG, where there is a -2 in GE:

$$GE = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -\mathbf{2} & 1 & 1 \end{bmatrix} \quad \text{but} \quad EG = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \mathbf{0} & 1 & 1 \end{bmatrix}.$$

Thus $EG \neq GE$. A random example would show the same thing—most matrices don't commute. Here the matrices have meaning. There was a reason for $EF = FE$, and a reason for $EG \neq GE$. It is worth taking one more step, to see what happens with *all three elimination matrices at once*:

$$GFE = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -\mathbf{1} & 1 & 1 \end{bmatrix} \quad \text{and} \quad EFG = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -\mathbf{1} & 1 & 1 \end{bmatrix}.$$

The product GFE is the true order of elimination. *It is the matrix that takes the original A to the upper triangular U.* We will see it again in the next section.

The other matrix EFG is nicer. In that order, the numbers -2 from E and 1 from F and G were not disturbed. They went straight into the product. It is the wrong order for elimination. But fortunately *it is the right order for reversing the elimination steps*—which also comes in the next section.

Notice that the product of lower triangular matrices is again lower triangular.