Chapter 2

Vector Spaces

2.1 Vector Spaces and Subspaces

Elimination can simplify, one entry at a time, the linear system $Ax = b$. Fortunately it also simplifies the theory. The basic questions of *existence* and *uniqueness*—Is there one solution, or no solution, or an infinity of solutions?—are much easier to answer after elimination, We need to devote one more section to those questions, to find every solution for an m by n system. Then that circle of ideas will be complete.

But elimination produces only one kind of understanding of $Ax = b$. Our chief object is to achieve a different and deeper understanding. This chapter may be more difficult than the first one. It goes to the heart of linear algebra.

For the concept of a ***vector space***, we start immediately with the most important spaces. They are denoted by $\mathbf{R}^1, \mathbf{R}^2, \mathbf{R}^3, \ldots$; the space $\mathbf{R}^n$ consists of *all column vectors with n components*. (We write $\mathbf{R}$ because the components are real numbers.) $\mathbf{R}^2$ is represented by the usual x-y plane; the two components of the vector become the x and y coordinates of the corresponding point. The three components of a vector in $\mathbf{R}^3$ give a point in three-dimensional space. The one-dimensional space $\mathbf{R}^1$ is a line.

The valuable thing for linear algebra is that the extension to n dimensions is so straightforward. For a vector in $\mathbf{R}^7$ we just need the seven components, even if the geometry is hard to visualize. Within all vector spaces, two operations are possible:

We can add any two vectors, and we can multiply all vectors by scalars.
In other words, we can take linear combinations.

Addition obeys the commutative law $x + y = y + x$; there is a "zero vector" satisfying $0 + x = x$; and there is a vector "$-x$" satisfying $-x + x = 0$. Eight properties (including those three) are fundamental; the full list is given in Problem 5 at the end of this section. ***A real vector space is a set of*** *vectors* ***together with rules for vector addition and multiplication by real numbers***. Addition and multiplication must produce vectors in the space, and they must satisfy the eight conditions.

Normally our vectors belong to one of the spaces $\mathbf{R}^n$; they are ordinary column vectors. If $x = (1,0,0,3)$, then $2x$ (and also $x+x$) has components 2, 0, 0, 6. The formal definition allows other things to be "vectors"-provided that addition and scalar multiplication are all right. We give three examples:

1. *The infinite-dimensional space* $\mathbf{R}^\infty$. Its vectors have infinitely many components, as in $x = (1,2,1,2,\ldots)$. The laws for $x+y$ and cx stay unchanged.

2. *The space of 3 by 2 matrices.* In this case the "vectors" are matrices! We can add two matrices, and $A+B = B+A$, and there is a zero matrix, and so on. This space is almost the same as $\mathbf{R}^6$. (The six components are arranged in a rectangle instead of a column.) Any choice of m and n would give, as a similar example, the vector space of all m by n matrices.

3. *The space of functions* $f(x)$. Here we admit all functions f that are defined on a fixed interval, say $0 \le x \le 1$. The space includes $f(x) = x^2$, $g(x) = \sin x$, their sum $(f+g)(x) = x^2 + \sin x$, and all multiples like $3x^2$ and $-\sin x$. The vectors are functions, and the dimension is somehow a larger infinity than for $\mathbf{R}^\infty$.

Other examples are given in the exercises, but the vector spaces we need most are somewhere else—***they are inside the standard spaces*** $\mathbf{R}^n$. We want to describe them and explain why they are important. Geometrically, think of the usual three-dimensional $\mathbf{R}^3$ and choose any plane through the origin. ***That plane is a vector space in its own right***. If we multiply a vector in the plane by 3, or -3, or any other scalar, we get a vector in the same plane. If we add two vectors in the plane, their sum stays in the plane. This plane through $(0,0,0)$ illustrates one of the most fundamental ideas in linear algebra; it is a ***subspace*** of the original space $\mathbf{R}^3$.

Definition. A ***subspace*** of a vector space is a nonempty subset that satisfies the requirements for a vector space: ***Linear combinations stay in the subspace***.

(i) If we add any vectors x and y in the subspace, $x+y$ is *in the subspace*.

(ii) If we multiply any vector x in the subspace by any scalar c, cx is *in the subspace*.

Notice our emphasis on the word ***space***. A sub*space* is a subset that is "closed" under addition and scalar multiplication. Those operations follow the rules of the host space, keeping us ***inside the subspace***. The eight required properties are satisfied in the larger space and will automatically be satisfied in every subspace. Notice in particular that ***the zero vector will belong to every subspace***. That comes from rule (ii): Choose the scalar to be $c = 0$.

The smallest subspace $\mathbf{Z}$ contains only one vector, the zero vector. It is a "zero-dimensional space," containing only the point at the origin. Rules (i) and (ii) are satisfied,

since the sum $0+0$ is in this one-point space, and so are all multiples $c0$. *This is the smallest possible vector space*: the empty set is not allowed. At the other extreme. the largest subspace is the whole of the original space. If the original space is $\mathbf{R}^3$, then the possible subspaces are easy to describe: $\mathbf{R}^3$ itself, any plane through the origin, any line through the origin, or the origin (the zero vector) alone.

The distinction between a subset and a subspace is made clear by examples. In each case, can you add vectors and multiply by scalars, without leaving the space?

Example 1. Consider all vectors in $\mathbf{R}^2$ whose components are positive or zero. This subset is the first quadrant of the x-y plane; the coordinates satisfy $x \geq 0$ and $y \geq 0$. It is *not a subspace*, even though it contains zero and addition does leave us within the subset. Rule (ii) is violated, since if the scalar is -1 and the vector is $[1 \;\; 1]$, the multiple $cx = [-1 \;\; -1]$ is in the third quadrant instead of the first.

If we include the third quadrant along with the first, scalar multiplication is all right. Every multiple cx will stay in this subset. However, rule (i) is now violated, since adding $[1 \;\; 2]+[-2 \;\; -1]$ gives $[-1 \;\; 1]$, which is not in either quadrant. The smallest subspace containing the first quadrant is the whole space $\mathbf{R}^2$.

Example 2. Start from the vector space of 3 by 3 matrices. One possible subspace is the set of *lower triangular matrices*. Another is the set of *symmetric matrices*. $A+B$ and cA are lower triangular if A and B are lower triangular, and they are symmetric if A and B are symmetric. Of course, the zero matrix is in both subspaces.

The Column Space of A

We now come to the key examples, the **column space** and the **nullspace** of a matrix A. ***The column space contains all linear combinations of the columns of*** A. It is a subspace of $\mathbf{R}^m$. We illustrate by a system of $m = 3$ equations in $n = 2$ unknowns:

$$\textbf{Combination of columns equals } b \qquad \begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{1}$$

With $m > n$ we have more equations than unknowns—and *usually there will be no solution*. The system will be solvable only for a very "thin" subset of all possible b's. One way of describing this thin subset is so simple that it is easy to overlook.

> **2A** The system $Ax = b$ is solvable if and only if the vector b can be expressed as a combination of the columns of A. Then b is in the column space.

This description involves nothing more than a restatement of $Ax = b$, *by columns*:

$$\textbf{Combination of columns} \qquad u \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} + v \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{2}$$

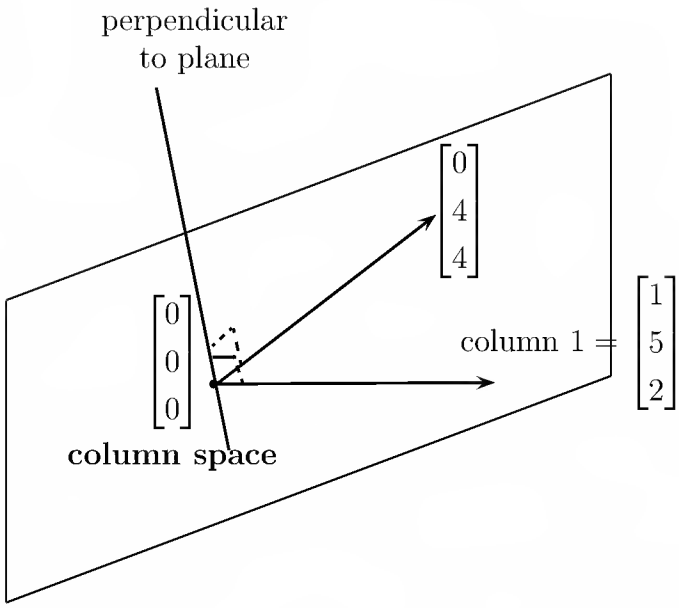

Figure 2.1: The column space $C(A)$, a plane in three-dimensional space.

These are the same three equations in two unknowns. Now the problem is: Find numbers u and v that multiply the first and second columns to produce b. The system is solvable exactly when such coefficients exist, and the vector (u,v) is the solution x.

We are saying that the attainable right-hand sides b are *all combinations of the columns of* A. One possible right-hand side is the first column itself; the weights are $u = 1$ and $v = 0$. Another possibility is the second column: $u = 0$ and $v = 1$. A third is the right-hand side $b = 0$. With $u = 0$ and $v = 0$, the vector $b = 0$ will always be attainable.

We can describe *all combinations* of the two columns geometrically: *$Ax = b$ can be solved if and only if b lies in the* ***plane*** *that is spanned by the two column vectors* (Figure 2.1). This is the thin set of attainable b. If b lies off the plane, then it is not a combination of the two columns. In that case $Ax = b$ has no solution.

What is important is that this plane is not just a subset of $\mathbf{R}^3$ it is a subspace. It is the ***column space*** of A, consisting of ***all combinations of the columns***. It is denoted by $C(A)$. Requirements (i) and (ii) for a subspace of $\mathbf{R}^m$ are easy to check:

(i) Suppose b and b' lie in the column space, so that $Ax = b$ for some x and $Ax' = b'$ for some x'. Then $A(x+x') = b+b'$, so that $b+b'$ is also a combination of the columns. The column space of all attainable vectors b is closed under addition.

(ii) If b is in the column space $C(A)$, so is any multiple cb. If some combination of columns produces b (say $Ax = b$), then multiplying that combination by c will produce cb. In other words, $A(cx) = cb$.

For another matrix A, the dimensions in Figure 2.1 may be very different. The smallest possible column space (one vector only) comes from the zero matrix $A = 0$. The

only combination of the columns is $b = 0$. At the other extreme, suppose A is the 5 by 5 identity matrix. Then $C(I)$ is the whole of $\mathbf{R}^5$; the five columns of I can combine to produce any five-dimensional vector b. This is not at all special to the identity matrix. *Any 5 by 5 matrix that is nonsingular will have the whole of* $\mathbf{R}^5$ *as its column space.* For such a matrix we can solve $Ax = b$ by Gaussian elimination; there are five pivots. Therefore every b is in $C(A)$ for a nonsingular matrix.

You can see how Chapter 1 is contained in this chapter. There we studied n by n matrices whose column space is $\mathbf{R}^n$. Now we allow singular matrices, and rectangular matrices of any shape. Then $C(A)$ can be somewhere between the zero space and the whole space $\mathbf{R}^m$. Together with its perpendicular space, it gives one of our two approaches to understanding $Ax = b$.

The Nullspace of A

The second approach to $Ax = b$ is "dual" to the first. We are concerned not only with attainable right-hand sides b, but also with the solutions x that attain them. The right-hand side $b = 0$ always allows the solution $x = 0$, but there may be infinitely many other solutions. (There always are, if there are more unknowns than equations, $n > m$.) ***The solutions to*** $Ax = 0$ ***form a vector space—the nullspace of*** A.

> The ***nullspace*** of a matrix consists of all vectors x such that $Ax = 0$. It is denoted by $N(A)$. It is a subspace of $\mathbf{R}^n$, just as the column space was a subspace of $\mathbf{R}^m$.

Requirement (i) holds: If $Ax = 0$ and $Ax' = 0$, then $A(x + x') = 0$. Requirement (ii) also holds: If $Ax = 0$ then $A(cx) = 0$. Both requirements fail if the right-hand side is not zero! Only the solutions to a *homogeneous* equation ($b = 0$) form a subspace. The nullspace is easy to find for the example given above; it is as small as possible:

$$\begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The first equation gives $u = 0$, and the second equation then forces $v = 0$. The nullspace contains only the vector $(0,0)$. This matrix has "independent columns"—a key idea that comes soon.

The situation is changed when a third column is a combination of the first two:

$$\textbf{Larger nullspace} \qquad B = \begin{bmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{bmatrix}.$$

B has the same column space as A. The new column lies in the plane of Figure 2.1; it is the sum of the two column vectors we started with. But the nullspace of B contains the