

## Subspaces of $\mathbb{R}^n$ and the Structure of Solving Equations

We have answered the principal question of linear algebra. We know how find the solution set for a system of linear equations, and we even have an algorithm for writing it down explicitly.

So we will add depth to our understanding, next. Part of this will involve looking more closely at the two geometric interpretations of a system: the column picture and the transformational picture. Another part will be to introduce and use the notion of a *subspace* to organize our thinking, and find the structure hiding in plain sight. The subspaces associated to a system/linear combination/matrix will give us the tools to answer all of the questions we set ourselves at the end of chapter 3, though maybe not all of them right away. We will encounter many new geometric concepts about vectors, including *spanning sets*, *linear dependence* and *linear independence*, and *basis*. Finally, we will see how the process of Gaussian Elimination can be refined a bit to get at the answers to these questions more quickly.

**How can we understand the structure of a system of linear equations?**

### NOTATIONAL REFRESHER

It will help us to have a uniform notation system for our work, so let us set one up. A system of  $m$  linear equations in  $n$  unknowns

SUBSPACES OF  $\mathbb{R}^n$  AND THE STRUCTURE OF SOLVING EQUATIONS  
in standard form looks like this.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \ddots \qquad \qquad \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

We can rewrite it as a linear combination of  $n$  vectors in  $\mathbb{R}^m$  like so:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

It will be convenient to have a short-hand for the names of these vectors, so for each  $j = 1, \dots, n$ , call the  $j$ th column  $u_j$ , and call the vector on the right-hand side  $b$ .

$$u_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Now we can write our system in this form

$$x_1 u_1 + \dots + x_n u_n = b.$$

Finally, we can also bundle together those vectors  $u_j$  as the columns of an  $m \times n$  matrix  $A$ , and put the unknowns into a vector  $x$

$$A = \begin{pmatrix} | & | & \dots & | \\ u_1 & u_2 & \dots & u_n \\ | & | & & | \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and then rewrite our system as a matrix-vector equation

$$Ax = b.$$

## THE IDEA OF A SUBSPACE

The new organizing idea for our study is that of a subspace. Roughly speaking, a subspace of  $\mathbb{R}^n$  is a collection of points in  $\mathbb{R}^n$  that has the same basic structure of linear combinations.

**Definition 1.** *Let  $S$  be a collection of vectors in  $\mathbb{R}^n$ . We say that  $S$  is a subspace of  $\mathbb{R}^n$  when the following conditions hold:*

- The zero vector is an element of  $S$ .*
- For any vectors  $w$  and  $v$  in  $S$ , and any scalars  $\lambda, \mu$ , the linear combination  $\lambda w + \mu v$  is also in  $S$ .*

We have already encountered some subspaces, but it helps to consider the most important classes of examples.

### SPANS AS SUBSPACES

Suppose you have a collection of  $m$ -vectors  $\{u_1, \dots, u_n\}$ . Since these are all  $m$ -vectors, they all live in  $\mathbb{R}^m$ . The set

$$\text{Span}\{u_1, \dots, u_n\} = \{x_1 u_1 + \dots + x_n u_n \mid x_1, \dots, x_n \in \mathbb{R}\}$$

consisting of all linear combinations of the given vectors is a subspace. This is not difficult to check. First, if we choose all of the coefficients  $x_i = 0$ , we see that the zero vector is an element of  $\text{Span}\{u_1, \dots, u_n\}$ . Then if we have two elements  $w$  and  $v$  in the span, and two numbers  $\lambda, \mu$ , we need to check that the linear combination  $\lambda w + \mu v$  is an element of  $\text{Span}\{u_1, \dots, u_n\}$ . Since  $w$  is an element of the span, we can write  $w$  as a linear combination of the  $u_j$ 's for some choice of coefficients  $s_j$ :

$$w = s_1 u_1 + s_2 u_2 + \dots + s_n u_n.$$

Similarly,  $v$  is an element of the span, so we can write  $v$  as a linear combination of the  $u_j$ 's for some choice of coefficients  $t_j$ :

$$v = t_1 u_1 + t_2 u_2 + \dots + t_n u_n.$$

## SUBSPACES OF $\mathbb{R}^n$ AND THE STRUCTURE OF SOLVING EQUATIONS

If we are careful and do a little rearranging, we learn that

$$\begin{aligned}\lambda w + \mu v &= \lambda(s_1 u_1 + \cdots + s_n u_n) + \mu(t_1 u_1 + \cdots + t_n u_n) \\ &= (\lambda s_1 + \mu t_1) u_1 + \cdots + (\lambda s_n + \mu t_n) u_n,\end{aligned}$$

so that the linear combination  $\lambda w + \mu v$  is also an element of  $\text{Span}\{u_1, \dots, u_n\}$ . Since we have checked both parts of the definition, we are done. So we don't lose it, let us recap with a firm statement.

**Theorem 2.** *If  $u_1, \dots, u_n$  are all  $m$ -vectors, the set  $\text{Span}\{u_1, \dots, u_n\}$  is a subspace of  $\mathbb{R}^m$ .*

Since we can write our systems of equations in the form of a linear combination of vectors equation, we see that subspaces built as spans will be relevant.

## HOMOGENEOUS EQUATIONS DEFINING SUBSPACES

The other natural way to build subspaces has come up, but is hidden deeper in our work so far.

**Definition 3.** *A system of  $m$  linear equations in  $n$  unknowns is called homogeneous if the right hand side consists of all zeros. That is, we say a system is homogeneous if it has the form*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

The fact that the right hand side is all zeros gives this system just a little extra help, and it is enough to make the solution set into a subspace of  $\mathbb{R}^n$ .

**Theorem 4.** *The solution set of a homogeneous system of  $m$  linear equations in  $n$  unknowns is a subspace of  $\mathbb{R}^n$ .*

*Proof.* There are two conditions to check. First, we must show that the zero vector is a solution to this system. But, if we set

$x_1 = \cdots = x_n = 0$ , of course the left hand side of every equation collapses to zero. Since this system is homogeneous, the right hand side of every equation is zero. Thus the zero vector is a solution to this system.

Next, we must show that if  $w$  and  $v$  are two solutions, and  $\lambda$  and  $\mu$  are scalars, then the linear combination  $\lambda w + \mu v$  is also a solution. Write

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

For the moment, consider just the first equation. Since  $w$  is a solution of this system, its components must satisfy the first equation. That is, we must have that

$$a_{11}w_1 + a_{12}w_2 + \cdots + a_{1n}w_n = 0.$$

Similarly, since  $v$  is a solution of the system, its components also satisfy the first equation.

$$a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n = 0.$$

But now we may multiply through the first equation by  $\lambda$  and the second equation by  $\mu$  and add to find (after some rearranging)

$$a_{11}(\lambda w_1 + \mu v_1) + a_{12}(\lambda w_2 + \mu v_2) + \cdots + a_{1n}(\lambda w_n + \mu v_n) = 0.$$

These entries are now just the components of the linear combination  $\lambda w + \mu v$ . So we deduce that  $\lambda w + \mu v$  satisfies the first equation.

But there is nothing special about the first equation in that argument. We can do exactly the same thing for all  $m$  equations. We deduce that  $\lambda w + \mu v$  is a solution of the homogeneous system of equations.

We have checked both parts of the definition, hence the solution set to our homogeneous system of equations is a subspace of  $\mathbb{R}^n$ .

□

## SUBSPACES OF $\mathbb{R}^n$ AND THE STRUCTURE OF SOLVING EQUATIONS DESCRIBING A SPAN WITH EQUATIONS

It seems now that we have two methods for making subspaces: spans, and solution sets to homogeneous equations. But stop for a moment and reflect on the work we did to describe a solution set. If we start with a homogeneous system of linear equations and apply the Gaussian elimination routine, we eventually produce a parametric description for the solution set. Those always look like this:

$$\text{Soln} = \{d + t_1v_1 + t_2v_2 + \cdots + t_kv_k | t_i \in \mathbb{R}\}$$

Where does that  $d$  come from? It is always the set of constants on the right hand side left over after the elimination steps. But if we start with all zeros there, every row operation we do will leave all zeros there. So we must have that for a homogeneous system  $d = 0$ . That means our solution set has the form

$$\text{Soln} = \{t_1v_1 + t_2v_2 + \cdots + t_kv_k | t_i \in \mathbb{R}\},$$

which is just  $\text{Span}\{v_1, \dots, v_k\}$ !

For a homogeneous system of equations, Gaussian elimination tells us how to turn our implicit description of the subspace of solutions into a parametric description as a span. So, the two types of examples are really just the same type, with different descriptions.

Naturally, we would like to know if we can pass the other way. Fortunately, the answer is “Yes.”

**Theorem 5.** *Let  $u_1, \dots, u_n$  be a collection of vectors in  $\mathbb{R}^m$ . Then the subspace  $\text{Span}\{u_1, \dots, u_n\}$  can be realized as the solution set to a homogeneous system of linear equations.*

*Proof.* The proof is constructive, in that we will show how to explicitly calculate the equations which are required to make the homogeneous system.

A vector  $b$  lies in  $\text{Span}\{u_1, \dots, u_n\}$  exactly when it is possible to solve the linear combination of vectors equation

$$x_1u_1 + x_2u_2 + \cdots + x_nu_n = b.$$

Equivalently,  $b$  lies in  $\text{Span}\{u_1, \dots, u_n\}$  exactly when it is possible to solve the system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \ddots \qquad \qquad \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

In this case, we don't actually care what the solution vector  $x$  looks like. Instead, we want to know what conditions the components  $b_i$  of  $b$  must satisfy for this to have a solution. So, think of the  $b_i$ 's as the variable names for the space of possible elements of the span. (This is a copy of  $\mathbb{R}^m$ , since  $b$  is an  $m$ -vector.) We will try to find equations on the unknowns  $b_i$ .

The key is to apply the Gaussian elimination process to the equations, and just carry along the  $b_i$ 's on the right hand side however it goes. Here, it is not necessary to do any step where you introduce a parameter, because all we care about is if the system of equations has a solution or not. We don't actually care to solve it.

But what is the test for solvability? The critical thing is that no equation can become inconsistent. So, do the Gaussian elimination process until the system is triangular. Then gather together all of the equations which have taken the form

$$0x_1 + \dots + 0x_n = \text{some combination of the } b_i\text{'s.}$$

The system is consistent exactly when each of these equations is trivially true, that is, when those combinations of  $b_i$ 's on the right hand sides really do equal zero. This forms a homogeneous system of equations on the  $m$  unknowns  $b_1, \dots, b_m$ , and that system is one we desire.

If there are no equations of this special type, then we can conclude that there are no conditions required of the components of  $b$ , so any vector  $b$  will work. That is,  $\text{Span}\{u_1, \dots, u_n\} = \mathbb{R}^m$  is the whole of  $\mathbb{R}^m$ .  $\square$

SUBSPACES OF  $\mathbb{R}^n$  AND THE STRUCTURE OF SOLVING EQUATIONS  
 THE COLUMN SPACE OF A MATRIX AND THE  
 SOLVABILITY QUESTION

Consider our standard system of linear equations written now in the linear combination of vectors form

$$x_1u_1 + \dots x_nu_n = b$$

or the matrix-vector equation form  $Ax = b$ . Recall that we basically chose to define the matrix-vector multiplication so that these two equations mean exactly the same thing. That is, we let  $A$  be the matrix with the  $u_i$ 's as columns, and then we define

$$Ax = \begin{pmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1u_1 + \dots + x_nu_n.$$

It is very convenient to have this compact notation, so we will start to use it more.

Our second big question in Chapter 3 was “Does a given equation have a solution? Can you tell in advance, without computing the answer?” We are now prepared to answer this. Sort of. We will pull a standard mathematician’s trick by making up a new object that helps us answer the question, and then show how to compute that object in a few ways.

An equation is solvable exactly when we can write  $b$  as a linear combination of the vectors  $u_i$ , because that is what it means to find a solution. Finding a solution of a linear combination equation means to find the coefficients  $x_i$  that realize  $b$  as a linear combination of the  $u_i$ 's. From the point of view of the matrix-vector equation, the  $u_i$ 's are the columns of the matrix  $A$ . This motivates our next definition, and explains the results which follow.

**Definition 6.** *Let  $A$  be an  $m \times n$  matrix. The column space of  $A$  is the subspace of  $\mathbb{R}^m$  which is the span of the columns of  $A$ . That is, if we write  $A$  as a bundle of columns, each of which is an*



$m$ -vector,

$$A = \begin{pmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{pmatrix}$$

then the column space is

$$\text{col}(A) = \text{Span}\{u_1, \dots, u_n\}.$$

**Theorem 7.** *The system of linear equations in matrix form  $Ax = b$  has a solution exactly when  $b$  is an element of  $\text{col}(A)$ , the column space of  $A$ .*

At this point, you might object that we haven't actually solved the problem, we just gave everything new names so we can pretend we solved the problem. There are two responses to this. First, yes, you are correct, but at least we have understood the problem better. Since that is somewhat unsatisfying, the second response is that since we now know that the way to solve the problem involves subspaces, we can at least try to use what we know about subspaces to finish things.

So,  $Ax = b$  is solvable if, and only if,  $b$  lies in  $\text{col}(A)$ . It may or may not be obvious that a particular  $b$  lies in the column space. One way we can find out is to *use the forward pass of our Gaussian elimination technique to see if the system is consistent*. That is usually a reasonable thing to do.

If, for some reason, it seems like Gaussian elimination should be avoided, you can also try this: Take the column space, which is described as a span, and find the a homogeneous system of linear equations which describe that column space as their common solution. Then check if  $b$  satisfies those equations.

Oh, wait. That process is essentially Gaussian elimination.

Huh. Seems that Gaussian elimination is our tool, either way.

## THE NULL SPACE OF A MATRIX AND THE SHAPE OF THE SOLUTION SET

Let us pick up questions 3 and 4 from the big list.

# SUBSPACES OF $\mathbb{R}^n$ AND THE STRUCTURE OF SOLVING EQUATIONS

3. Suppose that a given equation has a solution, how many solutions does it have?
4. Suppose that a given equation has a solution, what shape is the set of solutions? Can it be reasonably understood and described?

Again, our answers will rely on an appropriate choice of subspace, but things are a little more delicate. The key observation is that there is a relationship between the solutions of  $Ax = b$  and the solutions of a related system. But first, we need a little bit of algebraic understanding of matrix-vector multiplication.

**Theorem 8.** *Matrix-vector multiplication distributes through linear combinations of vectors. That is, if  $A$  is an  $m \times n$  matrix,  $v$  and  $w$  are two  $n$ -vectors, and  $\lambda$  and  $\mu$  are numbers, then*

$$A(\lambda v + \mu w) = \lambda Av + \mu Aw.$$

*Proof.* Write out  $A$  as a bundle of its columns

$$A = \begin{pmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{pmatrix},$$

and write out  $v$  and  $w$  in coordinates

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

We will compute both sides of the equation in the statement of the theorem and verify that they are identical.

First, we compute the right-hand side. The vector  $Av$  is

$$Av = v_1 u_1 + \dots + v_n u_n,$$

and the vector  $Aw$  is

$$Aw = w_1 u_1 + \dots + w_n u_n.$$

So, if we multiply through by the appropriate numbers, add, and rearrange, we see that

$$\begin{aligned}\lambda Av + \mu Aw &= \lambda(v_1u_1 + \cdots + v_nu_n) + \mu(w_1u_1 + \cdots + w_nu_n) \\ &= (\lambda v_1 + \mu w_1)u_1 + \cdots (\lambda v_n + \mu w_n)u_n\end{aligned}$$

Now, let us compute the left-hand side. The vector  $\lambda v + \mu w$  has coordinates

$$\lambda v + \mu w = \begin{pmatrix} \lambda v_1 + \mu w_1 \\ \vdots \\ \lambda v_n + \mu w_n \end{pmatrix}.$$

So, by definition of the matrix-vector product as a linear combination of the columns of  $A$ , we have

$$A(\lambda v + \mu w) = (\lambda v_1 + \mu w_1)u_1 + \cdots (\lambda v_n + \mu w_n)u_n.$$

It is now clear that the two expressions have the same value, so we deduce that

$$\lambda Av + \mu Aw = A(\lambda v + \mu w),$$

as desired. □

By the way, this property that matrix-vector multiplication distributes over a linear combination is really important. And this kind of property happens often enough in mathematics that we give it its own name, in a nod to this basic example. The operation of “multiply on the left by a matrix” is an example of a *linear* operation. You have probably already seen one other linear operation: taking derivatives!

Anyway, now that we have this crucial property of linearity established, we will use it without mercy in the next few results. Be on the lookout!

# SUBSPACES OF $\mathbb{R}^n$ AND THE STRUCTURE OF SOLVING EQUATIONS

**Definition 9.** Let  $Ax = b$  be a system of equations in matrix form  $Ax = b$ , where  $A$  is an  $m \times n$  matrix. The associated homogeneous system of equations is the system with matrix form  $Ax = 0$ . That is, we keep the same coefficient matrix, but make the right-hand side into the zero vector.

If we are already considering a homogeneous system  $Ax = 0$  (that is, if  $b$  is already the zero vector), then  $Ax = 0$  is its own associated homogeneous system.

**Theorem 10.** Let  $Ax = b$  be a system of linear equations written in matrix-vector form, where  $A$  is an  $m \times n$  matrix. Suppose that the  $n$ -vectors  $v$  and  $w$  are both solutions to  $Ax = b$ . Then the vector  $v - w$  is a solution to the associated homogeneous system  $Ax = 0$ .

*Proof.* Since  $v$  and  $w$  are solutions of the system, we know that they satisfy the equations. Thus, we have

$$Av = b \quad \text{and} \quad Aw = b.$$

Therefore, we see that

$$A(v - w) = Av - Aw = b - b = 0.$$

So, by definition,  $v - w$  is a solution of the associated homogeneous system  $Ax = 0$ .  $\square$

**Theorem 11.** Let  $Ax = b$  be a system of linear equations written in matrix-vector form, where  $A$  is an  $m \times n$  matrix. Suppose that  $p$  is a solution of the system  $Ax = b$ , and that  $v$  is a solution to the associated homogeneous system  $Ax = 0$ . Then the vector  $p + v$  is a solution to the equation  $Ax = b$ .

*Proof.* Since  $p$  is a solution to  $Ax = b$ , and  $v$  is a solution to  $Ax = 0$ , we learn that

$$Ap = b \quad \text{and} \quad Av = 0.$$

So, then the vector  $p + v$  satisfies

$$A(p + v) = Ap + Av = b + 0 = b,$$

which means that  $p + v$  is a solution of the equation  $Ax = b$ .  $\square$

If we put these together, we get a real structure result for the solution sets to systems of linear equations.

**Theorem 12.** *Let  $A$  be an  $m \times n$  matrix, and  $b$  an  $m$ -vector. Let  $S$  be the set of solutions to the system of equations  $Ax = b$  and let  $N$  be the set of solutions to the associated homogeneous system of equations  $Ax = 0$ .*

*Assume that the system  $Ax = b$  is consistent and, therefore, has at least one solution  $x = p$ . Then the solution sets  $S$  and  $N$  are related as follows:*

$$S = \{p + x \mid x \text{ is an element of } N\}.$$

*Proof.* For convenience, let  $p + N$  be a shorthand for the set

$$p + N = \{p + x \mid x \text{ is an element of } N\}.$$

We have to show that  $S$  is a subset of  $p + N$ , and that  $p + N$  is a subset of  $S$ .

Suppose that  $y$  is an element of  $S$ . Then  $y$  is a solution to  $Ax = b$ . By Theorem 10, the vector  $y - p$  is a solution of  $Ax = 0$ . That is,  $x = y - p$  is an element of  $N$ . If we rearrange, we see that  $y = x + p$ , so  $y$  is of the form required to be an element of  $p + N$ .

Next, suppose that  $y$  is an element of  $p + N$ . Then there is some element  $x$  of  $N$  so that  $y = p + x$ . By Theorem 11,  $y$  is a solution of  $Ax = b$  and hence an element of  $S$ .

Since  $S$  is a subset of  $p + N$  and  $p + N$  is a subset of  $S$ , these two sets must be the same.  $\square$

This is very useful. We can break the problem of solving  $Ax = b$  into two steps. First, find just any one solution  $p$ . This is usually

## SUBSPACES OF $\mathbb{R}^n$ AND THE STRUCTURE OF SOLVING EQUATIONS

called a *particular solution*, just to emphasize that we picked one. Second, we then solve the system  $Ax = 0$  to find  $N$ , which we hope is a bit easier because of the 0. We then put things together by shifting the set  $N$  over by  $p$  to sit where  $S$  is.

And we really mean “shift over.” To go from  $N$  to  $S$ , we add  $p$  to every element. Geometrically, that is just a translation by  $p$ .

Now, the process just says “find a particular solution  $p$ .” There might be many of those, but the argument works the same in any case. We just need one, we do not care which one. Any single one will do. The resulting set of solutions  $S$  will be the same.

We are now left with the process of finding the set  $N$  of solutions to a homogeneous system of equations  $Ax = 0$ . But we saw in Theorem 4 that such a set  $N$  is always a subspace. Let’s give it a name and a notation.

**Definition 13.** *Let  $A$  be an  $m \times n$  matrix. The set of vectors in  $\mathbb{R}^n$  which are solutions to the homogeneous equations  $Ax = 0$  is called the null space of  $A$ , and denoted*

$$\text{null}(A) = \{x \mid Ax = 0\}.$$

## COMPUTING THE NULL SPACE WITH THE RREF

### EFFICIENCY OF SPANNING SETS: LINEAR INDEPENDENCE AND LINEAR DEPENDENCE, BASIS

### A RENEWED LOOK AT SOLVING SYSTEMS