Subspaces

A 'subspace' of a vector space must follow two basic rules to be considered as a subspace:

① **Definition 8.7** (Subspace). Let V be a vector space over a field F. A ***subspace*** W of V is a subset of V which itself forms a vector space with the same addition and scalar multiplication as V.

i.e vectors need to represent the same thing, inside the subspace as well as outside.

② **Theorem 8.9** (Subspace criterion). *Let V be a vector space over a field F. Then a* non-empty *subset W of V is a subspace of V if the following hold:*

(a) $v + w \in W$ *for all* $v, w \in W$;

(b) $\alpha v \in W$ *for all* $v \in W$ *and* $\alpha \in F$.

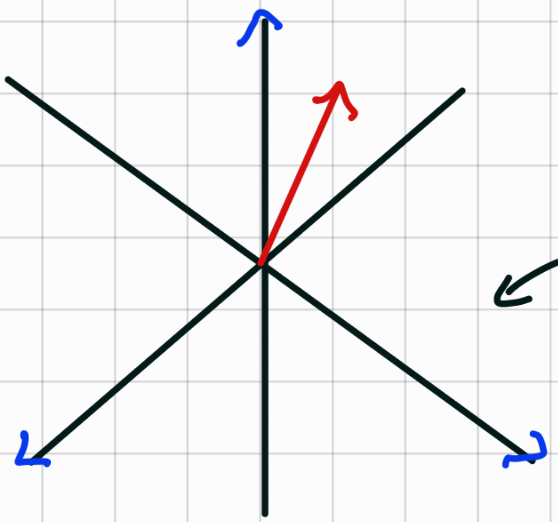
The important one:

- The subspace criteroen.

i.e vector subspaces be closed within their subspace under our fundamental operations.

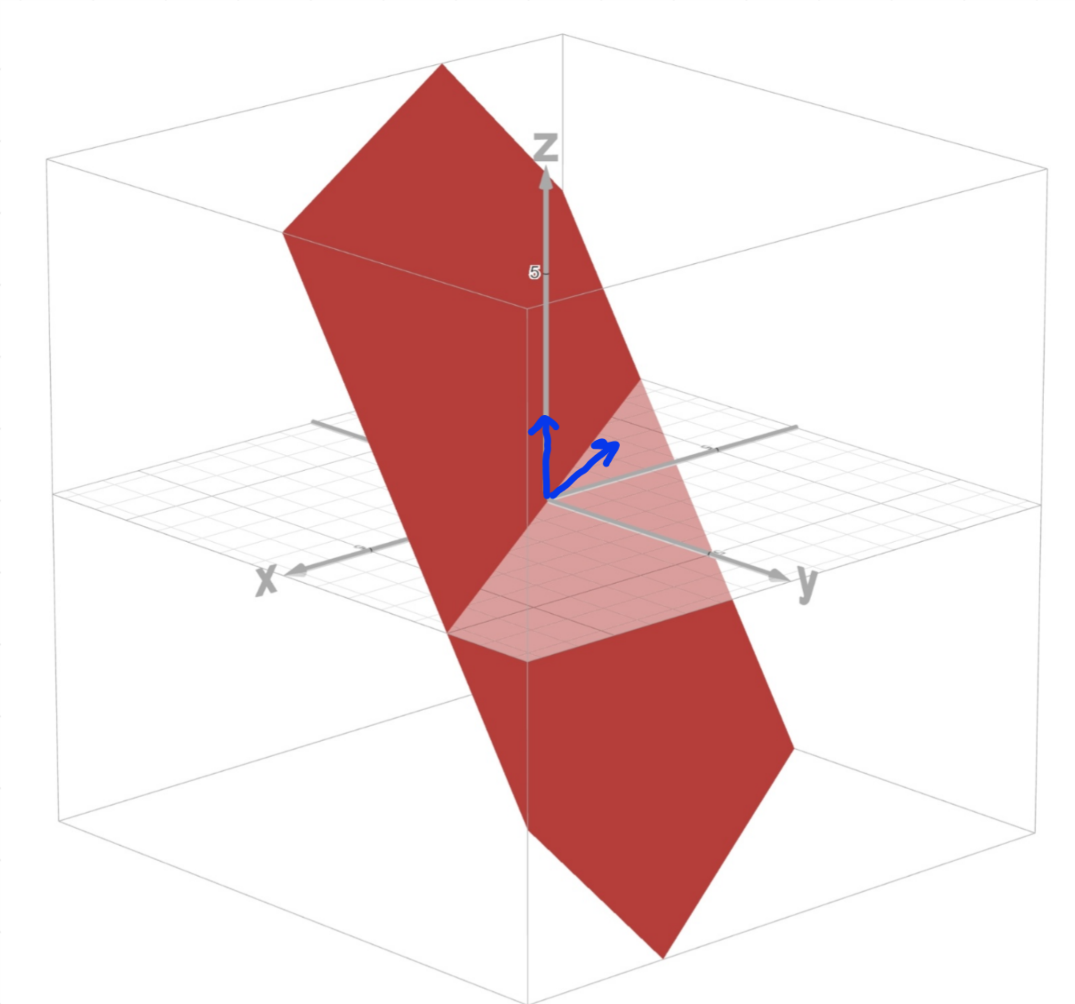
These two conditions mean that subspaces can be thought of as 'lower dimensional' spaces nested in higher dimensional ones.

To do an example, take a vector space, $V \in \mathbb{R}^3$. (i.e in 3D).



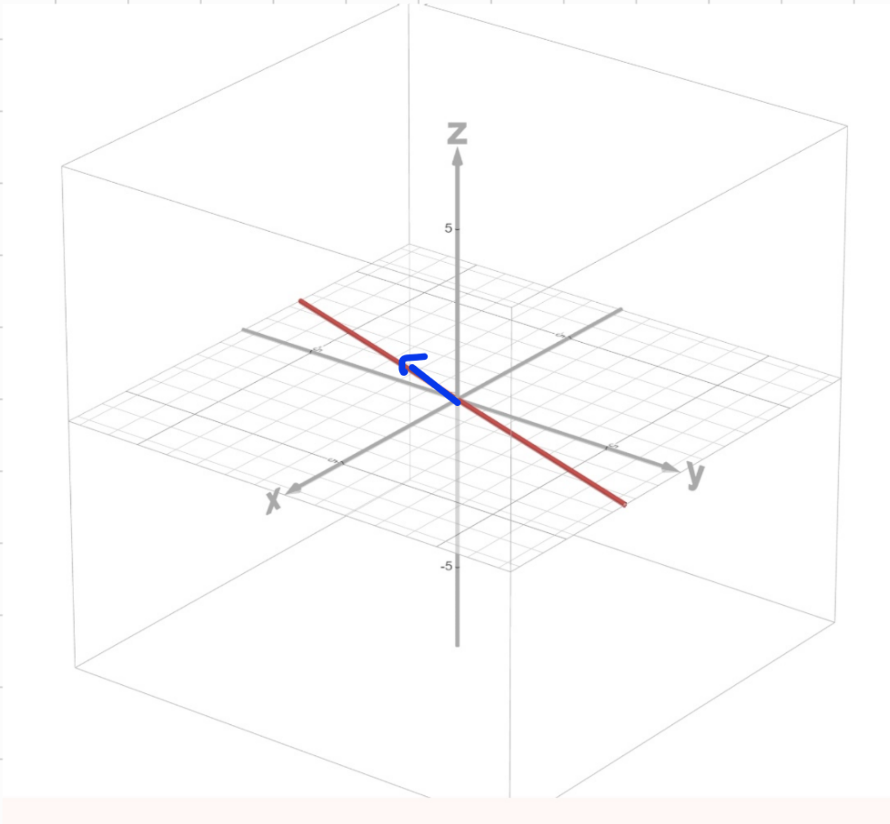
vectors live here.

Take the span of two vectors which don't lie on top of each. We know that they would span a plane.



The important thing to notice is that, if we were only allowed to pick vectors from this plane, there's no way to get off this plane. i.e, it's closed.

The same thing would be true of the span of two 2D vectors in $\mathbb{R}^3$:



- If you're only allowed to add and scale vectors on this line, there is no way to get of it.

- This gives us a lovely visual view of subspaces. For a space in $\mathbb{R}^n$, a subspace is any restricted space inside with $\dim(S) < n$, with the effect that this essentially 'traps' vectors.

This visual interpretation also means we can prove some easy facts.

1. If V is a vector space and $\underline{0}$ is the zero-vector in V, then $\underline{0}$ belongs to every subspace in V.

- Makes sense, all vectors must be rooted at $\underline{0}$, so $\underline{0}$ is in every subspace.

2. In any vector space V $\{0\}$ is always a subspace, as is V itself. As our definition doesn't stop the subspace being the same dimension as the space. The compromise is that the only subspace with the same dimension as the space, is the space itself. There's nothing to limit it.

i.e the only possible subspaces of $\mathbb{R}^3$ are: $\mathbb{R}^3$ itself, $\{0\}$, lines through the origin, and planes through the origin.

3 Principle of dimension monotincity

Theorem 11.7 (Dimension monotonicity). *Let V be a vector space of dimension n. Let U be a subspace of V. Then,*

(a) *The dimension of U is at most n.*

(b) *If $U \neq V$, the dimension of U is less than n.*

(c) *If the dimension of U is less than n, then $U \neq V$.*

4 The Span of any set of vectors in the vector space is always in the vector space.

$$W = \mathrm{Span}(v_1, \dots, v_k).$$

By definition, $W \subseteq V$.
Now, to show that W is a **subspace**, they use the **Subspace Criterion**:

- (a) Closed under addition.
- (b) Closed under scalar multiplication.

(b) Scalar multiplication closure

Suppose $w \in W$. Then w can be written as

$$w = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

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

Take a scalar $\beta \in F$. Then

$$\beta w = (\beta\alpha_1)v_1 + \cdots + (\beta\alpha_k)v_k,$$

which is **still a linear combination** of $v_1, \dots, v_k$.
So $\beta w \in W$.

(b) Scalar multiplication closure

Suppose $w \in W$. Then w can be written as

$$w = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

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

Take a scalar $\beta \in F$. Then

$$\beta w = (\beta\alpha_1)v_1 + \cdots + (\beta\alpha_k)v_k,$$

which is **still a linear combination** of $v_1, \dots, v_k$.
So $\beta w \in W$.

(a) Addition closure

Take two vectors $w, v \in W$. Write

$$w = \alpha_1 v_1 + \cdots + \alpha_k v_k, \quad v = \beta_1 v_1 + \cdots + \beta_k v_k.$$

Then

$$w + v = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_k + \beta_k)v_k,$$

which is again a linear combination of $v_1, \dots, v_k$.
So $w + v \in W$.

Conclusion

Since $W = \mathrm{Span}(\mathcal{A})$ is closed under **scalar multiplication** and **addition**, it satisfies the Subspace Criterion.
Therefore, $\mathrm{Span}(\mathcal{A})$ is a **subspace** of V.

Now, why we bring up the concept of subspaces should hopefully be quite evident. A $\det(A) = 0$ squishes space into a lower dimension of some kind, hence by understanding subspaces created by linear transforms, we can talk about the $\det(A) = 0$ cases more easily.

Subspaces from a matrix

① The Column Space.

Returning to the amount of squish transforms have, we can define some terms.

> The column space is the space that the matrix outputs to. This makes sense, because the span of the columns tells you where the vectors can land.

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

will map all vectors onto the span of

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

Another way of phrasing 'output space', you can also phrase this as the subspace by which:

$$\boxed{\begin{array}{c} A\underline{x} = \underline{b}, \quad \underline{b} \in \mathbb{F}^m \\ \text{has solutions.} \end{array}}$$

(i.e, it's all $\underline{b}$'s which have an $\underline{x}$ which is mapped onto them by A. which is equivalent to an output space.

Knowing this, we can prove the column space is a subspace:

Example 8.13. Let $A \in M_{m\times n}(F)$ and let

$$S := \{y \in F^m : \text{there is an } x \in F^n \text{ such that } Ax = y\} \subseteq F^m.$$

Show that S is a subspace of F^m.

Solution: Since $A0 = 0$ it follows that $0 \in S$, and so S is non-empty. Let $y, z \in S$. Then there are $x, t \in F^n$ so that

$$Ax = y \qquad \text{and} \qquad At = z.$$

We compute

$$y + z = Ax + At = A(x + t)$$

so $y + z \in S$. Let $\alpha \in F$. Then

$$\alpha y = \alpha(Ax) = A(\alpha x)$$

so $\alpha y \in S$. Hence S is a subspace of F^m. $\square$

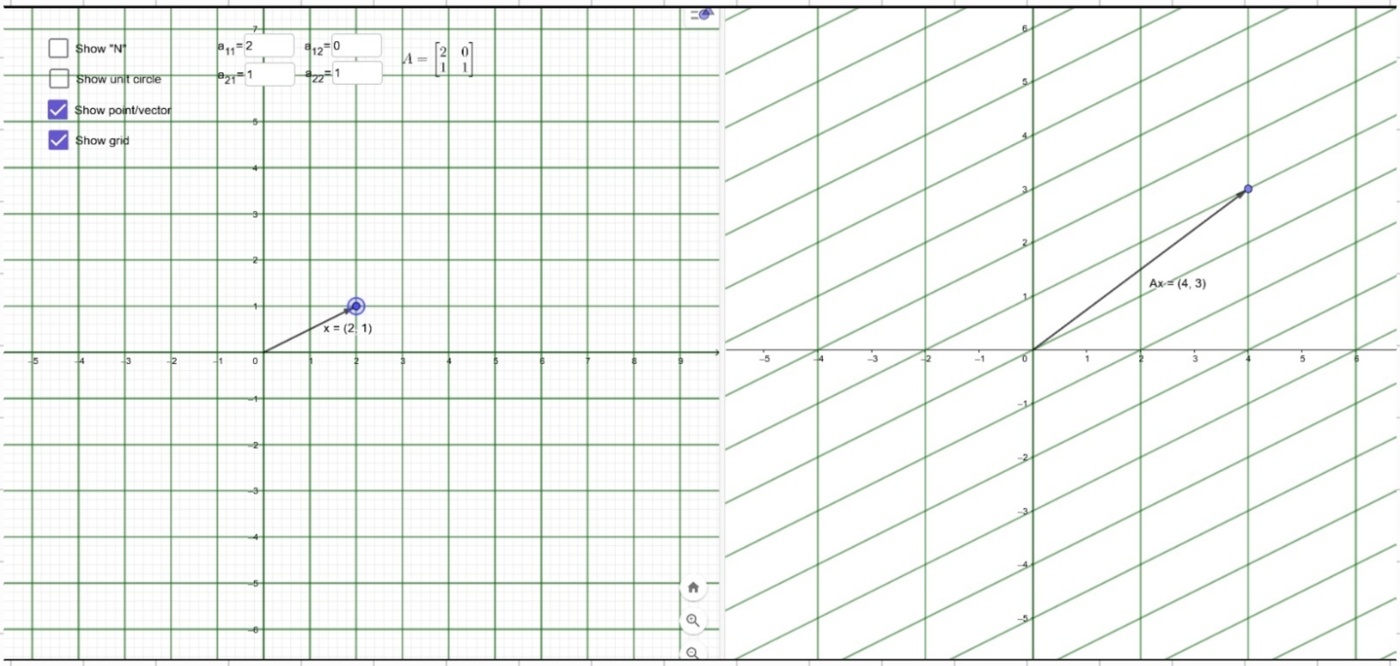
To really define how squished a trans-formation makes space, we can define a new term:

$$\boxed{\text{The rank is the dimension of the column space, so the amount of dimensions in the output.}}$$

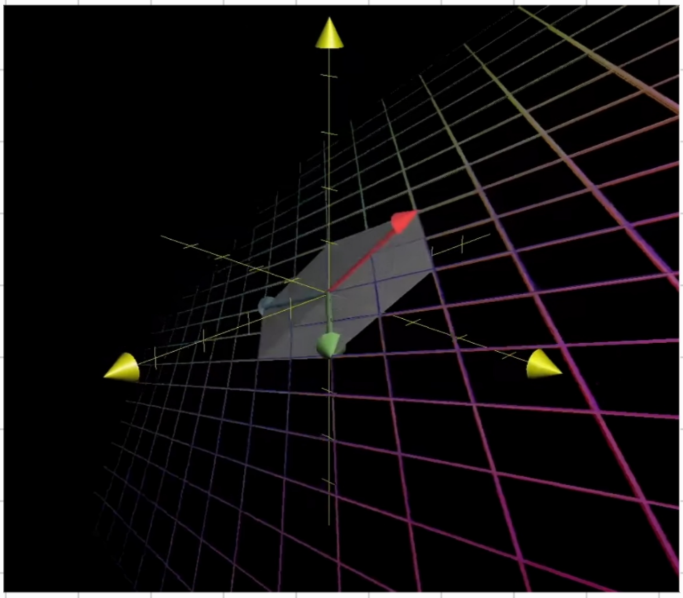
So, if the rank = 1, space is squished into a line. If rank = 2, space is a plane, etc.

For instance, in the case of 2×2 matricies, rank = 2 is the best that it can be, as the basis vectors continue to span $\mathbb{R}^2$ (i.e column space continues being $\mathbb{R}^2$.)

e.g



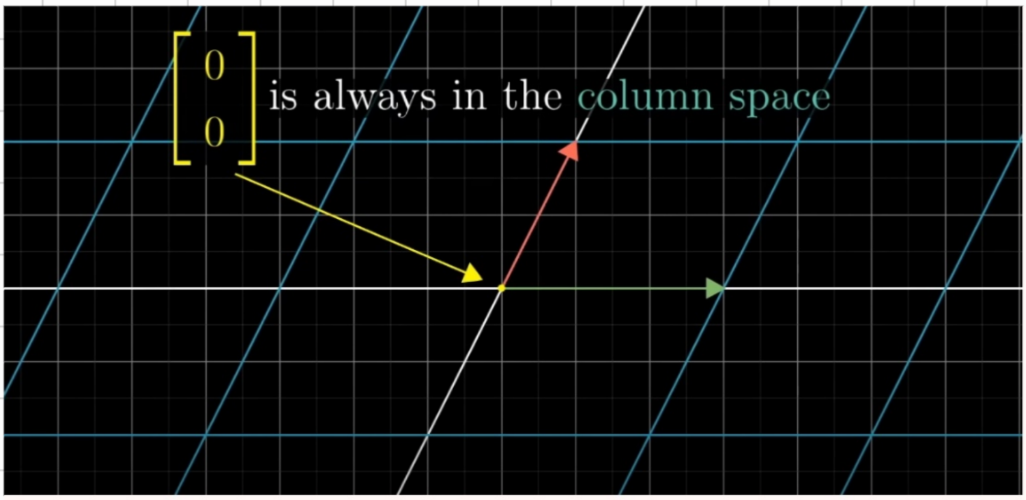
But for a 3×3 matrix, rank = 2 means we've collapsed to a plane.



But not as much as a rank = 1 would. If a 3×3 matrix has $\det(A) \neq 0$, it's column space will continue to be in $\mathbb{R}^3$, and it will have a rank = 3.

When the rank is as high as it can be, i.e all vectors are linearly independant, i.e rank = no. of columns, we call the matrix, full rank.

For a full-rank transform, the only vector that lands on the origin, is the $\underline{0}$ vector itself.



But for matricies which aren't full rank, you will have a whole bunch of vectors which will land on the $\underline{0}$ vector.

This leads to our next related subspace: The right null space.

② The Right Null Space.

The right null-space is the space of vectors in the input space that get mapped to $\underline{0}$.

i.e we can express it as the set of vectors $\underline{x} \in \mathbb{F}^n$, s.t

$$A\underline{x} = \underline{0}$$

(A system where the RHS is $\underline{0}$ is also called homogenous).

This is where the 'right' comes from, as we're trying to find the set of vectors, $\underline{x}$, that map to $\underline{0}$ when A is right-multiplied by $\underline{x}$.

We can also quite easily prove the right-null space is a subspace:

Example 8.12. Let $A \in M_{m\times n}(F)$ and let

$$S := \{x \in F^n : Ax = 0\} \subseteq F^n.$$

Show that S is a subspace of F^n.

Solution: S is non-empty since $A0 = 0$ and so $0 \in S$. For $x, y \in S$, $Ax = 0$ and $Ay = 0$ by definition. Then $A(x + y) = Ax + Ay = 0 + 0 = 0$, so $x + y \in S$. Let $\alpha \in \mathbb{R}$. Then $A(\alpha x) = \alpha(Ax) = \alpha 0 = 0$. Hence S is a subspace of F^n. □

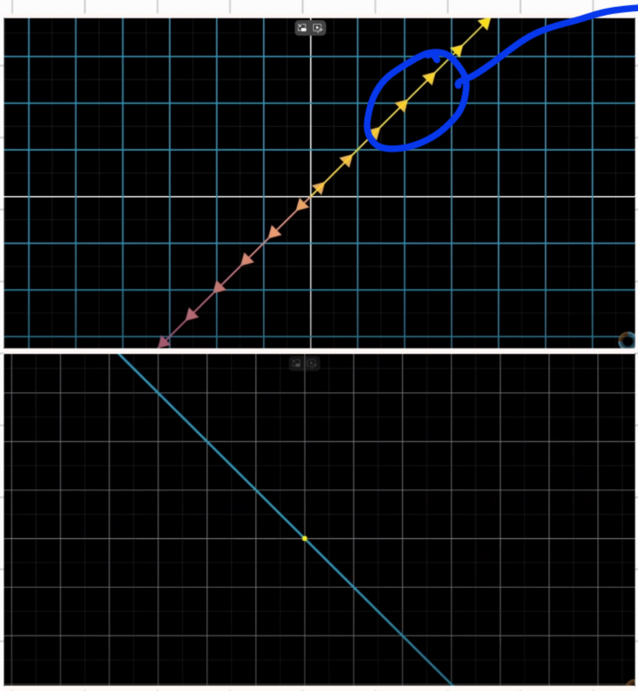
In other words, the set of solutions to a homogeneous system of linear equations is a subspace.

Now, in the case of a full-rank transform, the only vector mapped to $\underline{0}$, would be $\underline{0}$, as the null space is trivial (boring).

But, in the non-full rank cases, the dimensions being squished means that a certain direction of vectors must be squished to $\underline{0}$

e.g

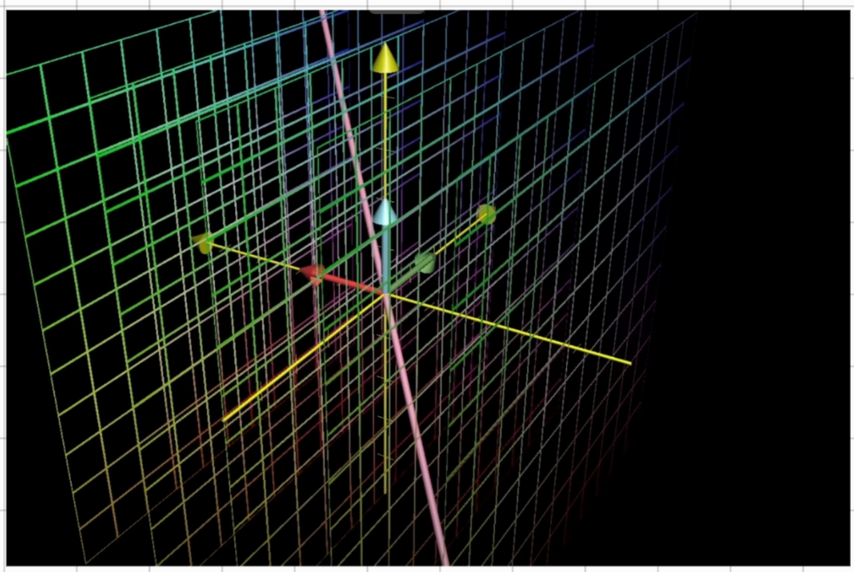
For a 2×2, rank 1 lin-transformation, that maps vectors to a Line, there will be a whole seperate Line of vectors in the input space which Land on $\underline{0}$.

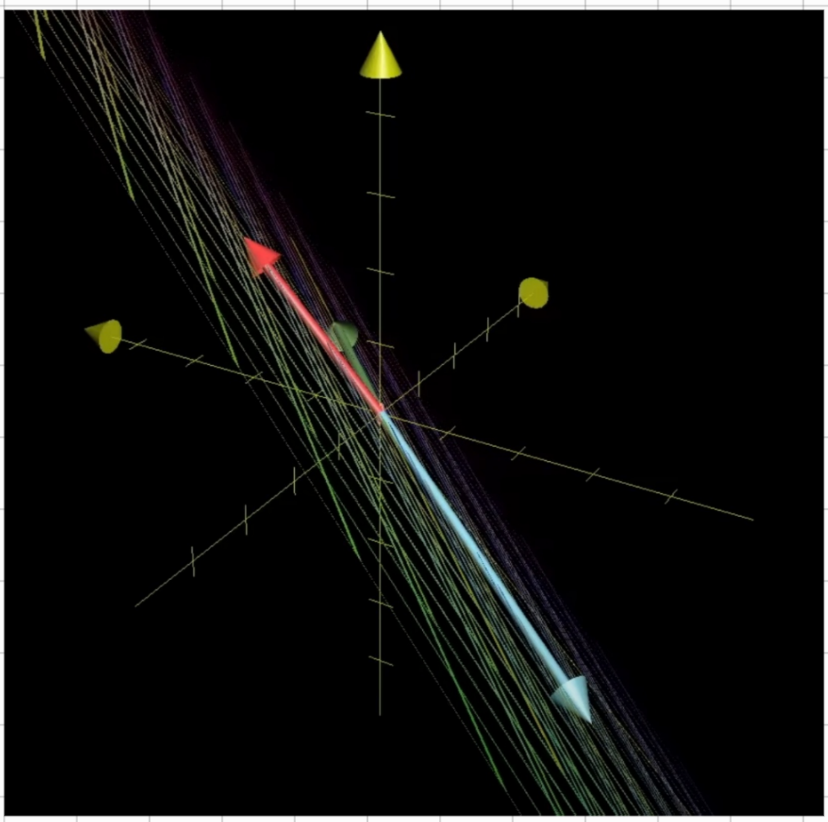


null space.

e.g

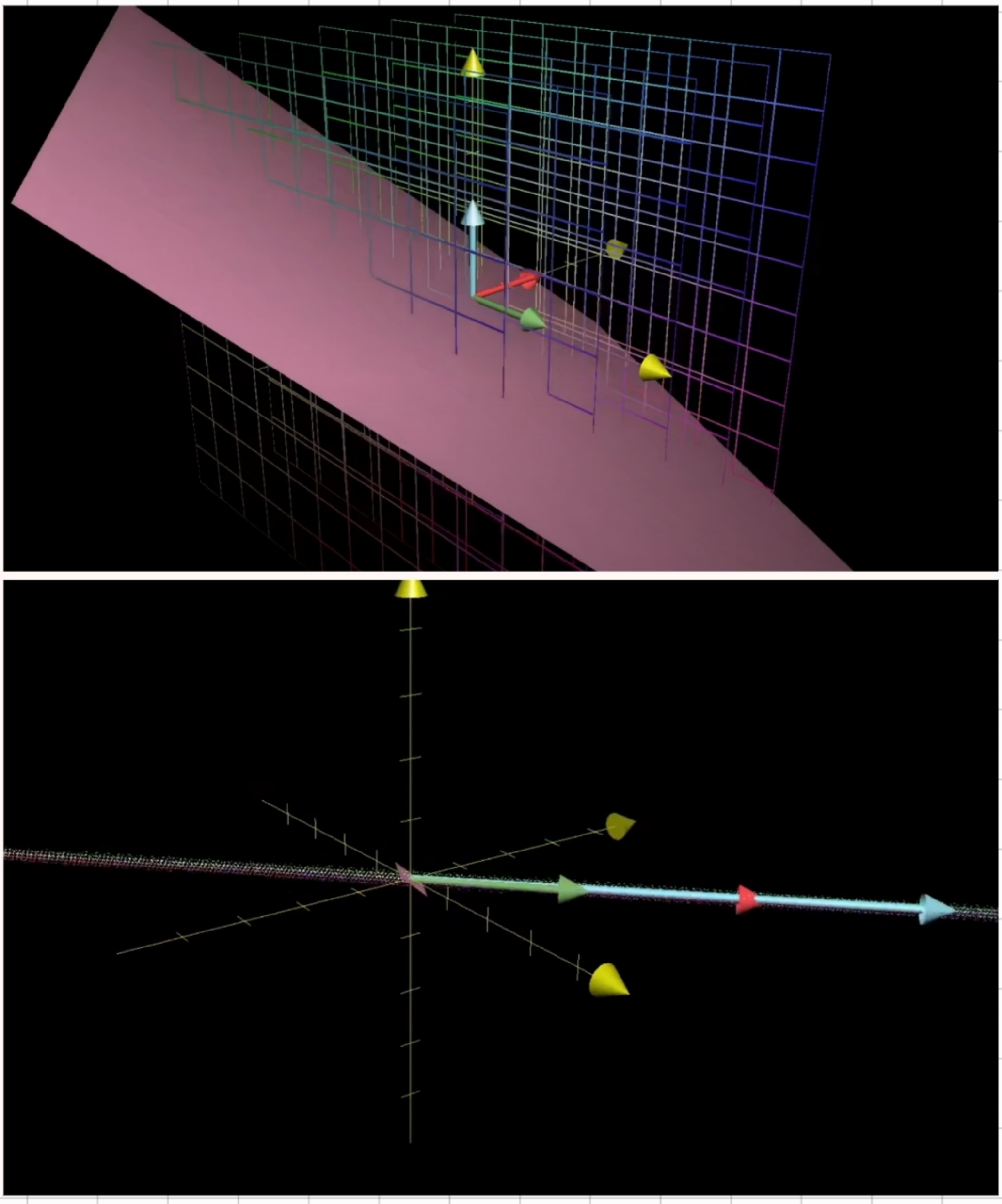
or if a 3×3 matrix is rank 2, i.e squishes into a plane, there will be a line of vectors, which get mapped to $\underline{0}$.





e.g again...

or if you have a 3×3 which squishes to a line, a whole plane of vectors will get squished to $\underline{0}$.



This is when the null-space becomes a useful construct.

The dimension of the right null space is referred to as the nullity.

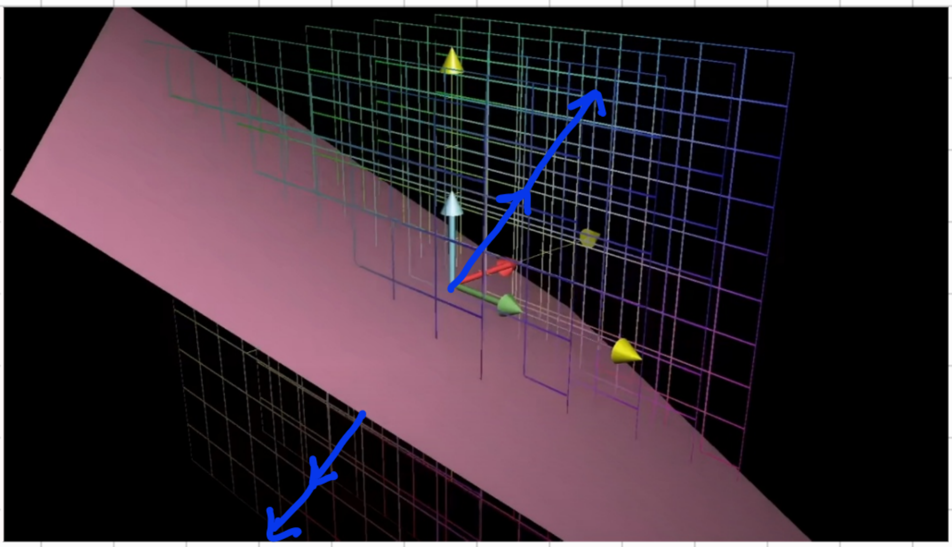
You may notice that the rank and nullity seem to have a correspondance.

This is known as the rank-nullity theorem.

Rank-Nullity

You may notice, that for every extra dimension flattened by the rank, the nullity grows by a dimension.

This makes sense, and we can see it visually:



Compressing down a dimension, means that a certain line of vectors must be crushed onto $\underline{0}$. Do this more, and you end up with a plane crushed to $\underline{0}$, & so on...

∴ the no. of dims lost = the null space.

$$\text{no. of dims lost} = \underset{\text{no. of columns}}{n} - \underset{\text{rank}}{\text{rank}(A)}$$

The makes sense as, if the matrix was full rank, all the columns would the LI, and hence each would correspond to it's own independent direction. By minusing rank(A), the no. of dims in the actual output, we find the amount of dims lost.

Because:

$$\text{no. of dims lost} = \text{nullity}(A)$$

$$\text{nullity}(A) = n - \text{rank}(A)$$

$$\boxed{\Rightarrow \text{nullity}(A) + \text{rank}(A) = n}$$

This is the rank-nullity theorem.

As we can see, the column space, rank, right null space, and nullity are useful subspaces to describe rank(A) = 0 systems.

Beyond this, there are also some other subspaces we can consider.

③ The Row Space

The column space of A^T.

A unique fact is that the dimension of the row space is the same as the dimension of the column space (the rank).

we will have to prove this later, but visually, it can be seen quite nicely.

It links back to what we were saying about having a 'row' perspective and 'column' perspective of matrix-vector multiplication.

The 'row' perspective instead interprets the output as a list representing the dot-product with each row. If a row is LI, then by extension, each number tells you how much the vector points in the unique direction specified by the row-vector. But if it is LD, we know this row-vector we're testing against is on the span of the others, and the dimension of the output drops. Since this describes the same transformation as the columns, the dimension must be the same.

$$\dim(\mathrm{colm}(A)) = \dim(\mathrm{row}(A))$$

$$= \mathrm{rank}(A).$$

4 The Left Null Space

The right null space of A^T.

via the rank-nullity theorem, $\dim(\text{Left Null}) = m - r$, where m is rows instead of colms.

Together, these form what are known as the four-fundamental subspaces.

Non-Square Systems

So far, we have only looked at systems which have the same number of equations as the number of variables. But there's no reason that we can't represent systems which don't fit this:

e.g
$$2x + 3y + 1z = 1$$
$$1x + 2y + 2z = 6$$

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

or e.g

$$x + y = 2$$
$$x - y = 0$$
$$y = -2$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$$

We can consider each of these possibilties by whether the matrix is overdetermined (more eq.'s than unknowns → long) or underdetermined (more unknowns than eq's → tall).

Overdetermined

(more eq. than unknowns)

This corresponds to 'tall' matricies.

$$x + y = 2$$
$$x - y = 0$$
$$y = -2$$

The reason why these are called 'overdetermined' is that, essentially, they have too many equations.

If we look at the line-view for this system: