

Related Topics in Linear Algebra

We discuss Related Topics in Linear Algebra

Subspaces and span

Assume that we find two vectors that solve $A\vec{x} = 0$. What other vectors also solve this equation? In our discussion of linear combinations, we saw that if \vec{x}_1 and \vec{x}_2 solve $A\vec{x} = 0$, then so does $A(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2)$ for any constants α_1 and α_2 . Thus, all linear combinations will also solve the equation. This leads to the definition of the span of a set of vectors.

Definition 1. The set of all linear combinations of a set of vectors is called their span.

$$\text{span}\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\} = \{\text{Set of all linear combinations of } \vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}.$$

Thus, if two vectors solve a homogeneous equation, so does everything in the span of those two vectors. The span of a collection of vectors is an example of a subspace, which is a common object in linear algebra. We say that a set S of vectors in \mathbb{R}^n is a *subspace* if whenever \vec{x} and \vec{y} are members of S and α is a scalar, then

$$\vec{x} + \vec{y}, \quad \text{and} \quad \alpha\vec{x}$$

are also members of S . That is, we can add and multiply by scalars and we still land in S . So every linear combination of vectors of S is still in S . That is really what a subspace is. It is a subset where we can take linear combinations and still end up being in the subset.

Example 1. If we let $S = \mathbb{R}^n$, then this S is a subspace of \mathbb{R}^n . Adding any two vectors in \mathbb{R}^n gets a vector in \mathbb{R}^n , and so does multiplying by scalars.

The set $S' = \{\vec{0}\}$, that is, the set of the zero vector by itself, is also a subspace of \mathbb{R}^n . There is only one vector in this subspace, so we only need to check for that one vector, and everything checks out: $\vec{0} + \vec{0} = \vec{0}$ and $\alpha\vec{0} = \vec{0}$.

The set S'' of all the vectors of the form (a, a) for any real number a , such as $(1, 1)$, $(3, 3)$, or $(-0.5, -0.5)$ is a subspace of \mathbb{R}^2 . Adding two such vectors, say $(1, 1) + (3, 3) = (4, 4)$ again gets a vector of the same form, and so does multiplying by a scalar, say $8(1, 1) = (8, 8)$.

We can apply these ideas to the vectors that live inside a matrix. The span of the rows of a matrix A is called the *row space*. The row space of A and the row space of the row echelon form of A are the same, because reducing the matrix A to its row echelon form involves taking linear combinations, which will preserve the span. In the example,

$$\begin{aligned} \text{row space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &= \text{span} \{ [1 \ 2 \ 3], [4 \ 5 \ 6], [7 \ 8 \ 9] \} \\ &= \text{span} \{ [1 \ 2 \ 3], [0 \ 1 \ 2] \}. \end{aligned}$$

Similarly to row space, the span of columns is called the *column space*.

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \right\}.$$

In particular, to find a set of linearly independent columns we need to look at where the pivots were. If you recall above, when solving $A\vec{x} = \vec{0}$ the key was finding the pivots, any non-pivot columns corresponded to free variables. That means we can solve for the non-pivot columns in terms of the pivot columns. Let's see an example.

Learning outcomes: Compute the rank of a matrix Find a maximal linearly independent subset of a set of vectors Compute a basis of a subspace and the dimension of that subspace Determine the kernel of a matrix using row reduction Understand the connection between rank and nullity in a given matrix Compute the inverse of a matrix using row reduction Use properties of the trace and determinant to analyze the eigenvalues of a matrix.

Author(s): Matthew Charnley and Jason Nowell

Example 2. Find the linearly independent columns of the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}.$$

Solution: We find a pivot and reduce the rows below:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 \\ 3 & 6 & 7 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -2 & -4 \end{bmatrix}.$$

We find the next pivot, make it one, and rinse and repeat:

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

The final matrix is the row echelon form of the matrix. Consider the pivots that we marked. The pivot columns are the first and the third column. All other columns correspond to free variables when solving $A\vec{x} = \vec{0}$, so all other columns can be solved in terms of the first and the third column. In other words

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

┘

We could perhaps use another pair of columns to get the same span, but the first and the third are guaranteed to work because they are pivot columns.

In the previous example, this means that only the first and third columns are “important” in the sense of generating the full column space as a span. We would like to have a way to talk about what these first and third columns do.

Definition 2 (Spanning set). *Let S be a subspace of a vector space. The set $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is a spanning set for the subspace S if each of these vectors are in S and the span of $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is equal to S .*

In the context of the previous example, for the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}$$

we know that

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

This means that both

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}$$

are spanning sets for this column space.

The idea also works in reverse. Suppose we have a bunch of column vectors and we just need to find a linearly independent set. For example, suppose we started with the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

These vectors are not linearly independent as we saw above. In particular, the span \vec{v}_1 and \vec{v}_3 is the same as the span of all four of the vectors. So \vec{v}_2 and \vec{v}_4 can both be written as linear combinations of \vec{v}_1 and \vec{v}_3 . A common thing that

comes up in practice is that one gets a set of vectors whose span is the set of solutions of some problem. But perhaps we get way too many vectors, we want to simplify. For example above, all vectors in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ can be written $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 + \alpha_4\vec{v}_4$ for some numbers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. But it is also true that every such vector can be written as $a\vec{v}_1 + b\vec{v}_3$ for two numbers a and b . And one has to admit, that looks much simpler. Moreover, these numbers a and b are unique. More on that later in this section.

To find this linearly independent set we simply take our vectors and form the matrix $[\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3 \ \vec{v}_4]$, that is, the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}.$$

We crank up the row-reduction machine, feed this matrix into it, and find the pivot columns and pick those. In this case, \vec{v}_1 and \vec{v}_3 .

Basis and dimension

At this point, we have talked about subspaces, and two other properties of sets of vectors: linear independence and being a spanning set for a subspace. In some sense, these two properties are in opposition to each other. If I add more vectors to a set, I am more likely to become a spanning set (because I have more options for adding to get other vectors), but less likely to be independent (because there are more possibilities for a linear combination to be zero). Similarly, the reverse is true; removing vectors means the set is more likely to be linearly independent, but less likely to span a given subspace. The question then becomes if there is a sweet spot where both things are true, and that leads to the definition of a basis.

Definition 3. If S is a subspace and we can find k linearly independent vectors in S

$$\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k,$$

such that every other vector in S is a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$, then the set $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is called a basis of S . In other words, S is the span of $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$. We say that S has dimension k , and we write

$$\dim S = k.$$

The next theorem illustrates the main properties and classification of a basis of a vector space.

Theorem 1. If $S \subset \mathbb{R}^n$ is a subspace and S is not the trivial subspace $\{\vec{0}\}$, then there exists a unique positive integer k (the dimension) and a (not unique) basis $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$, such that every \vec{w} in S can be uniquely represented by

$$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \dots + \alpha_k\vec{v}_k,$$

for some scalars $\alpha_1, \alpha_2, \dots, \alpha_k$.

We should reiterate that while k is unique (a subspace cannot have two different dimensions), the set of basis vectors is not at all unique. There are lots of different bases for any given subspace. Finding just the right basis for a subspace is a large part of what one does in linear algebra. In fact, that is what we spend a lot of time on in linear differential equations, although at first glance it may not seem like that is what we are doing.

Example 3. The standard basis

$$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n,$$

is a basis of \mathbb{R}^n (hence the name). So as expected

$$\dim \mathbb{R}^n = n.$$

On the other hand the subspace $\{\vec{0}\}$ is of dimension 0.

The subspace S'' from a previous example, that is, the set of vectors (a, a) is of dimension 1. One possible basis is simply $\{(1, 1)\}$, the single vector $(1, 1)$: every vector in S'' can be represented by $a(1, 1) = (a, a)$. Similarly another possible basis would be $\{(-1, -1)\}$. Then the vector (a, a) would be represented as $(-a)(-1, -1)$. In this case, the subspace S'' has many different bases, two of which are $\{(1, 1)\}$ and $\{(-1, -1)\}$, and the vector (a, a) has a different representation (different constant) for the different bases.

Row and column spaces of a matrix are also examples of subspaces, as they are given as the span of vectors. We can use what we know about row spaces and column spaces from the previous section to find a basis.

Example 4. Earlier, we considered the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix}.$$

Using row reduction to find the pivot columns, we found

$$\text{column space of } A \left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} \right) = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

What we did was we found the basis of the column space. The basis has two elements, and so the column space of A is two dimensional.

We would have followed the same procedure if we wanted to find the basis of the subspace X spanned by

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

We would have simply formed the matrix A with these vectors as columns and repeated the computation above. The subspace X is then the column space of A .

Example 5. Consider the matrix

$$L = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$

Conveniently, the matrix is in reduced row echelon form. The column space is the span of the pivot columns, because the pivot columns always form a basis for the column space of a matrix. It is the 3-dimensional space

$$\text{column space of } L = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^3.$$

The row space is the 3-dimensional space

$$\text{row space of } L = \text{span} \{ [1 \ 2 \ 0 \ 0 \ 3], [0 \ 0 \ 1 \ 0 \ 4], [0 \ 0 \ 0 \ 1 \ 5] \}.$$

As these vectors have 5 components, we think of the row space of L as a subspace of \mathbb{R}^5 .

Rank

In that last example, we noticed that the dimension of the row space and the column space were the same. It turns out that this is not a coincidence. In order to describe this in more detail, we need to define one more term.

Definition 4. Given a matrix A , the maximal number of linearly independent rows is called the rank of A , and we write “rank A ” for the rank.

For example,

$$\text{rank} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} = 1.$$

The second and third row are multiples of the first one. We cannot choose more than one row and still have a linearly independent set. But what is

$$\text{rank} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = ?$$

That seems to be a tougher question to answer. The first two rows are linearly independent, so the rank is at least two. If we would set up the equations for the α_1 , α_2 , and α_3 , we would find a system with infinitely many solutions. One solution is

$$[1 \ 2 \ 3] - 2[4 \ 5 \ 6] + [7 \ 8 \ 9] = [0 \ 0 \ 0].$$

So the set of all three rows is linearly dependent, the rank cannot be 3. Therefore the rank is 2.

But how can we do this in a more systematic way? We find the row echelon form!

$$\text{Row echelon form of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \text{ is } \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

The elementary row operations do not change the set of linear combinations of the rows (that was one of the main reasons for defining them as they were). In other words, the span of the rows of the A is the same as the span of the rows of the row echelon form of A . In particular, the number of linearly independent rows is the same. And in the row echelon form, all nonzero rows are linearly independent. This is not hard to see. Consider the two nonzero rows in the example above. Suppose we tried to solve for the α_1 and α_2 in

$$\alpha_1 \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

Since the first column of the row echelon matrix has zeros except in the first row means that $\alpha_1 = 0$. For the same reason, α_2 is zero. We only have two nonzero rows, and they are linearly independent, so the rank of the matrix is 2. This also tells us that if we were trying to solve the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= a \\ 4x_1 + 5x_2 + 6x_3 &= b \\ 7x_1 + 8x_2 + 9x_3 &= c \end{aligned}$$

we would get that one row of the reduced augmented matrix has all zeros on the left side, and so this system either has a free variable or is inconsistent, because only two equations here are relevant.

Referring back to the examples from earlier in this section, we could carry out the same calculations to say that

$$\text{rank} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 2$$

and

$$\text{rank} \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix} = 3.$$

We know how to find the set of linearly independent rows, but sometimes it may also be useful to find the linearly independent columns as well. It is a tremendously useful fact that the number of linearly independent columns is always the same as the number of linearly independent rows:

Theorem 2. $\text{rank } A = \text{rank } A^T$

Or, in the context of the row and column spaces that we have already discussed:

Theorem 3 (Rank). *The dimension of the column space and the dimension of the row space of a matrix A are both equal to the rank of A .*

This relates to the statement at the start of this section; since the number of vectors that we needed to take to get a basis of linearly independent columns was always the same as the number of pivots, and the number of pivots is the rank, we get that above theorem.

Kernel

The set of solutions of a linear equation $L\vec{x} = \vec{0}$, the kernel of L , is a subspace: If \vec{x} and \vec{y} are solutions, then

$$L(\vec{x} + \vec{y}) = L\vec{x} + L\vec{y} = \vec{0} + \vec{0} = \vec{0}, \quad \text{and} \quad L(\alpha\vec{x}) = \alpha L\vec{x} = \alpha\vec{0} = \vec{0}.$$

So $\vec{x} + \vec{y}$ and $\alpha\vec{x}$ are solutions. The dimension of the kernel is called the *nullity* of the matrix.

The same sort of idea governs the solutions of linear differential equations. We try to describe the kernel of a linear differential operator, and as it is a subspace, we look for a basis of this kernel. Much of this book is dedicated to finding such bases.

The kernel of a matrix is the same as the kernel of its reduced row echelon form. For a matrix in reduced row echelon form, the kernel is rather easy to find. If a vector \vec{x} is applied to a matrix L , then each entry in \vec{x} corresponds to a column of L , the column that the entry multiplies. To find the kernel, pick a non-pivot column make a vector that has a -1 in the entry corresponding to this non-pivot column and zeros at all the other entries corresponding to the other non-pivot columns. Then for all the entries corresponding to pivot columns make it precisely the value in the corresponding row of the non-pivot column to make the vector be a solution to $L\vec{x} = \vec{0}$. This procedure is best understood by example.

Example 6. Consider

$$L = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}.$$

This matrix is in reduced row echelon form, the pivots are marked. There are two non-pivot columns, so the kernel has dimension 2, that is, it is the span of 2 vectors. Let us find the first vector. We look at the first non-pivot column, the 2nd column, and we put a -1 in the 2nd entry of our vector. We put a 0 in the 5th entry as the 5th column is also a non-pivot column:

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

Let us fill the rest. When this vector hits the first row, we get a -2 and 1 times whatever the first question mark is. So make the first question mark 2. For the second and third rows, it is sufficient to make it the question marks zero. We are really filling in the non-pivot column into the remaining entries. Let us check while marking which numbers went where:

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

Yay! How about the second vector. We start with

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

We set the first question mark to 3, the second to 4, and the third to 5. Let us check, marking things as previously,

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

There are two non-pivot columns, so we only need two vectors. We have found the basis of the kernel. So,

$$\text{kernel of } L = \text{span} \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \\ 5 \\ -1 \end{bmatrix} \right\}$$

What we did in finding a basis of the kernel is we expressed all solutions of $L\vec{x} = \vec{0}$ as a linear combination of some given vectors.

The procedure to find the basis of the kernel of a matrix L :

- Find the reduced row echelon form of L .
- Write down the basis of the kernel as above, one vector for each non-pivot column.

The rank of a matrix is the dimension of the column space, and that is the span of the pivot columns, while the kernel is the span of vectors in the non-pivot columns. So the two numbers must add to the number of columns.

Theorem 4 (Rank–Nullity). *If a matrix A has n columns, rank r , and nullity k (dimension of the kernel), then*

$$n = r + k.$$

The theorem is immensely useful in applications. It allows one to compute the rank r if one knows the nullity k and vice versa, without doing any extra work.

Let us consider an example application, a simple version of the so-called *Fredholm alternative*. A similar result is true for differential equations. Consider

$$A\vec{x} = \vec{b},$$

where A is a square $n \times n$ matrix. There are then two mutually exclusive possibilities:

- (a) A nonzero solution \vec{x} to $A\vec{x} = \vec{0}$ exists.
- (b) The equation $A\vec{x} = \vec{b}$ has a unique solution \vec{x} for every \vec{b} .

How does the Rank–Nullity theorem come into the picture? Well, if A has a nonzero solution \vec{x} to $A\vec{x} = \vec{0}$, then the nullity k is positive. But then the rank $r = n - k$ must be less than n . In particular it means that the column space of A is of dimension less than n , so it is a subspace that does not include everything in \mathbb{R}^n . So \mathbb{R}^n has to contain some vector \vec{b} not in the column space of A . In fact, most vectors in \mathbb{R}^n are not in the column space of A .

The idea of a kernel also comes up when defining and discussing eigenvectors. In order to find this vector, we are looking for a vector \vec{v} so that

$$(A - \lambda I)\vec{v} = \vec{0}.$$

This means that we are looking for a vector \vec{v} that is in the kernel of the matrix $(A - \lambda I)$. Since the kernel is also a subspace, this means that the set of all eigenvectors of a matrix A with a certain eigenvalue is a subspace, so it has a dimension. This dimension is number of linearly independent eigenvectors with that eigenvalue, so it is the geometric multiplicity of this eigenvalue. This also motivates why this is sometimes called the *eigenspace* for a given eigenvalue. Finding a basis of this subspace (which is also finding the kernel of the matrix $A - \lambda I$) is the exact same as the process of finding the eigenvectors of the matrix A .

Computing the inverse

If the matrix A is square and there exists a unique solution \vec{x} to $A\vec{x} = \vec{b}$ for any \vec{b} (there are no free variables), then A is invertible.

In particular, if $A\vec{x} = \vec{b}$ then $\vec{x} = A^{-1}\vec{b}$. Now we just need to compute what A^{-1} is. We can surely do elimination every time we want to find $A^{-1}\vec{b}$, but that would be ridiculous. The mapping A^{-1} is linear and hence given by a matrix, and we have seen that to figure out the matrix we just need to find where does A^{-1} take the standard basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$.

That is, to find the first column of A^{-1} we solve $A\vec{x} = \vec{e}_1$, because then $A^{-1}\vec{e}_1 = \vec{x}$. To find the second column of A^{-1} we solve $A\vec{x} = \vec{e}_2$. And so on. It is really just n eliminations that we need to do. But it gets even easier. If you think about it, the elimination is the same for everything on the left side of the augmented matrix. Doing n eliminations separately we would redo most of the computations. Best is to do all at once.

Therefore, to find the inverse of A , we write an $n \times 2n$ augmented matrix $[A \mid I]$, where I is the identity matrix, whose columns are precisely the standard basis vectors. We then perform row reduction until we arrive at the reduced row echelon form. If A is invertible, then pivots can be found in every column of A , and so the reduced row echelon form of $[A \mid I]$ looks like $[I \mid A^{-1}]$. We then just read off the inverse A^{-1} . If you do not find a pivot in every one of the first n columns of the augmented matrix, then A is not invertible.

This is best seen by example.

Example 7. Find the inverse of the matrix

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

Solution: We write the augmented matrix and we start reducing:

$$\begin{aligned}
 & \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -4 & -5 & -2 & 1 & 0 \\ 0 & -5 & -9 & -3 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & -5 & -9 & -3 & 0 & 1 \end{array} \right] \rightarrow \\
 & \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & -11/4 & -1/2 & -5/4 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array} \right] \rightarrow \\
 & \rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 5/11 & -5/11 & 12/11 \\ 0 & 1 & 0 & 3/11 & -9/11 & 5/11 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/11 & 3/11 & 2/11 \\ 0 & 1 & 0 & 3/11 & -9/11 & 5/11 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array} \right].
 \end{aligned}$$

So

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -1/11 & 3/11 & 2/11 \\ 3/11 & -9/11 & 5/11 \\ 2/11 & 5/11 & -4/11 \end{bmatrix}.$$

└

Not too terrible, no? Perhaps harder than inverting a 2×2 matrix for which we had a formula, but not too bad. Really in practice this is done efficiently by a computer.

Trace and Determinant of Matrices

The next thing to add into our toolbox of matrices is the idea of the trace of a matrix, and how it and the determinant relate to the eigenvalues of said matrix.

Definition 5. Let A be an $n \times n$ square matrix. The trace of A is the sum of all diagonal entries of A .

For example, if we have the matrix

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

the trace is $1 + 2 + 3 = 6$.

The trace is important in our context because it also tells us something about the eigenvalues of a matrix. To work this out, let's consider the generic 2×2 matrix and how we would find the eigenvalues. If we have a 2×2 matrix of the form

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we can write out the expression $\det(A - \lambda I)$ in order to find the eigenvalues. In this case, we would get

$$\det(A - \lambda I) = \det \left(\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} \right) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc).$$

However, the coefficients in this polynomial look familiar. $(ad - bc)$ is just the determinant of the matrix A , and $a + d$ is the trace. Therefore, for any 2×2 matrix, we could write the characteristic polynomial as

$$\det(A - \lambda I) = \lambda^2 - T\lambda + D \tag{1}$$

where T is the trace of the matrix and D is the determinant. On the other hand, assume that r_1 and r_2 are the two eigenvalues of this matrix (whether they be real, complex, or repeated). In that case, we know that this polynomial has r_1 and r_2 as roots. Therefore, it is equal to

$$\det(A - \lambda I) = (\lambda - r_1)(\lambda - r_2) = \lambda^2 - (r_1 + r_2)\lambda + r_1 r_2. \tag{2}$$

Matching up the coefficient of λ and the constant term in (1) and (2) gives the relation that

$$T = r_1 + r_2 \quad D = r_1 r_2,$$

that is, the trace of the matrix is the sum of the eigenvalues, and the determinant of the matrix is the product of the eigenvalues. We only showed this fact for 2×2 matrices, but it does hold for matrices of all sizes, giving us the following theorem.

Theorem 5. Let A be an $n \times n$ square matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, written with multiplicity if needed. Then

- (a) The trace of A is $\lambda_1 + \lambda_2 + \dots + \lambda_n$.
- (b) The determinant of A is $(\lambda_1)(\lambda_2) \cdots (\lambda_n)$.

From the above statement, we note that if any of the eigenvalues is zero, the product of all eigenvalues will be zero, and so the matrix will have zero determinant. This gives an extra follow-up fact, and addition to Theorem ??.

Theorem 6. A matrix A is invertible if and only if all of its eigenvalues are non-zero.

Example 8. Use the facts above to analyze the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

Solution: From the matrix A , we can compute that the trace of A is $1+4 = 5$, and the determinant is $(1)(4) - (2)(5) = -6$. Based on the theorem above, we know that the two eigenvalues of this matrix must add to 5 and multiply to -6 . While you could probably guess the numbers here, the important take-aways from this example are what we can learn.

The main fact to point out is that this is enough information, in the 2×2 case, to tell us that the eigenvalues have to be real and distinct. Since their product is a negative number, we can eliminate the other two options. If we have two complex roots, they must be of the form $x + iy$ and $x - iy$, and so the product is

$$(x + iy)(x - iy) = x^2 + ixy - ixy - i^2y^2 = x^2 + y^2$$

which is always positive, no matter what x and y are. Similarly, if we have a repeated eigvalue, the product will be that number squared, which is also positive. Therefore, if the determinant of a 2×2 matrix is negative, the eigenvalues must be real and distinct, with one being positive and one negative (otherwise the product can not be negative). These facts will be important when we start to analyze the solutions to systems of differential equations in Chapter ??.

Example 9. What can be said about the eigenvalues of the matrix

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 2 & 0 \\ -7 & -3 & -1 \end{bmatrix}?$$

Solution: We can find the same information as the previous example. The trace of A is 1, and the determinant, by cofactor expansion along column 3, is $(-1)(0+2) = -2$. Therefore, the sum of the *three* eigenvalues is 1, and the product of them is -2 . We don't actually have enough information here to determine what the eigenvalues are. The issue is that with three eigenvalues, there are many different ways to get to a product being negative. There could be three negative eigenvalues, two positive and one negative, or one negative real with two complex eigenvalues. However, the one thing we do know for sure is that there must be one negative real eigenvalue. For this particular example, we can compute that the eigenvalues are $-1, 1 + i$, and $1 - i$, so we did end up in the complex case.

Exercise 1 Imagine that we have a 3×3 matrix with a positive determinant (it doesn't matter what the trace is). Think about all the scenarios and verify that at least one eigenvalue must be real and positive for this to happen.

Extension of Previous Theorem

With all of the new definitions and properties that have been stated, we can add a few more equivalent statements to Theorem ??.

Theorem 7. Let A be an $n \times n$ matrix. The following are equivalent:

- (a) A is invertible.
- (b) $\det(A) \neq 0$.

- (g) *The rank of A is n .*
- (h) *The rows of A are linearly independent.*
- (i) *The nullity of the matrix is 0.*
- (j) *None of the eigenvalues of A are 0, or equivalently, the product of the eigenvalues of A is non-zero.*
- (k) *The columns of A are a basis of \mathbb{R}^n .*
- (l) *The rows of A are a basis of \mathbb{R}^n .*

Proof Most of these follow from the components of Theorem ?? . If A is invertible, then we know that the columns are linearly independent. But there are n columns, so that number must be the rank. This implies that the rows are linearly independent, and if the rank plus the nullity must be n , we must have the nullity equal to zero. On that same train of thought, if we have n linearly independent vectors in \mathbb{R}^n , then they must be a basis, giving (k) and (l). Finally, since the determinant is the product of the eigenvalues, if the determinant is non-zero, that implies fact (j). ■