Subspaces, Bases and Dimension

Ryan C. Daileda

Trinity University

Linear Algebra

Introduction

To unify our work with spans and linear independence, we introduce the notion of *subspace*.

We will see that every span is a subspace and that every subspace is a span.

A *minimal* spanning set for a subspace is a *basis*.

The size of a basis turns out to be an important invariant of a subspace known as its *dimension*.

Subspaces

Definition

A subset $H \subseteq \mathbb{R}^n$ is called a *subspace* if:

- $\mathbf{0} \in H$;
- if $\mathbf{u}, \mathbf{v} \in H$, then $\mathbf{u} + \mathbf{v} \in H$;
- if $\mathbf{v} \in H$ and $c \in \mathbb{R}$, then $c\mathbf{v} \in H$.

In this case we write $H \leq \mathbb{R}^n$.

Examples.

- $H = \{\mathbf{0}\}$ and $H = \mathbb{R}^n$ are both subspaces of $\mathbb{R}^n$.
- For any finite $\mathcal{S} \subseteq \mathbb{R}^n$, $\operatorname{Span} \mathcal{S} \leq \mathbb{R}^n$ by an earlier result.
- If A is an $m \times n$ matrix, then $\operatorname{Col} A \leq \mathbb{R}^m$ and $\operatorname{Null} A \leq \mathbb{R}^n$.

Nullspaces are Subspaces

The final statement requires proof. Since $A\mathbf{0} = \mathbf{0}$, we have $\mathbf{0} \in \operatorname{Null} A$.

Suppose $\mathbf{x}, \mathbf{y} \in \operatorname{Null} A$ and $c \in \mathbb{R}$. Then

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

so $\mathbf{x} + \mathbf{y} \in \operatorname{Null} A$, too.

We also have

$$A(c\mathbf{x}) = c(A\mathbf{x}) = c\mathbf{0} = \mathbf{0},$$

which means $c\mathbf{x} \in \operatorname{Null} A$.

These three facts show that $\operatorname{Null} A$ is a subspace of $\mathbb{R}^n$.

Spans are the prototypical examples of subspaces. In fact, they are the *only* examples.

We will say that a subspace $H \leq \mathbb{R}^n$ is *finitely generated* if

$$H = \operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$$

for some vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ (which must necessarily belong to H).

We claim that every subspace $H \leq \mathbb{R}^n$ is finitely generated.

Suppose this is *not* the case. Then for any finite linearly independent set

$$\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq H$$

we must have $\operatorname{Span} \mathcal{S} \subseteq H$ (because H is a subspace) but $\operatorname{Span} \mathcal{S} \neq H$.

So we can choose $\mathbf{v}_{k+1} \in H$ that is *not* in $\operatorname{Span} \mathcal{S}$.

Since $\mathbf{v}_{k+1} \notin \operatorname{Span} \mathcal{S}$ and $\mathcal{S}$ is linearly independent, then so is $\mathcal{S}' = \mathcal{S} \cup \{\mathbf{v}_{k+1}\} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ (there's no way any vector can be a linear combination of those preceding it).

This means that if we start with $\mathcal{S} = \varnothing$, say, we can build linearly independent subsets of H with as many vectors as we like.

But $H \leq \mathbb{R}^n$, and in $\mathbb{R}^n$ linearly independent subsets can have no more than n vectors.

So H *cannot* fail to be finitely generated.

This proves:

Theorem 1

Let $H \leq \mathbb{R}^n$. Then H is finitely generated. That is,

$$H = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

for some $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$.

That is, the *only* subspaces of $\mathbb{R}^n$ are the spans!

Recall that by removing vectors from a spanning set that are linear combinations of the others, we can always arrive at a spanning set that is linearly independent.

So we immediately obtain the following corollary:

Corollary 1

Let $H \leq \mathbb{R}^n$. *Then*

$$H = \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$$

for some linearly independent $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, $m \leq n$.

This leads to the following definition.

Definition

Let $H \leq \mathbb{R}^n$. We say that $\mathcal{B} \subseteq \mathbb{R}^n$ is a *basis* for H provided:

1. $\mathcal{B}$ is linearly independent.
2. $H = \operatorname{Span} \mathcal{B}$.

The corollary to Theorem 1 can now be rephrased as follows: every subspace of $\mathbb{R}^n$ has a basis with at most n vectors.

Example 1

Find a basis for $\operatorname{Null} A$, where

$$A = \begin{pmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{pmatrix}.$$

Solution. We have

$$A \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore the solutions to $A\mathbf{x} = \mathbf{0}$ are given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_3 - x_4 \\ -2x_3 + 4x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 4 \\ 0 \\ 1 \end{pmatrix}$$

with x_3 and x_4 free. Thus

$$\operatorname{Null} A = \operatorname{Span} \underbrace{\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ 0 \\ 1 \end{pmatrix} \right\}}_{\mathcal{B}}.$$

Since the vectors on the RHS are not multiples of one another, they are linearly independent. Thus $\mathcal{B}$ is a basis for $\operatorname{Null} A$. $\square$

Completing a Linearly Independent Set

Let $H \leq \mathbb{R}^n$. Starting with a spanning set for H, we can remove appropriate vectors to obtain a basis for H.

On the other hand, suppose we start with a linearly independent set $\mathcal{S} \subseteq H$.

Then $\operatorname{Span} \mathcal{S} \subseteq H$ (since H is closed under vector addition and scalar multiplication).

If $\operatorname{Span} \mathcal{S} \neq H$, we can choose a vector $\mathbf{v} \in H$ that *does not* belong to $\operatorname{Span} \mathcal{S}$.

Then the set $\mathcal{S}' = \mathcal{S} \cup \{\mathbf{v}\}$ must be linearly independent as well (no vector is a linear combo. of those preceding it), and $\operatorname{Span} \mathcal{S}' \subseteq H$.

Now replace $\mathcal{S}$ with $\mathcal{S}'$ and repeat.

Because $|\mathcal{S}| \leq n$, this process cannot go on forever. That is, eventually we will have $\operatorname{Span} \mathcal{S} = H$ with $\mathcal{S}$ linearly independent.

Theorem 2

Let $H \leq \mathbb{R}^n$.

1. *If $\operatorname{Span} \mathcal{S} = H$, then $\mathcal{S}$ contains a basis for H.*
2. *If $\mathcal{S} \subseteq H$ is linearly independent, then H has a basis containing $\mathcal{S}$.*

Remarks

Let $H \leq \mathbb{R}^n$. Theorem 2 states that:

- Every spanning set (of H) contains a basis (for H).
- Every linearly independent set (in H) can be *completed* to a basis (for H).

These two (complementary) facts can be extremely useful!

Dimension

Every subspace of $\mathbb{R}^n$ has a basis. As we will now see, the number of vectors in a basis is invariant.

Let $H \leq \mathbb{R}^n$ and let $\mathcal{B}$ be a basis for H with $|\mathcal{B}| = m \leq n$.

Suppose $\mathcal{S} \subseteq H$ is a linearly independent subset of H.

Then we know that

$$[\mathcal{S}]_{\mathcal{B}} \subseteq \mathbb{R}^m$$

is a linearly independent subset of $\mathbb{R}^m$.

This means that $|\mathcal{S}| = |[\mathcal{S}]_{\mathcal{B}}| \leq m$. Thus:

Theorem 3

Let $H \leq \mathbb{R}^n$. If H has a basis of size m, and $\mathcal{S} \subseteq H$ is linearly independent, then

$$|\mathcal{S}| \leq m.$$

Suppose that $H \leq \mathbb{R}^n$ has a basis $\mathcal{B}$ with $|\mathcal{B}| \leq n$.

Suppose $\mathcal{C} \subseteq H$ is another basis for H. Note that $\mathcal{C}$ cannot contain more than n vectors.

Because $\mathcal{B}$ is a basis and $\mathcal{C}$ is linearly independent, Theorem 3 tells us that

$$|\mathcal{C}| \leq |\mathcal{B}|.$$

Likewise, since $\mathcal{C}$ is a basis and $\mathcal{B}$ is linearly independent, we must have

$$|\mathcal{B}| \leq |\mathcal{C}|.$$

That is, $|\mathcal{B}| = |\mathcal{C}|$.

Theorem 4

Let $H \leq \mathbb{R}^n$. Then H has a finite basis, and all bases of H have the same size $m \leq n$.

The number m in the theorem is called the *dimension* of H:

$$\dim H = \ \# \text{ of vectors in every basis of } H.$$

Example 2

Compute $\dim \operatorname{Null} A$ where

$$A = \begin{pmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{pmatrix}.$$

Solution. We saw in Example 1 that Null A has a basis with two vectors. Thus

$$\boxed{\dim \operatorname{Null} A = 2}.$$

□

Example 3

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

Solution. If $\mathbf{x} \in \mathbb{R}^n$, notice that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Let

$$\mathbf{e}_j = (\delta_{ij}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

which has a 1 in the jth entry and zeros elsewhere. Then

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \in \mathrm{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}.$$

That is,

$$\mathbb{R}^n = \mathrm{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}.$$

Since the matrix $A = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{pmatrix}$ is in reduced echelon form and has a pivot in every column, we conclude that $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is linearly independent.

Hence $\mathcal{B}$ is a basis for $\mathbb{R}^n$, so that

$$\dim \mathbb{R}^n = |\mathcal{B}| = n.$$

□

Remark. The basis

$$\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$$

is called the *standard basis* for $\mathbb{R}^n$.

Dimensions of Column and Null Spaces

Given an matrix A, we will be particularly interested in computing

$$\operatorname{rank} A = \dim \operatorname{Col} A \quad \text{and} \quad \dim \operatorname{Null} A.$$

To do this, we need to find bases for the subspaces $\operatorname{Col} A$ and $\operatorname{Null} A$.

Since $\operatorname{Col} A$ is the span of the columns of A, we know that by discarding certain columns we will be left with a basis.

But which columns do we discard?

This can be determined through row reduction!

Let U be the reduced echelon form of A. Then the equations

$$U\mathbf{x} = \mathbf{0} \quad \text{and} \quad A\mathbf{x} = \mathbf{0}$$

have *exactly the same solutions*.

This tells us two things:

- $\operatorname{Null} A = \operatorname{Null} U$.
- The columns of U and the columns of A have the same dependence relations.

The pivot columns of U are clearly linearly independent, and every non-pivot column is a linear combination of the pivot columns to its left.

This means the same is true of the columns of A. So if we discard the non-pivot columns of A we will be left with a basis for $\operatorname{Col} A$.

Theorem 5

Let A be an $m \times n$ matrix. The pivot columns of A form a basis for $\operatorname{Col} A$. Thus:

$$\operatorname{rank} A = \dim \operatorname{Col} A = \# \text{ pivot columns of } A.$$

Remark. Because $\operatorname{Col} A \leq \mathbb{R}^m$, we must have $\operatorname{rank} A \leq m$.

Example

Let's illustrate with an example. We have

$$A = \begin{pmatrix} 1 & 3 & 2 & -6 & -6 \\ 3 & 9 & 1 & 5 & 10 \\ 2 & 6 & -1 & 9 & 14 \\ 5 & 15 & 0 & 14 & 24 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 3 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = U$$

The second column of U is a linear combination (multiple) of the first, and the final column of U is a linear combination of the first, third and fourth.

And the pivot columns of U are clearly linearly independent.

So the same statements are true of the columns of A.

So to get a basis for $\operatorname{Col} A$ we can take the first, third and fourth columns of A:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ 5 \\ 9 \\ 14 \end{pmatrix} \right\}.$$

Remark. We must use the pivot columns of A, *not* those of its reduced echelon form, to get a basis for $\operatorname{Col} A$.

What about the null space of A?

Any nonzero row of U has the form

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & c_{i+1} & \cdots & c_n \end{pmatrix},$$

where the 1 is in the ith column.

In the (reduced) equation $U\mathbf{x} = \mathbf{0}$, this corresponds to an equation of the form

$$x_i + c_{i+1}x_{i+1} + \cdots c_n x_n = 0 \quad \Leftrightarrow \quad x_i = -c_{i+1}x_{i+1} - \cdots - c_n x_n.$$

This means that every basic variable x_i can be expressed in terms of the free variables x_j with $j > i$.

So when we write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

parametrically, the free variable x_j can only occur in the entries for which $i \leq j$. That is, x_j can only occur among the first j entries. Hence

$$\mathbf{x} = \sum_{\substack{j \\ x_j \text{ is free}}} x_j \begin{pmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j\text{th position.} \\ \\ \\ \\ \end{matrix}$$

Let

$$\mathbf{v}_j = \begin{pmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{th position.}$$

Because the x_j are free, we conclude that

$$\operatorname{Null} A = \operatorname{Span}\{\mathbf{v}_j \mid x_j \text{ is free}\}.$$

Because $\mathbf{v}_j$ has a nonzero jth entry, but no earlier $\mathbf{v}_i$ does, $\mathbf{v}_j$ cannot be a linear combination of the $\mathbf{v}_i$ preceding it.

This means the set of $\mathbf{v}_j$ is linearly independent! So we have a basis for $\operatorname{Null} A$.

Theorem 6

Let A be an $m \times n$ matrix. Then $\dim \operatorname{Null} A$ is the number of free variables in $A\mathbf{x} = \mathbf{0}$. Equivalently,

$$\dim \operatorname{Null} A = \# \text{ of non-pivot columns of } A.$$

Example

Let's return to the matrix

$$A = \begin{pmatrix} 1 & 3 & 2 & -6 & -6 \\ 3 & 9 & 1 & 5 & 10 \\ 2 & 6 & -1 & 9 & 14 \\ 5 & 15 & 0 & 14 & 24 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 3 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = U$$

It has two non-pivot columns, and therefore

$$\dim \operatorname{Null} A = 2.$$

The solutions to $A\mathbf{x} = \mathbf{0}$ are given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -3x_2 - 2x_5 \\ x_2 \\ x_5 \\ -x_5 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -2 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

So a basis for Null A is

$$\mathcal{B} = \left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

□

We now make a trivial, but fundamental, observation. If A is $m \times n$, then

$$\begin{aligned} n &= \#\text{ cols. of } A \\ &= (\#\text{ pivot cols. of } A) + (\#\text{ non-pivot cols. of } A) \\ &= \operatorname{rank} A + \dim \operatorname{Null} A. \end{aligned}$$

This is known as the *rank theorem*.

The Rank Theorem

Theorem 7 (The Rank Theorem)

If A is an $m \times n$ matrix, then

$$n = \operatorname{rank} A + \dim \operatorname{Null} A.$$

Example 4

Let A be a 5×7 matrix. Suppose the equation $A\mathbf{x} = \mathbf{0}$ has 3 linearly independent solutions. How many linearly independent columns can A have?

Solution. The hypothesis implies that

$$3 \leq \dim \operatorname{Null} A.$$

The Rank Theorem then tells us that

$$7 = \operatorname{rank} A + \dim \operatorname{Null} A \geq \operatorname{rank} A + 3.$$

Hence

$$\operatorname{rank} A \leq 7 - 3 = 4.$$

So A can have no more than 4 linearly independent columns. □

Remark. In general, the Rank Theorem shows that a *lower* bound for $\dim \operatorname{Null} A$ yields an *upper* bound for $\operatorname{rank} A$, and vice versa.

The Basis Theorem

Let $H \leq \mathbb{R}^n$ with $\dim H = m \leq n$, and suppose $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq H$.

If $\operatorname{Span} \mathcal{S} = H$, then we can remove vectors from $\mathcal{S}$ (if necessary) to obtain a basis for H.

But every basis for H has exactly m vectors, so $\mathcal{S}$ must already be a basis.

Likewise, if $\mathcal{S}$ is linearly independent, we can add vectors to $\mathcal{S}$ (if necessary) to obtain a basis for H.

The same reasoning shows that $\mathcal{S}$ is a basis for H.

Thus:

Theorem 8 (The Basis Theorem)

Let $H \leq \mathbb{R}^n$ and suppose $\mathcal{S} \subseteq H$.

1. *If $|\mathcal{S}| = \dim H$ and $\operatorname{Span} \mathcal{S} = H$, then $\mathcal{S}$ is a basis for H.*
2. *If $|\mathcal{S}| = \dim H$ and $\mathcal{S}$ is linearly independent, then $\mathcal{S}$ is a basis for H.*

Remark. This result tells us that if $\mathcal{S} \subseteq H \leq \mathbb{R}^n$ contains the "right number" of vectors, then we only have to check "half" of the definition to show that $\mathcal{S}$ is a basis for H.

Example 5

Let $H \leq \mathbb{R}^n$. Show that if $\dim H = n$, then $H = \mathbb{R}^n$.

Solution. Let $\mathcal{B}$ be a basis for H.

Then $|\mathcal{B}| = \dim H = n$.

Since $\mathcal{B}$ is linearly independent and $|\mathcal{B}| = n = \dim \mathbb{R}^n$, the Basis Theorem tells us that $\mathcal{B}$ is a basis for $\mathbb{R}^n$.

Thus

$$H = \operatorname{Span} \mathcal{B} = \mathbb{R}^n.$$

□