LINEAR ALGEBRA. LECTURE 1

The Geometry of Linear Equations

The fundamental problem of linear algebra is how to solve a system of linear equations. Let's start with a case when we have the same number of equations and unknowns. That's the tipical case. First, we describe the Row. That's one equation at a time. The second is the Column. And both are the rows and columns form a matrix. The matrix is the algebra way to look at the problem of solving a system of linear equations. Let's take an example with two equations and two unknowns.

$$\begin{cases} 2x - y = 1 \\ x + y = 5 \end{cases} \tag{1.1}$$

Now let's see that's the matrix and that's the coefficient of the matrix. The matrix is just a rectangular array of numbers. Here there are two rows and two columns:

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{matrix} (2, -1) \\ (1, \;\; 1) \end{matrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{1.2}$$

That's the matrix A. We also have the right-hand vector $b = (1, 5)$ and the vector of two unknowns $X = (x, y)$. So finaly we can rewrite (1.1) in the vector-matrix form:

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix} \tag{1.3}$$

or $A \cdot X = b$. Here the vector of unknowns has only got two unknowns. Later we'll have any number of unknowns.

We can look at that system *by rows* or *by columns*. We want to see them both. The first approach concentrates on the separate equations (the *rows*). That is the most familiar, and in two dimensions we can do it quickly. The equation $2x - y = 1$ is represented by a *straight line* in the *x*-*y* plane. The line goes through the points $x = 1$, $y = 1$ and $x = 1/2$, $y = 0$ (and also through (2;3) and all intermediate points). The second equation $x + y = 5$ produces a second line (Figure 1.a). Its slope is $dy/dx = -1$ and it crosses the first line at the solution. The point of intersection lies on both lines. It is the only solution to both equations. That point $x = 2$ and $y = 3$ can also be found by "elimination."

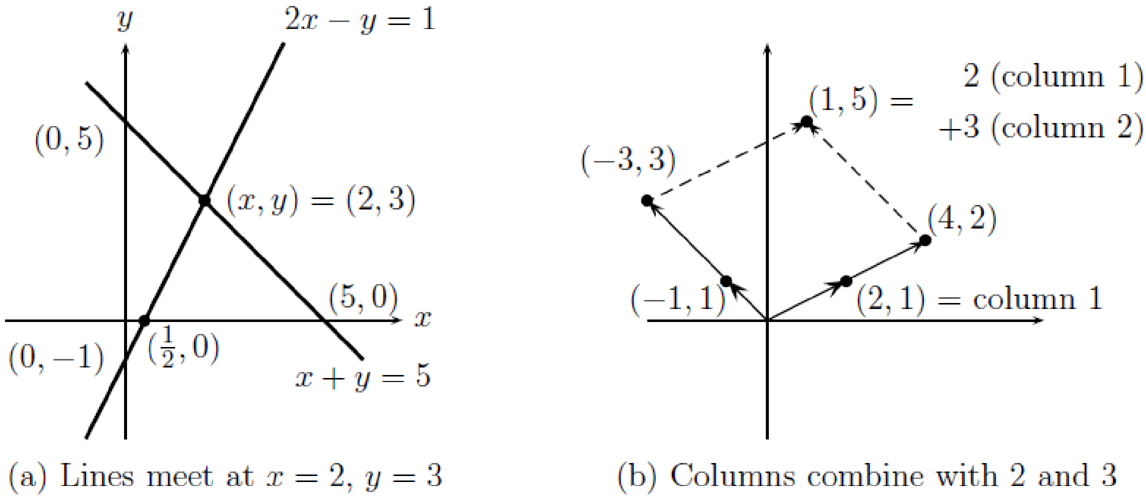

Figure 1. Row picture (two lines) and column picture (combine columns).

The second approach looks at the *columns* of the linear system. The two separate equations are really *one vector equation*:

$$x\begin{bmatrix}2\\1\end{bmatrix}+y\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1\\5\end{bmatrix} \tag{1.4}$$

The problem is *to find the* ***linear combination*** *of the column vectors on the left side that produces the vector on the right side*. Those vectors (2, 1) and (-1, 1) are represented by the bold lines in Figure 1.b. The unknowns are the numbers x and y that multiply the column vectors. The whole idea can be seen in that figure, where 2 times column 1 is added to 3 times column 2. Geometrically this produces a famous parallelogram. Algebraically it produces the correct vector (1, 5), on the right side of our equations. The column picture confirms that $x = 2$ and $y = 3$.

This idea of linear combination is crucial. What are all the combinations? If we take all the *x-s* and all the *y-s*, all the combinations, what would be the results? And, actually, the result would be that we could get any right-hand side in (1.4). The combinations of *the column vectors* **would fill the whole plane**. This idea of what linear combination *the column vectors* gives the **any right-hand *b*** that's going to be basic.

More time could be spent on that example, but I would rather move forward to $n = 3$. Three equations are still manageable, and they have much more variety:

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases} \tag{1.5}$$

Just let's write the matrix form (1.5), because that is easy. It is a three by three matrix: three equations and three unknowns:

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} \tag{1.6}$$

Again we can study the rows or the columns, and we start with the rows. Each equation describes a *plane* in three dimensions. The first plane is $2u+v+w=5$, and it is sketched in Figure 2. It contains the points (2.5, 0, 0) and (0, 5, 0) and (0, 0, 5). It is determined by any three of its points—provided if they do not lie on a line.

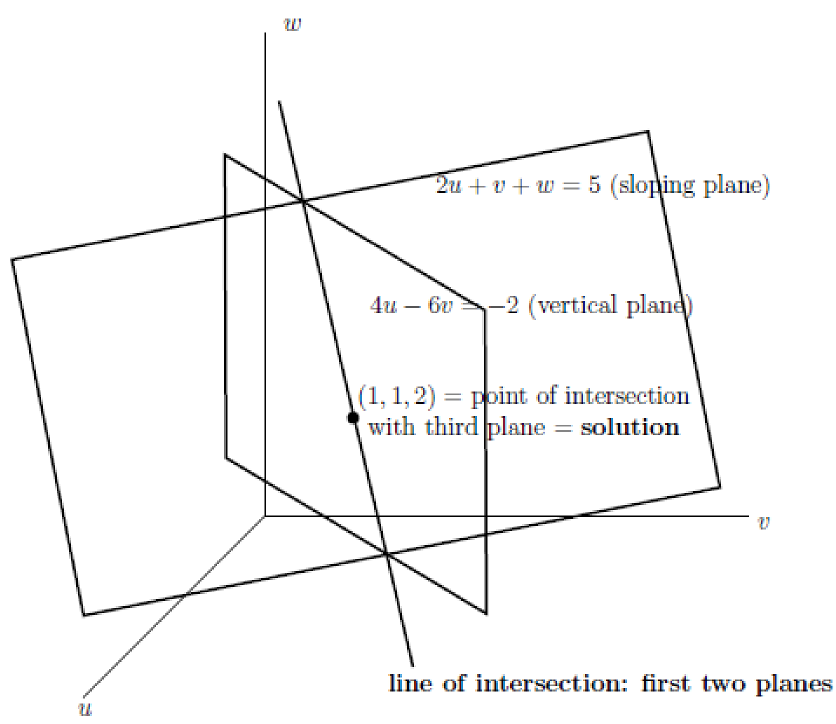

Figure 2. The row picture: three intersecting planes from three linear equations.

Changing *5* to *10*, the plane $2u+v+w=10$ would be parallel to this one. It contains (5;0;0) and (0,10,0) and (0,0,10), twice as far from the origin—which is the center point $u=0$, $v=0$, $w=0$. Changing the right side moves the plane parallel to itself, and the plane $2u+v+w=0$ goes through the origin.

The second plane is $4u-6v=-2$. It is drawn vertically, because w can take any value. The coefficient of w is zero, but this remains a plane in 3-space. (The equation $4u=3$, or even the extreme case $u=0$, would still describe a plane.) The figure shows the intersection of the second plane with the first. That intersection is a line. In three dimensions a line requires two equations; in n dimensions it will require $n-1$.

Finally the third plane intersects this line in a point. The plane (not drawn) represents the third equation $-2u+7v+2w=9$, and it crosses the line at $u=1$, $v=1$, $w=2$. That triple intersection point (1,1,2) solves the linear system. Now you can see that this row picture is getting a little hard to see. The row picture was a cinch when we looked at two lines meeting. When we look at three planes meeting, it's not so clear and in four dimensions probably a little less clear.

We turn to the columns. This time the vector equation (the same equation as (1.5)-(1.6)) is

$$u\begin{bmatrix}2\\4\\-2\end{bmatrix}+v\begin{bmatrix}1\\-6\\7\end{bmatrix}+w\begin{bmatrix}1\\0\\2\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}=b \tag{1.7}$$

Those are *three-dimensional column vectors*. The vector b is identified with the point whose coordinates are (5,-2,9). Every point in three-dimensional space is matched to a vector, and vice versa. That was the idea of Descartes, who turned geometry into algebra by working with the coordinates of the point. We can write the vector in a column, or we can list its components as b = (5,-2,9), or we can represent it geometrically by an arrow from the origin. You can choose *the arrow*, or *the point*, or *the three numbers*. In six dimensions it is probably easiest to choose the six numbers.

We use parentheses and commas when the components are listed horizontally, and square brackets (with no commas) when a column vector is printed vertically. What really matters is *addition of vectors* and *multiplication by a scalar* (a number). In Figure 3a you see a vector addition, component by component:

$$\begin{bmatrix}5\\0\\0\end{bmatrix}+\begin{bmatrix}0\\-2\\0\end{bmatrix}+\begin{bmatrix}0\\0\\9\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}=b \tag{1.8}$$

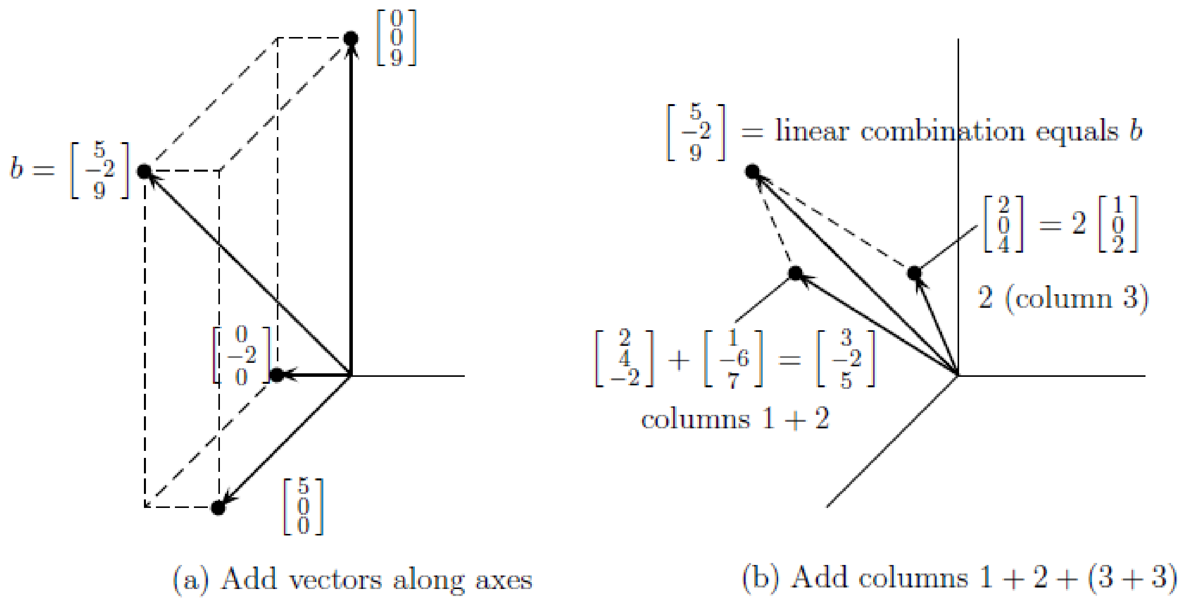

Figure 3. The column picture: linear combination of columns equals *b*.

In the right-hand figure is one of the central ideas of linear algebra. It uses *both* of the basic operations; vectors are *multiplied by numbers:* 2(1,0,2)=(2,0,4) *and then added*. The result is called a *linear combination*, and this combination solves our equation:

$$1\cdot\begin{bmatrix}2\\4\\-2\end{bmatrix}+1\cdot\begin{bmatrix}1\\-6\\7\end{bmatrix}+2\cdot\begin{bmatrix}1\\0\\2\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}=b \tag{1.9}$$

Equation (1.7) asked for multipliers *u*, *v*, *w* that produce the right side *b*. Those numbers are *u=1*, *v=1*, *w=2*. They give the correct combination of the columns. They also gave the point (1,1,2) in the row picture (where the three planes intersect).

Now let's think about all possible right-hand sides. Can we solve these equations for every right-hand side? Can we solve $A\cdot X=b$ for every *b*? Is there a solution? And, actually, the answer for this matrix will be yes. **For this matrix *A* - for these columns, the answer is yes**. This matrix - that we chose for an example is **a good matrix**. ***A* is non-singular and invertible**

matrix. Those will be the matrices that we like best. There could be other and we will see other matrices where the answer becomes, no.

Suppose we are again in three dimensions, and the three planes in the row picture *do not intersect*. What can go wrong? One possibility is that two planes may be parallel. The equations $2u+v+w=5$ and $4u+2v+2w=11$ are inconsistent—and parallel planes give no solution (Figure 4a).

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ -2 & 7 & 2 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ 9 \end{bmatrix} \tag{1.10}$$

In (1.10) we see that second row equal the first row multiplied by two. That makes parallel these two planes. In two dimensions, parallel lines are the only possibility for breakdown. But three planes in three dimensions can be in trouble without being parallel.

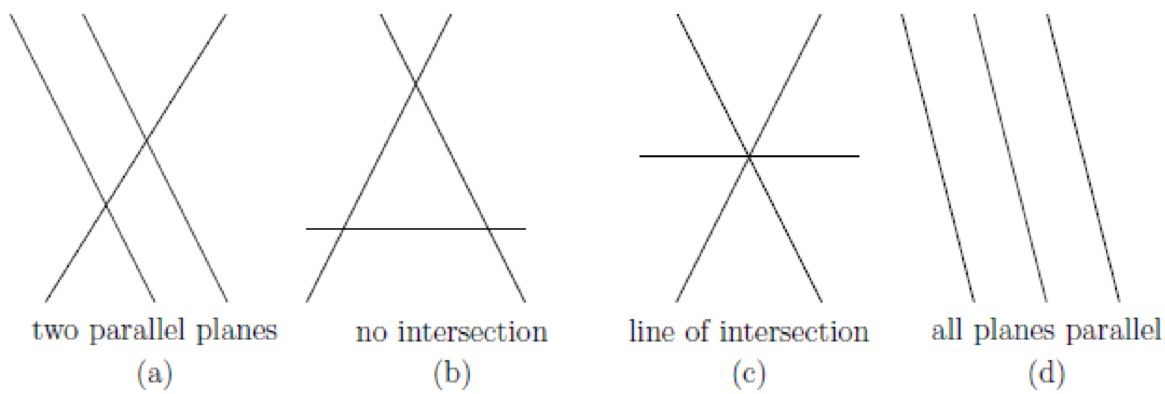

Figure 4. Singular cases: no solution for (a), (b), or (d), an infinity of solutions for (c).

The most common difficulty is shown in Figure 4b. From the end view the planes form a triangle. Every pair of planes intersects in a line, and those lines are parallel. The third plane is not parallel to the other planes, but it is parallel to their line of intersection. This corresponds to a singular system with *b*=(2;5;6):

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \\ 3 & 1 & 4 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} \tag{1.11}$$

The first two left sides add up to the third. On the right side that fails: $2+5 \neq 6$. Equation 1 plus equation 2 minus equation 3 is the impossible statement 0=1. Thus the equations are *inconsistent*.

Another singular system, close to this one, has an infinity of solutions. When the 6 in the last equation becomes 7, the three equations combine to give 0 = 0. Now the third equation is the sum of the first two. In that case the three planes have a whole *line in common* (Figure 4c). Changing the right sides will move the planes in Figure 4b parallel to themselves, and for b=(2;5;7) the figure is suddenly different. The lowest plane moved up to meet the others, and there is a line of solutions. Problem 4c is still singular, but now it suffers from *too many solutions* instead of too few.

The extreme case is three parallel planes. For most right sides there is no solution (Figure 4d). For special right sides (like b=(0;0;0)!) there is a whole plane of solutions—because the three parallel planes move over to become the same.

What happens to the *column picture* when the system is singular? It has to go wrong; the question is how. There are still three columns on the left side of the equations, and we try to combine them to produce b. Stay with equation (1.11):

$$u\begin{bmatrix}1\\2\\3\end{bmatrix}+v\begin{bmatrix}1\\0\\1\end{bmatrix}+w\begin{bmatrix}1\\3\\4\end{bmatrix}=b \tag{1.12}$$

For b=(2,5,7) this was possible; for b=(2,5,6) it was not. The reason is that *those three columns lie in a plane*. Then every combination is also in the plane (which goes through the origin). If the vector b is not in that plane, no solution is possible (Figure 5). That is by far the most likely event; a singular system generally has no solution. But there is a chance that b does lie in the plane of the columns. In that case there are too many solutions; the three columns can be combined in *infinitely many ways* to produce b. That column picture in Figure 5b corresponds to the row picture in Figure 4c.

How do we know that the three columns lie in the same plane? One answer is to find a combination of the columns that adds to zero. After some calculation, it is u=3, v=1, w=-2. Three times column 1 equals column 2 plus twice column 3. Column 1 is in the plane of columns 2 and 3. Only two columns are independent.

The vector b=(2;5;7) is in that plane of the columns - it is column 1 plus column 3 - so (1,0,1) is a solution. *We can add a multiple of the combination* (3,-1;-2) *that gives* b=0. So there is a whole line of solutions—as we know from the row picture. The truth is that we *knew* the columns would combine to give zero, because the rows did. That is a fact of mathematics, not of computation—and it remains true in dimension *n*. ***If the n planes have no point in common, or infinitely many points, then the n columns lie in the same plane*.**

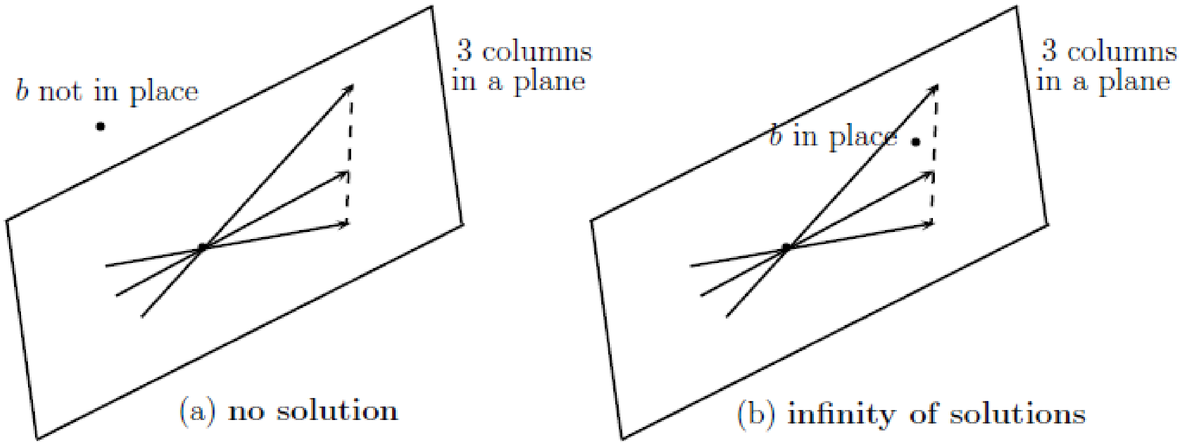

Figure 5. Singular cases: *b* outside or inside the plane with all three columns.

Gaussian Elimination

We begin wih a system of equations that's going to be our example to work with:

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases} \tag{1.13}$$

The problem is to find the unknown values of *u*, *v*, and *w*, and we shall apply Gaussian Elimination The method starts by subtracting multiples of the first equation from the other equations. The goal is to eliminate u from the last two equations. This requires that we:

1) subtract 2 times the first equation from the second;

2) subtract -1 times the first equation from the third and get the **equivalent system**.

$$\begin{cases} 2u + v + w = 5 \\ -8v - 2w = -12 \\ 8v + 3w = 14 \end{cases} \tag{1.14}$$

The coefficient 2 is the first factor. Elimination is constantly dividing the factor into the numbers underneath it, to find out the right multipliers. Our purpose is to knock out the *u* part of equations two and three.

The factor for the second stage of elimination is -8. We now ignore the first equation. A multiple of the second equation will be subtracted from the remaining equations (in this case there is only the third one) so as to eliminate *v*. We add the second equation to the third or, in other words, we

3) subtract -1 times the second equation from the third.

The elimination process in the “forward” direction is complete and we have the **triangular system**:

$$\begin{cases} 2u + v + w = 5 \\ 8v + 2w = 12 \\ w = 2 \end{cases} \tag{1.15}$$

This system can be solved backward, bottom to top. The last equation gives $w = 2$. Substituting into the second equation, we find $v = 1$. Then the first equation gives $u = 1$. This process is called back-substitution.

To repeat: Forward elimination produced the factors 2, -1, 1. It subtracted multiples of each row from the rows beneath, It reached the “triangular” system (3), which is can be solved now in reverse order. Just substitute each newly computed value into the equations from down to up.

In a larger problem, forward elimination takes most of the effort. We use multiples of the first equation to produce zeros below the first factor. Then the second column is cleared out below the second factor. The forward step is finished when the system is triangular; equation n contains only the last unknown multiplied by the last factor. Back substitution yields the complete solution in the opposite order — beginning with the last unknown, then solving for the next to last, and eventually for the first.

By definition, **factors cannot be zero**. We need to divide by them.

The Breakdown of Elimination

Under what circumstances could the process break down? Something must go wrong in the singular case, and something might go wrong in the nonsingular case. The possibility of breakdown makes more clear the method itself.

The answer is: With a full set of *n* factors, there is only one solution. The system is non singular, and it is solved by forward elimination and back-substitution. But *if a zero appears* in a factor position, elimination has to stop—either temporarily or permanently. The system might or might not be singular.

If the first coefficient is zero, in the upper left corner, the elimination of u from the other equations will be impossible. The same is true at every intermediate stage. Notice that a zero can appear in a factor position, even if the original coefficient in that place was not zero. Roughly speaking, *we do not know whether a zero will appear until we try*, by actually going through the elimination process.

In many cases this problem can be resolved, and elimination can proceed. Such a system still counts as nonsingular; it is only the algorithm that needs repair. In other cases a breakdown is unavoidable. Those incurable systems are singular, they have no solution or else infinitely many, and a full set of factors cannot be found.

Example 1. Nonsingular (cured by exchanging equations 2 and 3)

$$\begin{cases} u+v+w= \\ 2u+2v+5w= \\ 4u+6v+8w= \end{cases} \Rightarrow \begin{cases} u+v+w= \\ 3w= \\ 2v+4w= \end{cases} \Rightarrow \begin{cases} u+v+w= \\ 2v+4w= \\ 3w= \end{cases}$$

The system is now triangular, and back-substitution will solve it.

Example 2. Singular (incurable)

$$\begin{cases} u+v+w= \\ 2u+2v+5w= \\ 4u+4v+8w= \end{cases} \Rightarrow \begin{cases} u+v+w= \\ 3w= \\ 4w= \end{cases}$$

There is no exchange of equations that can avoid zero in the second pivot position. The equations themselves may be solvable or unsolvable. If the last two equations are $3w=6$ and $4w=7$, there is no solution. If those two equations happen to be consistent as in $3w=6$ and $4w=8$ then this singular case has an infinity of solutions. We know that $w=2$, but the first equation cannot decide both u and v.

Matrix Notation and Matrix Multiplication

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:

$$\begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 1 & 4 \\ 1 & 9 \end{bmatrix}$$

$$2 \times \begin{bmatrix} 1 & 2 & 4 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 8 \\ 0 & 4 & 2 \end{bmatrix} \tag{1.16}$$

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: nine matrix coefficients; vector of three unknowns; vector of right-hand side.

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases} \quad \Leftrightarrow \quad A\vec{X} = \vec{b} \tag{1.17}$$

On the right-hand side (1.17) is the column vector $\vec{b}$. On the left-hand side is vector of three unknowns $\vec{X} = (u, v, w) = \{u, v, w\} = [u, v, w] = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$. Also on the left-hand side are nine coefficients of *square* matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$ (one of which happens to be zero). 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: $[2,1,1] \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 2u + v + w = 5$ and the same for second and third components. ***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.

There are two ways to multiply a matrix A **and a vector** $\vec{X}$. One way is a row at a time. Each row of $\boldsymbol{A}$ combines with $\vec{X}$ to give a component of $A\vec{X}$. There are three inner products when $\boldsymbol{A}$ has three rows as in (1.17). That is how $A\vec{X}$ 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 $A\vec{X}$ is found all at once, as ***a combination of the three columns of A***:

$$u\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + v\begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} + w\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} \tag{1.18}$$

It corresponds to the "column picture" of linear equations. Therefore $A\vec{X}$ is ***a combination of the columns of A***. The coefficients are the components of $\vec{X}$.

So far we have a convenient shorthand $A\vec{X} = \vec{b}$ for the original system of equations (1.17). What about the operations that we are carried out during elimination? In our example, the first step subtracted *2* times the first equation from the second. ***The same result is achieved if we multiply A by this elementary matrix*** **(or** ***elimination matrix*****)** $\boldsymbol{E_{21}}$:

$$E_{21}A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}$$

Next step was subtraction *-1* times the first equation from the third and gets us the matrix:

$$E_{31}\left(E_{21}A\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix}$$

And the last step was subtraction *-1* times the second equation from the third and gets us the final upper triangal matrix *U*:

$$U = E_{32}\left(E_{31}\left(E_{21}A\right)\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

All these steps can be written down in one short form: $\left(E_{32}E_{31}E_{21}\right)A = U$ with the resulting elimination matrix: $E_{32}E_{31}E_{21}$.

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

(I) Each entry of *AB* is the product of a *row* and a *column*: $(AB)_{ij}$ = (row *i* of *A*) times (column *j* of *B*).

(II) Each column of *AB* is the product of a matrix and a column: column *j* of *AB* = *A* times (column *j* of *B*).

(III) Each row of AB is the product of a row and a matrix: row i of AB = (row i of A) times B.

This leads hack to the key properties 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:

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

Matrix operations are distributive: $A(B+C) = AB+AC$ and $(B+C)D = BD+CD$

Matrix multiplication is not commutative: Usually $AB \neq BA$.