Lesson 9

Reading: Larson, Section 4.4, Spanning Sets and Linear Independence; Section 4.5, Basis and Dimension; Section 4.6, Rank of a Matrix and Systems of Linear Equations

Suggested exercises: Larson, Section 4.4: 1a, 1d, 3a, 3c, 9, 17, 21, 23, 29, 31, 35, 37, 49, 57; Section 4.5: 7, 11, 19, 31, 33, 35, 39, 47, 63, 67; Section 4.6: 7, 9, 11, 13, 15, 19, 21, 25, 27, 29, 31, 37, 39, 43, 47, 49, 51, 55

Submit: Lesson 9: Linear independence, bases, and rank

Section 4.4: Spanning Sets and Linear Independence

Two fundamental ideas are introduced in this section. The first is the idea of the *span* of a set of vectors. The span of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the set of all vectors that can be realized as a linear combination of these vectors, in other words, all vectors of the form

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n,$$

for every possible choice of scalars $c_1, \ldots, c_n$.

If you are given any set of vectors in a vector space V, their span is a subspace of V. Conversely, given any subspace W of a vector space V (including the possibility that $W = V$), there is a set of vectors whose span is equal to W.

For example, in R^3, the span of $\{(1,1,2),(0,1,1)\}$ is a subspace of R^3. This subspace is the plane through the origin that contains both of these vectors. We can discover an equation of this plane as follows: a typical member of the span is a vector (x, y, z) that is a linear combination of the two given vectors, so it is of the form

$$(x, y, z) = a(1,1,2) + b(0,1,1)$$

for some choice of scalars a and b. This gives us three equations

$$\begin{aligned} x &= a \\ y &= a + b \\ z &= 2a + b, \end{aligned}$$

and we can eliminate the parameters a, b, c to find our equation. First, from the first equation, we can replace a with x in the second and third equations to get

$$\begin{aligned} y &= x + b \\ z &= 2x + b, \end{aligned}$$

or

$$\begin{aligned} y - x &= b \\ z - 2x &= b. \end{aligned}$$

Equating the two expressions that are equal to b gives us

$$y - x = z - 2x,$$

or

$$x + y - z = 0.$$

We will see how to find a basis for the span of a set of vectors in the next section (when we find a basis for the row space of a matrix).

The second fundamental idea introduced in this section is the idea of linear independence. By definition: a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent when

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

only if $c_1 = \cdots = c_n = 0$. This condition is just a formalized way of stating that none of the vectors in the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linear combination of any of the others. For example, if we had a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for which $\mathbf{v}_2 = 2\mathbf{v}_1 - 5\mathbf{v}_3$, then we could write

$$2\mathbf{v}_1 - \mathbf{v}_2 - 5\mathbf{v}_3 = \mathbf{0},$$

allowing us to write the zero vector as a nontrivial linear combination (that is, one in which not all coefficients are zero) of the three vectors.

A set of vectors that fails to be linearly independent is called *linearly dependent*. A linearly dependent set of vectors is a set in which at least one of the vectors is a linear combination of the others. (Typically, if one vector is a linear combination of the others, then other vectors are, as well. We have to state this so carefully because of the extreme case where one of the vectors is the zero vector. A set containing the zero vector is always linearly dependent, even if it is the only vector in the set. For example, in R^3, the set

$$\{(1,0,0),(0,1,0),(0,0,0)\}$$

is linearly dependent: we have

$$(0,0,0) = 0 \cdot (1,0,0) + 0 \cdot (0,1,0),$$

so $(0,0,0)$ is a linear combination of the other vectors, but it is the only one in the set that is a linear combination of the others. Note that we can write

$$0 \cdot (1,0,0) + 0 \cdot (0,1,0) + 1 \cdot (0,0,0) = (0,0,0),$$

which shows that we have a nontrivial linear combination of our three vectors giving us the zero vector in R^3, which shows that this set fails to be linearly independent.

In general, testing a set of vectors for linear independence is a matter of setting up

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0},$$

which is just a homogeneous linear system of equations, and finding out if there are nonzero solutions or not. Let's look at an example. Suppose we want to find out if $\{(1,2),(3,4),(5,6)\}$ is linearly independent. We set up the equation

$$c_1(1,2) + c_2(3,4) + c_3(5,6) = (0,0)$$

and then write out the linear system of equations in the components:

$$\begin{aligned} c_1 + 3c_2 + 5c_3 &= 0 \\ 2c_1 + 4c_2 + 6c_3 &= 0. \end{aligned}$$

The augmented matrix of this system is

$$\begin{bmatrix} 1 & 3 & 5 & 0 \\ 2 & 4 & 6 & 0 \end{bmatrix},$$

which reduces to

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

Our solution is

$$\begin{aligned} c_1 &= c_3 \\ c_2 &= -2c_3, \end{aligned}$$

which means we have nontrivial solutions. Therefore, this set of vectors is linearly dependent.

This tells us that, for example, the vector $(5, 6)$ is a linear combination of $(1, 2)$ and $(3, 4)$: choose $c_3 = 1$ in our solution to get $c_1 = 1$ and $c_2 = -2$, which gives us

$$(1, 2) - 2(3, 4) + (5, 6) = (0, 0),$$

so that

$$(5, 6) = (-1) \cdot (1, 2) + 2 \cdot (3, 4).$$

As a result of this, when looking at the span of this set of vectors, we don't really need all three vectors to generate all possible linear combinations. The set of all linear combinations of $(1, 2)$ and $(3, 4)$ will suffice, since adding on a scalar multiple of $(5, 6)$ doesn't really add anything new, it just adds another linear combination of $(1, 2)$ and $(3, 4)$. That is, if we have a linear combination

$$a(1, 2) + b(3, 4) + c(5, 6),$$

we can use our result above and write this as

$$a(1, 2) + b(3, 4) + c\left[(-1) \cdot (1, 2) + 2 \cdot (3, 4)\right],$$

or

$$(a - c) \cdot (1, 2) + (b + 2c) \cdot (3, 4),$$

which is just a linear combination of $(1, 2)$ and $(3, 4)$.

Thus, our list of vectors contains more vectors than it really needs, if we are interested in looking at the span of these vectors. A linearly independent set, on the other hand, doesn't have any wasted vectors: if you are interested in the span of the set of vectors, you cannot omit any of the vectors and still have the same span. This leads us into the next section.

4.5: Basis and Dimension

A *basis* of a vector space V is a set of vectors that spans the space and is linearly independent. This means that (1) every vector in V is a linear combination of the vectors in the set, and (2) the linear combination is unique. Why does it have to be unique? Suppose our set of vectors is $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. If we have two linear combinations of these vectors equal to each other, then we have

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n,$$

or

$$(a_1 - b_1)\mathbf{v}_1 + \cdots + (a_n - b_n)\mathbf{v}_n = \mathbf{0},$$

and since our set of vectors is linearly independent, we must therefore have all coefficients equal to zero, so

$$a_1 - b_1 = \cdots = a_n - b_n = 0.$$

Therefore, $a_1 = b_1, \ldots, a_n = b_n$. In other words, the two linear combinations are actually the same thing!

A vector space can have many different bases; infinitely many, in fact. In R^n, you just need a set of n linearly independent vectors to have a basis, and there are many such sets to choose from. There is a *standard basis* for R^n, which is designed to make things as simple as possible. In R^2, the standard basis is $\{(1,0),(0,1)\}$, and in R^3, the standard basis is $\{(1,0,0),(0,1,0),(0,0,1)\}$. I think you can imagine how it goes from here to the general case of R^n.

These vectors are chosen to be the standard basis because it is easy to figure out how to write any vector as a linear combination of the basis vectors. In R^3, for example, the vector (a,b,c) is a linear combination like this:

$$a(1,0,0) + b(0,1,0) + c(0,0,1).$$

That's as easy as it gets!

On the other hand, if you use the basis $\{(1,2,-4),(3,1,6),(-1,2,2)\}$, then finding a linear combination for, say, $(5,-6,11)$ is not as pleasant: you have to set up and solve a linear system.

Of course, the situation is not really that bad, now that you know about matrix inverses. Let's make use of what we've learned so far. We are trying to find constants c_1, c_2, c_3 such that

$$c_1 \begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 1 \\ 6 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 11 \end{bmatrix}.$$

Remember that the left-hand side can be written as a matrix product:

$$\begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 2 \\ -4 & 6 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 11 \end{bmatrix}.$$

The matrix containing our basis vectors in its columns is invertible, so

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 2 \\ -4 & 6 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ -6 \\ 11 \end{bmatrix}.$$

Once you have the inverse computed, then you can plug in as many vectors as you want on the right-hand side and get their coefficients without too much trouble.

Even though a vector space can have infinitely many different bases, all of these bases have one thing in common: for a given vector space, every basis must contain the same number of vectors. For example, every basis of R^3 has to contain precisely three vectors. This number is how we define the *dimension* of a vector space: the dimension of a vector space V is the number of vectors in any basis of V.

You should be aware that if V is a vector space of dimension n, then any set containing more than n vectors *must* be linearly dependent.

This idea of dimension holds for subspaces of a vector space, since a subspace is a subset of a vector space that can be considered as a vector space in its own right.

For example, in R^3, the plane $x+2y+3z=0$ is a subspace of R^3. Think of the equation defining the plane as a linear system. The variables y and z are free variables, so we can assign a parameter to each of them: set $y=s$ and $z=t$, which gives us $x=-2s-3t$. A typical vector in this plane is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2s-3t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix},$$

which shows that the set of vectors

$$\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

spans the subspace. This set of vectors is linearly independent (check!), so we have a basis of the subspace. Therefore, the subspace is (no surprise) two-dimensional.

4.6: Rank of a Matrix and Systems of Linear Equations

We are introduced to several important subspaces associated with a matrix in this section: the row space, the column space, and the nullspace.

The *row space* of a matrix A is the span of the row vectors of A (the rows, considered a vectors), and the *column space* of A is the span of the column vectors of A (the columns, considered as vectors).

The row space of a matrix is not affected by performing elementary row operations on the matrix, so you can reduce a matrix to reduced row-echelon form to find a simple basis for the row space. This will reveal whether your set of vectors is linearly independent or not: if reduced row-echelon form has fewer nonzero rows than the original matrix, then the set is linearly dependent. If it has the same number of nonzero rows, then your set of vectors is linearly independent. This is a good way to test a set of vectors for linear independence, by the way.

For example, consider the set $\{(1, -3, 1), (1, 1, 2), (2, -10, 1)\}$. Create the matrix

$$\begin{bmatrix} 1 & -3 & 1 \\ 1 & 1 & 2 \\ 2 & -10 & 1 \end{bmatrix},$$

and row reduce, obtaining

$$\begin{bmatrix} 1 & 0 & 7/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix}.$$

This gives us a basis $\{(1, 0, 7/4), (0, 1, 1/4)\}$ for the row space of the matrix, which is the same thing as the span of $\{(1, -3, 1), (1, 1, 2), (2, -10, 1)\}$. It also tells us that the set of three vectors is linearly dependent, because its span is two-dimensional.

Recall that a linear system $A\mathbf{x} = \mathbf{b}$ has a solution only if $\mathbf{b}$ is a linear combination of the columns of A (refer back to section 2.1 if necessary). In other words, $\mathbf{b}$ must be in the column space of A.

To find a basis for the column space of a matrix, transpose the matrix, making rows out of the columns, and reduce to reduced row-echelon form. Since the column space of A is the same thing as the row space of A^T, this gives us a basis for the column space.

Important fact: the row space and column space of a matrix A always have the same dimension. This common dimension is called the *rank* of the matrix A. Therefore, to find the rank of A: reduce to row-echelon form, and count the number of nonzero rows. For example, consider the matrix

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

This matrix reduces to

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 2 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

so its rank is 2. Therefore, both the row space and column space for this matrix are two-dimensional. The row space is a two-dimensional subspace of R^5; the column space is a two-dimensional subspace of R^4.

Another subspace associated with a matrix A is its nullspace: the subspace of solutions of $A\mathbf{x} = \mathbf{0}$. In solving homogeneous systems, the number of free variables is the dimension of the nullspace. For example, consider the system

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

The coefficient matrix is

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & -1 \end{bmatrix},$$

which reduces to

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

This shows that z is a free variable, and we can introduce a parameter k by setting $z = t$, and our reduced system

$$\begin{aligned} x + z &= 0 \\ y - z &= 0 \end{aligned}$$

has solutions

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -t \\ t \\ t \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

so we have a basis vector $(-1, 1, 1)$ for our one-dimensional nullspace.

If a matrix A has reduced row-echelon form

$$\begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -2 \end{bmatrix},$$

then we have two free variables, so we expect a two-dimensional nullspace. Considering A as the coefficient matrix of the system $A\mathbf{x} = \mathbf{0}$, we have the equations

$$\begin{aligned} x_1 - x_2 \qquad + 2x_4 &= 0 \\ x_3 - 2x_4 &= 0, \end{aligned}$$

and introducing parameters s and t for our free variables x_2 and x_4 by setting $x_2 = s$ and $x_4 = t$ gives us solutions

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} s - 2t \\ s \\ 2t \\ t \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix}.$$

Therefore, a basis for our two-dimensional nullspace is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

The *nullity* of a matrix is the dimension of its nullspace, and there is an important relationship between the rank and the nullity of a matrix, which is commonly referred to as the *rank-nullity theorem*. This theorem is given in the text, although it is not referred to by this name. The result is this: if A is an $m \times n$ matrix, then

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n.$$

Therefore, if you know either the rank or the nullity of a matrix, you can find the other value.

For example, for the 4×5 matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 & 2 \\ 0 & 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

we observe that the rank is 3, and we have five columns, so

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = 3 + \operatorname{nullity}(A) = 5,$$

from which we deduce that $\operatorname{nullity}(A) = 2$.

If you think about the matrix above, you can see the idea behind the proof of the rank-nullity theorem: the number of free variables gives us the dimension of the null space, and therefore the nullity of the matrix. The number of dependent variables (that is, the number of leading ones in the reduced row-echelon form) gives us the rank of the matrix. Add these together and you have accounted for all of the variables, which corresponds to the number of columns of the matrix.

The final topic in this section is the solution of a nonhomogeneous linear system $A\mathbf{x} = \mathbf{b}$.

First, such a system might not have a solution: remember, $\mathbf{b}$ must be in the column space of A in order for a solution to exist. If there is a solution, then the general form of the solution is always

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_p,$$

where $\mathbf{x}_p$ is a particular solution of $A\mathbf{x} = \mathbf{b}$ (*any* solution of the nonhomogeneous system) and $\mathbf{x}_h$ is the solution of the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. We have $A\mathbf{x}_p = \mathbf{b}$ and $A\mathbf{x}_h = \mathbf{0}$, so

$$A\mathbf{x} = A(\mathbf{x}_p + \mathbf{x}_h) = A\mathbf{x}_p + A\mathbf{x}_h = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

For example, consider the system $A\mathbf{x} = \mathbf{b}$ as follows:

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

The augmented matrix is

$$\begin{bmatrix} 1 & -3 & 1 & 4 \\ 1 & -1 & 1 & 2 \end{bmatrix},$$

which reduces to

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix},$$

representing the system

$$\begin{aligned} x \quad + z &= 1 \\ y \quad &= -1. \end{aligned}$$

The variable z is a free variable, and setting $z = t$ gives us a solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1-t \\ -1 \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Here,

$$\mathbf{x}_p = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

and

$$\mathbf{x}_h = t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

You can, and should, check that $A\mathbf{x}_p = \mathbf{b}$ and $A\mathbf{x}_h = \mathbf{0}$ for these vectors.

This section concludes with an expanded version of a set of equivalent conditions for square matrices: the ideas of rank, row space, and column space have now been incorporated.