Orientation:

This course roughly goes as follows:

(1) Solve $Ax=b$ ✓

(2) Almost solve $Ax=b$ ☆

(3) Solve $Ax=\lambda x$

(4) SVD

We've done pretty much all we'll do for (1).

Now we start on (2). The method of least squares is a geometric construction, so we'll spend the next 3 weeks on geometry (subspaces, orthogonality, projections).

Please bear with me for a couple of weeks of theory before we start to see applications on the horizon.

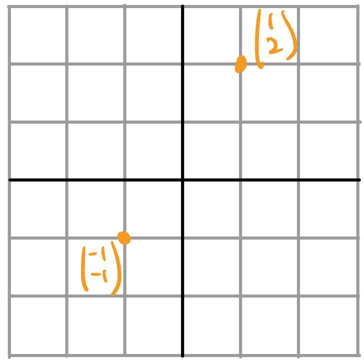
Geometry of Vectors

Recall: A vector in $\mathbb{R}^n$ is just a list of n numbers:

$$v = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

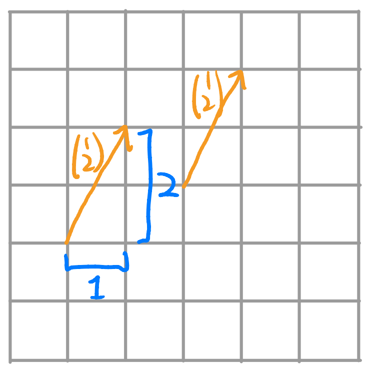
We can draw a vector as a dot / point in Euclidean space:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\text{-coordinate} \\ y\text{-coordinate} \end{pmatrix}$$



We will often consider a vector as an arrow: it measures the difference between two points.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\text{-displacement} \\ y\text{-displacement} \end{pmatrix}$$



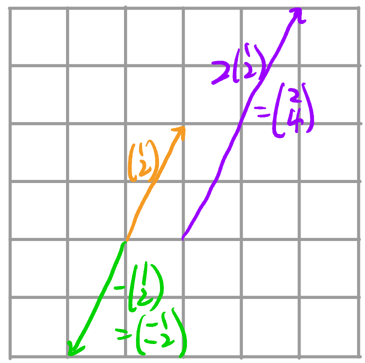
NB: The tail of the arrow can be anywhere. By default we'll draw it at the origin.

[DEMO]

How do algebraic operations behave geometrically?
We'll describe them by drawing arrows.

Scalar Multiplication

- The length of $c \cdot v$ is
  $|c| \times (\text{length of } v)$
- The direction of $c \cdot v$ is
  → the same as v if $c > 0$
  → opposite if $c < 0$



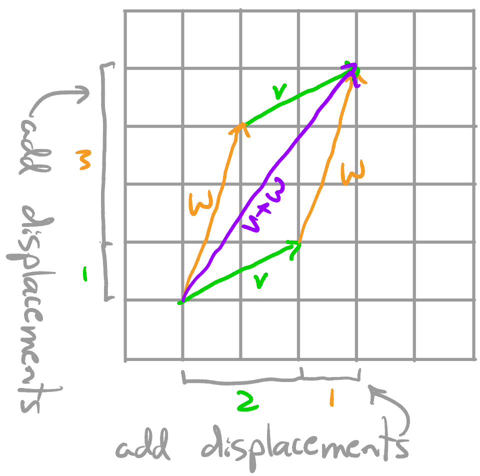
[DEMO]

Vector Addition

Adding $v+w$ just adds the displacements.

$$v = \binom{2}{1} \qquad w = \binom{1}{3} \qquad v + w = \binom{3}{4}$$

Parallelogram Law: to draw $v+w$, move the tail of v to the head of w (or vice-versa). Now $v+w$ is the arrow from the tail of w to the head of v.



[DEMO]

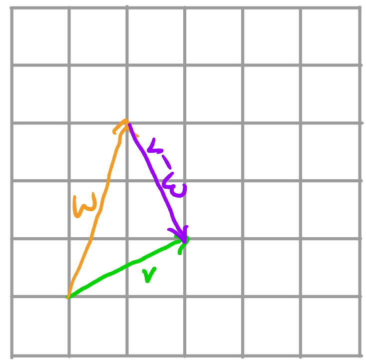
Vector Subtraction

$w + (v - w) = v$

so $v-w$ starts at the head of w and ends at the head of v.
(According to the parallelogram law above.)

$v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ $v - w = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

$w = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$



[DEMO]

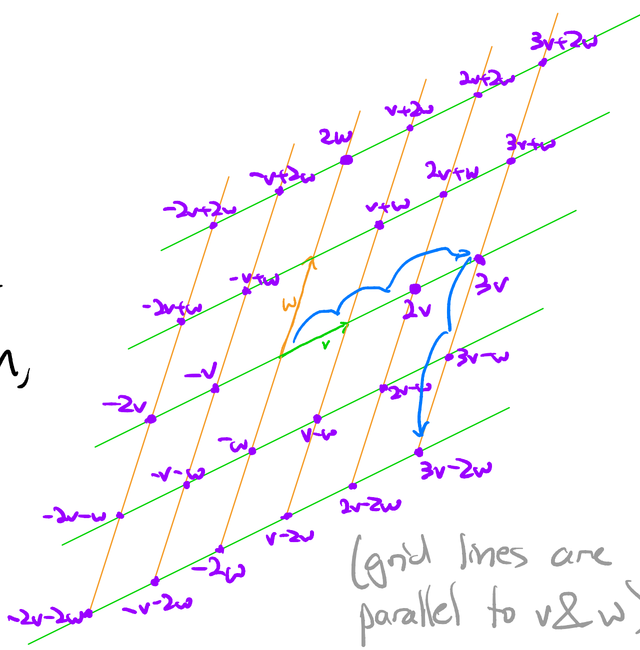
Linear Combinations

First scale, then add.

This is like giving directions:

"To get to $3v - 2w$, first go $3\times$ length of v in the v-direction, then go $2\times$ length of w in the $(-w)$-direction."

$v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ $w = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$



(grid lines are parallel to v & w)

[DEMO 1] [DEMO 2] [DEMO 3]

Spans

Recall: If $Ax=b$ is consistent, then the solution set has the form

$$x = \begin{pmatrix} \text{constant} \\ \text{vector} \end{pmatrix} + \begin{pmatrix} \text{all linear combinations of} \\ \text{some vectors from PVF} \end{pmatrix}$$

For which b is $Ax=b$ consistent (has a solution)?
Rewrite as a vector equation:

$$Ax=b \quad \equiv \quad x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = b$$

columns of A

This has a solution (there exists some choice of weights $x_1, \ldots, x_n$) exactly when b is a linear combination of $v_1, v_2, \ldots, v_n$.

Said differently, $Ax=b$ is consistent for

$$b \in \begin{pmatrix} \text{all linear combinations of} \\ \text{the columns of } A \end{pmatrix}$$

This is the column picture criterion for consistency, which we'll see many times.

Let's study & draw "all linear combinations of some vectors."

Def: The span of a list of vectors is the set of all linear combinations of those vectors:

$$\text{Span}\{v_1, v_2, \dots, v_n\} = \{x_1v_1 + x_2v_2 + \cdots + x_nv_n : x_1, \dots, x_n \in \mathbb{R}\}$$

"the set of" "all things of this form" "such that" "these conditions hold"

This is set builder notation

Translation of the previous page using spans:

If $Ax=b$ is consistent, then the solution set has the form

$$x = \begin{pmatrix} \text{constant} \\ \text{vector} \end{pmatrix} + \text{Span}\{\text{some vectors from PVF}\}$$

and $Ax=b$ is consistent if and only if

$$b \in \text{Span}\{\text{columns of } A\}.$$

Visualizing Spans

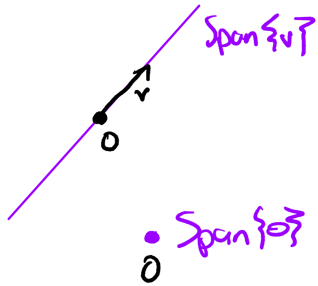
Span$\{v_1, v_2, \dots, v_n\}$ is the smallest "linear space" containing the vectors $v_1, v_2, \dots, v_n$ and 0.

One Vector:

$$\text{Span}\{v\} = \{x \cdot v : x \in \mathbb{R}\}$$

→ This is the line thru 0 and v unless

→ $v = 0$, in which case it's the point $\{0\}$.



[DEMO 1] [DEMO 2]

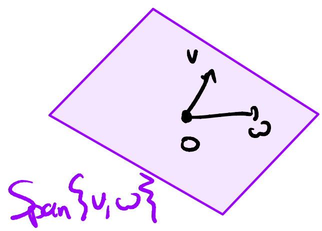
2 vectors:

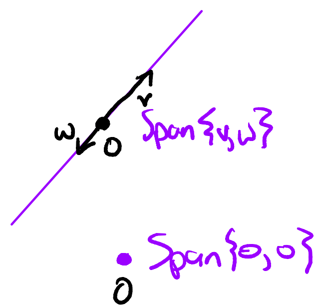
$$\text{Span}\{v, w\} = \{x_1 v + x_2 w : x_1, x_2 \in \mathbb{R}\}$$

→ This is the plane thru $0, v, w$ unless

→ $0, v, w$ lie on the same line, unless

→ $0 = v = w$, in which case it's the point $\{0\}$.

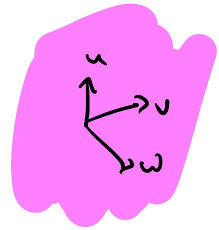




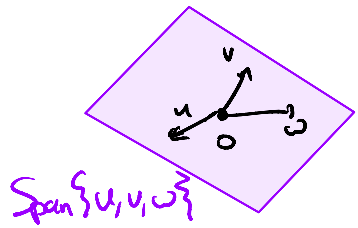
[DEMO 1] [DEMO 2]

3 vectors:

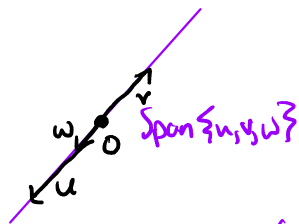
$$\text{Span}\{u, v, w\} = \{x_1 u + x_2 v + x_3 w : x_1, x_2, x_3 \in \mathbb{R}\}$$



→ This is the space containing $0, u, v, w$ unless



→ $0, u, v, w$ lie on the same plane, unless



→ $0, u, v, w$ lie on the same line, unless

→ $0 = u = v = w$, in which case it's the point $\{0\}$.

[DEMO]

What about the span of zero vectors?

$$\boxed{\text{Span}\{\} = \{0\}}$$

(by convention)

Potential Cause of Confusion:

Be careful to distinguish between these 2 sets:

$\{\}$ is the empty set. It has no vectors at all.
(eg: the solution set of an inconsistent system)

$\{0\}$ is a point. It contains (only) the zero vector.

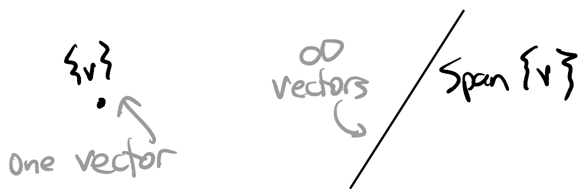
The (only) difference between them is:

$\{0\}$ contains 0 and $\{\}$ does not.

Another Potential Cause of Confusion:

$\{v_1, v_2, \ldots, v_n\}$ is a set with n vectors in it.

$\mathrm{Span}\{v_1, v_2, \ldots, v_n\}$ is a linear space: it has infinitely many vectors. (unless all $v_i = 0$...)



Taking a Span is a parametric description of a linear space thru the origin:

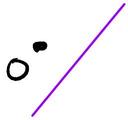
parameters

$$\mathrm{Span}\{v, w\} = \{x_1 v + x_2 w : x_1, x_2 \in \mathbb{R}\}$$

It is one way to describe an infinite set using finite data. (Then you can do computations!)

$$\boxed{\textbf{NB: } \text{Every span contains } 0!}$$

because $0 = 0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n$.

Non-Eg:  This line is a "linear space" but it is not a span because it does not contain 0.

Non-Eg: $\{\}$ does not contain $0 \Longrightarrow$ not a span.

Computational Question: Is $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ in $\text{Span}\left\{\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}\right\}$?

Translate: is $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ a linear combination of $\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}$?

Translate: Can we solve the equation

$$x_1\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} + x_2\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 6 & -1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix} \ ?$$

We know how to do this!

$$\left(\begin{array}{cc|c} 1 & -1 & 8 \\ 2 & -2 & 16 \\ 6 & -1 & 3 \end{array}\right) \xrightarrow{\text{rref}} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & -9 \\ 0 & 0 & 0 \end{array}\right) \rightsquigarrow \begin{array}{l} x_1 = -1 \\ x_2 = -9 \end{array}$$

so $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix} = -\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} - 9\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} \in \text{Span}\left\{\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}\right\}$ ✓

This is a specific instance of the:

Column Picture Criterion for Consistency

$Ax = b$ is consistent $\Longleftrightarrow$ $b \in \text{Span}\{\text{columns of } A\}$

(has at least one solution)

I can't over-emphasize how important it is for you to internalize this concept!

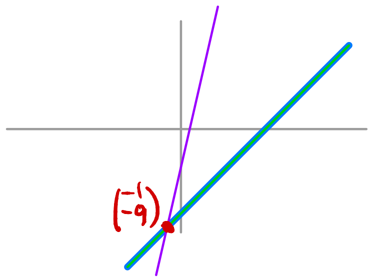
Eg: $\begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 6 & -1 \end{pmatrix} x = \begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ is consistent because

$$\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix} \in \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} \right\}$$

[DEMO]

row picture:

$$x_1 - x_2 = 8$$
$$2x_1 - 2x_2 = 16$$
$$6x_1 - x_2 = 3$$



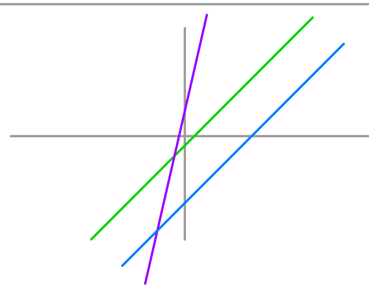
Eg: $\begin{pmatrix} 1 & -1 \\ 2 & -2 \\ 6 & -1 \end{pmatrix} x = \begin{pmatrix} 7 \\ 1 \\ -1 \end{pmatrix}$ is inconsistent because

$$\begin{pmatrix} 7 \\ 1 \\ -1 \end{pmatrix} \notin \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} \right\}$$

[DEMO]

row picture:

$$x_1 - x_2 = 7$$
$$2x_1 - 2x_2 = 1$$
$$6x_1 - x_2 = -1$$



Homogeneous Equations

"hoe-moe-GEE-nee-us"

Def: A system of equations $Ax=b$ is homogeneous if $b=0$.

Eg:
$$\begin{aligned} x_1+2x_2+2x_3+x_4&=0\\ 2x_1+4x_2+x_3-x_4&=0 \end{aligned}$$
is homogeneous

A homogeneous system is always consistent because $x=0$ is a solution: $A\cdot 0=0$

Def: The zero vector is called the trivial solution of the homogeneous system $Ax=0$.

It's "trivial" in the sense that it has nothing to do with A.

NB: Conversely, if $x=0$ is a solution of $Ax=b$ then $b=A\cdot 0=0$.

Eg: Let's find the PVF of the solutions of the system

$$\begin{cases} x_1+2x_2+2x_3+x_4=0\\ 2x_1+4x_2+x_3-x_4=0 \end{cases} \rightsquigarrow \left(\begin{array}{cccc|c} 1&2&2&1&0\\ 2&4&1&-1&0 \end{array}\right)$$

$$\xrightarrow{R_2 -= 2R_1} \left(\begin{array}{cccc|c} 1&2&2&1&0\\ 0&0&-3&-3&0 \end{array}\right)$$

$$\xrightarrow{R_2 \div= 3} \left(\begin{array}{cccc|c} 1&2&2&1&0\\ 0&0&1&1&0 \end{array}\right)$$

$$\xrightarrow{R_1 -= 2R_2} \left(\begin{array}{cccc|c} 1&2&0&-1&0\\ 0&0&1&1&0 \end{array}\right)$$

free

$$\rightsquigarrow \begin{aligned} x_1 &= 0 - 2x_2 + x_4 \\ x_2 &= \quad\ \ x_2 \\ x_3 &= 0 \qquad\quad - x_4 \\ x_4 &= \qquad\qquad\ x_4 \end{aligned} \quad \overset{\text{PVF}}{\rightsquigarrow} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Observations:

(1) The augmented column is always zero.

If your system is homogeneous, there's no real need to write it down!

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 0 \\ 2 & 4 & 1 & -1 & 0 \end{array}\right) \rightsquigarrow \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 1 & -1 \end{pmatrix}$$

(2) The particular solution is zero.

(3) The solution set is a span:

$$\begin{pmatrix} \text{solution} \\ \text{set} \end{pmatrix} = \text{Span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\} \quad \begin{pmatrix} \text{the weights} \\ \text{are the} \\ \text{free variables} \end{pmatrix}$$

These observations are true for any homogeneous system:

$$\boxed{\begin{pmatrix} \text{solution set} \\ \text{of } Ax = 0 \end{pmatrix} = \text{Span}\left\{ \begin{matrix} \text{vectors from} \\ \text{PVF} \end{matrix} \right\}}$$

Inhomogeneous Equations

Def: A system of equations $Ax=b$ is **inhomogeneous** if $b \neq 0$.

Eg:
$$\begin{aligned} x_1+2x_2+2x_3+x_4 &= 1 \\ 2x_1+4x_2+x_3-x_4 &= -1 \end{aligned}$$
is inhomogeneous

An inhomogeneous system may be inconsistent.

Def: If $Ax=b$ is inhomogeneous, its **associated homogeneous equation** is $Ax=0$.

Eg: Let's solve this inhomogeneous system, with its associated homogeneous equation alongside.

INHOMOGENEOUS

$$\begin{cases} x_1+2x_2+2x_3+x_4 = 1 \\ 2x_1+4x_2+x_3-x_4 = -1 \end{cases}$$

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 2 & 4 & 1 & -1 & -1 \end{array}\right) \xrightarrow{R_2 -= 2R_1} \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & -3 & -3 & -3 \end{array}\right)$$

$$\xrightarrow{R_2 \div= 3} \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right) \xrightarrow{R_1 -= 2R_2} \left(\begin{array}{cccc|c} 1 & 2 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right)$$

↖ free ↗

HOMOGENEOUS

$$\begin{cases} x_1+2x_2+2x_3+x_4 = 0 \\ 2x_1+4x_2+x_3-x_4 = 0 \end{cases}$$

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 0 \\ 2 & 4 & 1 & -1 & 0 \end{array}\right) \xrightarrow{R_2 -= 2R_1} \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & -3 & -3 & 0 \end{array}\right)$$

$$\xrightarrow{R_2 \div= 3} \left(\begin{array}{cccc|c} 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right) \xrightarrow{R_1 -= 2R_2} \left(\begin{array}{cccc|c} 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right)$$

↖ free ↗

$$\rightsquigarrow \begin{aligned} x_1 &= -1 - 2x_2 + x_4 \\ x_2 &= x_2 \\ x_3 &= 1 - x_4 \\ x_4 &= x_4 \end{aligned} \qquad\qquad \rightsquigarrow \begin{aligned} x_1 &= 0 - 2x_2 + x_4 \\ x_2 &= x_2 \\ x_3 &= 0 - x_4 \\ x_4 &= x_4 \end{aligned}$$

PVF $\rightsquigarrow$ $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ PVF $\rightsquigarrow$ $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$

solution set:

$$\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mathrm{Span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\} \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathrm{Span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

different!

same!

The only difference is the particular solution!

This means they're parallel planes!

In general:

$$\begin{pmatrix} \text{solution set} \\ \text{of } Ax = 0 \end{pmatrix} = (\text{zero}) + \mathrm{Span}\left\{ \begin{matrix} \text{vectors from} \\ \text{PVF} \end{matrix} \right\}$$

this is a span!

same vectors!

$$\begin{pmatrix} \text{solution set} \\ \text{of } Ax = b \end{pmatrix} = \begin{pmatrix} \text{nonzero} \\ \text{vector} \end{pmatrix} + \mathrm{Span}\left\{ \begin{matrix} \text{vectors from} \\ \text{PVF} \end{matrix} \right\}$$

$b \neq 0$

(if $Ax=b$ is consistent)

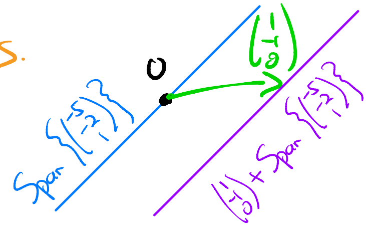
this is not a span! it is a translate of a span.

Eg: The solution set of...

... $\begin{pmatrix} 2 & 1 & 12 \\ 1 & 2 & 9 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is $\mathrm{Span}\left\{\begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}\right\}$

... $\begin{pmatrix} 2 & 1 & 12 \\ 1 & 2 & 9 \end{pmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \mathrm{Span}\left\{\begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}\right\}$

These are parallel lines.



[DEMO]

Upshot: the solution set of $Ax=b$ is a translate of the solution set of $Ax=0$. (or it is empty)

$$\boxed{\text{The solution set of } Ax=b \text{ looks exactly the same for different values of } b \in \mathrm{Span}\{\text{cols of } A\}.}$$

(if $b \notin \mathrm{Span}\{\text{cols of } A\}$ then $Ax=b$ is inconsistent.)

NB: In fact, to get from $\begin{pmatrix} \text{solns of} \\ Ax=0 \end{pmatrix}$ to $\begin{pmatrix} \text{solns of} \\ Ax=b \end{pmatrix}$, you can translate by any one solution of $Ax=b$.

$\rightarrow$ if $Ap=b$ and $Ax=0$ then

$$A(p+x) = Ap + Ax = b + 0 = b \quad \checkmark$$

So you don't really need to use the particular solution from PVF.

Reiterate: Expressing a solution set as a (translate of a) span means giving a parametric description:

$$\begin{pmatrix}\text{solution}\\ \text{set}\end{pmatrix} = \begin{pmatrix}-1\\0\\1\\0\end{pmatrix} + \text{Span}\left\{\begin{pmatrix}-2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\1\end{pmatrix}\right\}$$

means

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

parameters = free variables

So think:

Spans $\equiv$ Parametric Description

Summary: Row & Column Picture

We have discussed two conceptual properties of systems of linear equations, namely, what happens when you allow the x vector & the b vector to vary.

Column Picture: vary the b vector in $\mathbb{R}^m$

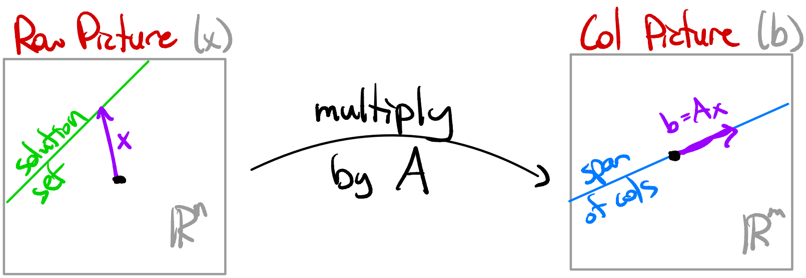
$$\boxed{Ax=b \text{ is consistent} \iff b \in \text{Span}\{\text{cols of } A\}}$$

Row Picture: the x vectors in $\mathbb{R}^n$

$$\boxed{\begin{pmatrix}\text{all solutions} \\ \text{of } Ax=b\end{pmatrix} = \begin{pmatrix}\text{some solution} \\ \text{of } Ax=b\end{pmatrix} + \begin{pmatrix}\text{all solutions} \\ \text{of } Ax=0\end{pmatrix}}$$

or no solutions if $b \notin \text{Span}\{\text{Cols of } A\}$

Picture of both at the same time:



[SUPER IMPORTANT DEMO]

The solution set lives in the... row picture!
The b-vectors live in the... column picture!
The columns all live in the... column picture!
That's how you keep them straight!

NB: We're thinking of a matrix A as a function:

$x \in \mathbb{R}^n$ is the input (row picture)

$b = Ax \in \mathbb{R}^m$ is the output (column picture)

Visualizing a matrix means geometrically understanding the relationship between x and Ax.