

Lecture 17: Vector Spaces and Subspaces

Generalizing the notion of a “vector” (§3.5)

Differential Equations & Linear Algebra

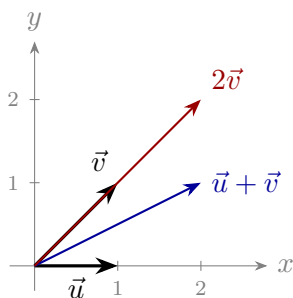
1 Beyond arrows

We’ve been using the language of linear algebra and matrices to talk about n -dimensional arrows (“vectors”) and linear operators on them.

In 2360 we do this because the same formalism can be applied to *functions* (“vectors”) and to solutions of linear differential equations.

Today: generalize our notion of “vectors” beyond arrows.

1.1 Arrows as “vectors”



If

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are vectors, we can use addition and scalar multiplication to construct others, e.g.

$$\vec{u} + \vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad 2\vec{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

1.2 Polynomials as “vectors”

Let $\mathbb{P}_2$ be the set of polynomials of degree ≤ 2 (lines and parabolas). Let

$$u(t) = t^2, \quad v(t) = t^2 + t \quad \in \mathbb{P}_2.$$

Then

$$u + v = 2t^2 + t \quad \text{is also in } \mathbb{P}_2, \quad \text{and} \quad 2v = 2t^2 + 2t \quad \text{is in } \mathbb{P}_2.$$

If we wanted to, we could use shorthand and write each polynomial as a list of its coefficients:

$$u = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{array}{l} \leftarrow t^2 \text{ coeff.} \\ \leftarrow t \text{ coeff.} \\ \leftarrow \text{const. part} \end{array} \quad v = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{array}{l} \leftarrow t^2 \text{ coeff.} \\ \leftarrow t \text{ coeff.} \\ \leftarrow \text{const. part} \end{array}$$

Then notice that

$$u + v = 2t^2 + t = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad 2v = 2t^2 + 2t = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix},$$

exactly as if they were arrows in $\mathbb{R}^3$.

Remark (Fun fact, and probably the single most important take-away message from this section). Recall we said that any linear operator can be represented as a matrix (see section 1.3 of my first notes from chapter 3). Well, $\frac{d}{dt}$ is linear, and it takes in and spits out polynomials, so it must be possible to represent it as some matrix $\mathbf{D}$ such that the matrix multiplication $\mathbf{D}\vec{v}$ is exactly equivalent to $\frac{d}{dt}v(t)$ for any 2nd-order polynomial $v(t)$. Notice that

$$\mathbf{D}\vec{v} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \longleftrightarrow 2t + 1 = \frac{d}{dt}(t^2 + t).$$

or more generally, for arbitrary $w(t) = at^2 + bt + c \leftrightarrow \vec{w} = [a \ b \ c]^T$, we can see that

$$\mathbf{D}\vec{w} = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 2a \\ b \end{bmatrix} \longleftrightarrow 0t^2 + 2at + b = \frac{d}{dt}(at^2 + bt + c).$$

This is one of the biggest reasons for generalizing from thinking about vectors as just arrows to vectors sometimes being functions!! This “fun fact” alone underlies a huge amount of what linear algebra and diff eq actually looks like in the “real world” for STEM professionals. If “vectors” can only be arrows, then the language of matrix operations/multiplication is only useful for linear operators that act on arrows and return arrows. But now we can use matrix methods as a language for modeling more complex things, including derivatives and many differential equations (including *partial* differential equations, i.e., diff eqs where you have derivatives of multiple different things like x, y, z, and t, including the heat equation, Navier-Stokes equation for fluids, Schroedinger’s Equation for quantum mechanics, and more). My favorite example: once you get the hang of the fundamentals of linear algebra *and* this notion of functions as vectors and therefore derivatives and differential equations are just fancy matrix operations, then a whole bunch of quantum mechanics is WAY simpler than you’d think.

2 Vector spaces

If we want to generalize, and let things other than arrows be “vectors”, what are the ground rules?

Definition (vector space)

A **vector**, generally, is any element of some collection $\mathbb{V}$, called a **vector space**, that comes equipped with notions of “vector addition” and “scalar multiplication” satisfying the following, for any $\vec{u}, \vec{v}, \vec{w} \in \mathbb{V}$ and $c, d \in \mathbb{R}$ (or $\mathbb{C}$):

$$\begin{array}{ll} \text{Closure} & \vec{u} + \vec{v} \in \mathbb{V} \\ & c\vec{u} \in \mathbb{V} \end{array}$$

$$\text{Additive identity} \quad \text{there is a “}\vec{0}\text{”} \in \mathbb{V} \text{ such that } \vec{0} + \vec{u} = \vec{u} \text{ for all } \vec{u} \in \mathbb{V}$$

$$\text{Additive inverse} \quad \text{there is a “}\vec{-u}\text{” such that } \vec{u} + (\vec{-u}) = \vec{0}$$

$$\text{Associativity} \quad (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}) \quad \text{and} \quad c(d\vec{u}) = (cd)\vec{u}$$

$$\text{Commutativity} \quad \vec{u} + \vec{v} = \vec{v} + \vec{u}$$

$$\text{Multiplicative identity} \quad 1 \cdot \vec{u} = \vec{u}$$

$$\text{Distributivity} \quad c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v} \quad \text{and} \quad (c + d)\vec{u} = c\vec{u} + d\vec{u}$$

3 Examples (and non-examples)

3.1 $\mathbb{R}^2$ with the usual operations

Example 3.1. Take $\mathbb{V} = \mathbb{R}^2$, where we take “+” to mean *component-wise addition* and “ $c\vec{u}$ ” to mean *component-wise multiplication*:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}, \quad c \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}.$$

- **Closure:** sums and scalar multiples of pairs of real numbers are pairs of real numbers, so both stay in $\mathbb{R}^2$. ✓
- **Additive identity:** $\vec{0} = [0, 0]^T$, since by component-wise addition

$$\vec{u} + \vec{0} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u_1 + 0 \\ u_2 + 0 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \vec{u}. \quad \checkmark$$

- **Additive inverse:** $-\vec{u} = [-u_1, -u_2]^T$, since

$$\vec{u} + (-\vec{u}) = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -u_1 \\ -u_2 \end{bmatrix} = \begin{bmatrix} u_1 + (-u_1) \\ u_2 + (-u_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}. \quad \checkmark$$

- **Associativity, commutativity, multiplicative identity, distributivity:** each holds component by component, because it holds for ordinary arithmetic of real numbers. ✓

So $\mathbb{R}^2$ is a vector space.

3.2 $\mathbb{R}_+^2$: a vector space with *weird* operations

The operations don't have to be the familiar ones. As long as the criteria hold, *any* notion of “addition” and “scalar multiplication” will do.

Example 3.2. Let $\mathbb{R}_+^2$ be the set of vectors in $\mathbb{R}^2$ with *strictly positive* entries, and let $c \in \mathbb{R}$. Define “addition” $\oplus$ as element-wise *multiplication*, and “scalar multiplication” $\odot$ as element-wise *exponentiation*:

$$\vec{u} \oplus \vec{v} \equiv \begin{bmatrix} u_1 v_1 \\ u_2 v_2 \end{bmatrix}, \quad c \odot \vec{u} \equiv \begin{bmatrix} u_1^c \\ u_2^c \end{bmatrix}.$$

For instance, $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \oplus \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 3 \end{bmatrix}$ and $2 \odot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$.

Checking each criterion, for any $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}_+^2$ and $c, d \in \mathbb{R}$ (below, $[a_i]$ is shorthand for the vector whose entries are a_1 and a_2):

Closure $u_1 v_1 > 0$, $u_2 v_2 > 0$, and $u_i^c > 0$ for any real c , so $\vec{u} \oplus \vec{v} \in \mathbb{R}_+^2$ and $c \odot \vec{u} \in \mathbb{R}_+^2$. ✓

Additive identity $\vec{0} = [1, 1]^T$, since $\vec{0} \oplus \vec{u} = [1 \cdot u_1, 1 \cdot u_2]^T = \vec{u}$. ✓

Additive inverse $-\vec{u} = [1/u_1, 1/u_2]^T \in \mathbb{R}_+^2$, since $\vec{u} \oplus (-\vec{u}) = [1, 1]^T = \vec{0}$. ✓

Associativity $(u_i v_i) w_i = u_i (v_i w_i)$, and $c \odot (d \odot \vec{u}) = [(u_i^d)^c] = [u_i^{cd}] = (cd) \odot \vec{u}$. ✓

Commutativity $u_i v_i = v_i u_i$. ✓

Multiplicative identity $1 \odot \vec{u} = [u_i^1] = \vec{u}$. ✓

Distributivity $c \odot (\vec{u} \oplus \vec{v}) = [(u_i v_i)^c] = [u_i^c v_i^c] = (c \odot \vec{u}) \oplus (c \odot \vec{v})$,
 $(c + d) \odot \vec{u} = [u_i^{c+d}] = [u_i^c u_i^d] = (c \odot \vec{u}) \oplus (d \odot \vec{u})$. ✓

So $\mathbb{R}_+^2$, with these operations, is a vector space — even though its “zero vector” is $[1, 1]^T$!

3.3 A non-example: $\mathbb{R}^2$ with “addition” as subtraction

Example 3.3. Take $\mathbb{V} = \mathbb{R}^2$ with ordinary scalar multiplication, but define “addition” $\boxplus$ to be component-wise *subtraction*:

$$\vec{u} \boxplus \vec{v} \equiv \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \end{bmatrix}.$$

Closure holds: $u_1 - v_1$ and $u_2 - v_2$ are real numbers, so $\vec{u} \boxplus \vec{v} \in \mathbb{R}^2$; and $c \vec{u} \in \mathbb{R}^2$ as before. ✓

Commutativity fails. Swapping the order flips the sign of every entry:

$$\vec{v} \boxplus \vec{u} = \begin{bmatrix} v_1 - u_1 \\ v_2 - u_2 \end{bmatrix} = - \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \end{bmatrix} = -(\vec{u} \boxplus \vec{v}),$$

so the two agree only when $\vec{u} = \vec{v}$. For instance, with $\vec{u} = [3, 1]^T$ and $\vec{v} = [1, 2]^T$,

$$\vec{u} \boxplus \vec{v} = \begin{bmatrix} 3-1 \\ 1-2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \neq \vec{v} \boxplus \vec{u} = \begin{bmatrix} 1-3 \\ 2-1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}. \quad \times$$

So $\mathbb{R}^2$ with this “addition” is **not** a vector space, even though it is closed. Failing a single criterion is enough.

3.4 $\mathbb{M}_{22}$: the set of all 2×2 matrices

Example 3.4. Let $\mathbb{M}_{22}$ be the set of all 2×2 matrices with real entries, with the usual entry-by-entry addition and scalar multiplication. Take any

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \in \mathbb{M}_{22}, \quad k \in \mathbb{R}.$$

Closure:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} \in \mathbb{M}_{22}, \quad k\mathbf{A} = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix} \in \mathbb{M}_{22}. \quad \checkmark$$

Both results are again 2×2 matrices with real entries.

One can show in the same way that $\mathbb{M}_{22}$ satisfies all the other criteria too (its zero vector is the zero matrix $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, and $-\mathbf{A}$ is the matrix of negated entries). So $\mathbb{M}_{22}$ is a vector space, and each 2×2 matrix is a “vector” in it.

Note: matrix multiplication $\mathbf{AB} = \mathbf{C}$ is a perfectly well-defined operation on 2×2 matrices, but it is *irrelevant* to the notion of a vector space. As in: yes, 2×2 matrices can be multiplied together. That’s great. But for the purposes of asking the question, “does $\mathbb{M}_{22}$ qualify as a vector space?” we don’t care, all we care about is the fact that we can add two matrices together and multiply a matrix by a scalar.

The same goes for polynomials like in section 1.2 above. Yes, polynomials can be multiplied together, but for purposes of discussing polynomials as vectors in a vector space, that’s irrelevant. (*Exception: down the road you might learn about generalizing dot products AKA scalar products from arrows to functions and things like “Hilbert spaces”, and then the generalized scalar product on functions typically includes something along the lines of multiplying the functions together and then integrating the result.*)

3.5 A non-example: matrices with at least one zero entry

Example 3.5. Let $\mathbb{W}$ be the set of 2×2 matrices with *at least one* zero entry. Then

$$\vec{u} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{W}, \quad \vec{v} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{W},$$

but

$$\vec{u} + \vec{v} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \notin \mathbb{W},$$

since it has no zero entries. $\mathbb{W}$ is **not closed under addition**, so it is **not** a vector space. ✗

3.6 Matrices whose first entry is zero

Example 3.6. Now let $\mathbb{W}$ be the set of 2×2 matrices whose *first* (top-left) entry is zero:

$$\mathbb{W} = \left\{ \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} : b, c, d \in \mathbb{R} \right\}.$$

For any two such matrices and any $k \in \mathbb{R}$,

$$\begin{bmatrix} 0 & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & f \\ g & h \end{bmatrix} = \begin{bmatrix} 0 & b+f \\ c+g & d+h \end{bmatrix} \in \mathbb{W}, \quad k \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & kb \\ kc & kd \end{bmatrix} \in \mathbb{W}.$$

The top-left entry stays 0 no matter what, so $\mathbb{W}$ is **closed** under both operations. ✓

We will return to this example below: it is a vector space that lives *inside* $\mathbb{M}_{22}$.

3.7 Other important vector spaces

- $\mathbb{R}^n$: column vectors with n real entries.
- $\mathbb{C}^n$: column vectors with n complex entries (with scalars from $\mathbb{R}$ or from $\mathbb{C}$).
- $\mathbb{M}_{mn}$: all $m \times n$ matrices.
- $\mathbb{P}$: all polynomials; $\mathbb{P}_n$: polynomials of degree $\leq n$.
- $C(I)$: all continuous functions on an interval I ; $C^n(I)$: all functions on I with n continuous derivatives.

(Don't confuse $\mathbb{C}^n$, lists of complex numbers, with $C^n(I)$, functions with n continuous derivatives.)

Solutions to linear homogeneous equations! For any (system of) linear homogeneous differential equations, the set of solutions is a vector space.

Briefly: if $L[y_1] = 0$ and $L[y_2] = 0$ for a linear operator L , then for any c ,

$$L[y_1 + c y_2] = L[y_1] + c L[y_2] = 0 + 0 = 0,$$

so the solution set is closed. This is *why* we are emphasizing linear algebra in this course: every tool we develop for vector spaces applies directly to solutions of linear DEs.

4 Subspaces

Definition (subspace)

A **subspace** $\mathbb{W}$ of a vector space $\mathbb{V}$ is a (nonempty) subset $\mathbb{W} \subseteq \mathbb{V}$ that *also* satisfies the criteria for a vector space, using the same addition and scalar multiplication as $\mathbb{V}$.

Most importantly, $\mathbb{W}$ must remain **closed**: for any $\vec{u}, \vec{v} \in \mathbb{W}$ and $c \in \mathbb{R}$,

$$\vec{u} + c\vec{v} \in \mathbb{W}.$$

Combining vectors in $\mathbb{W}$ must never land in a part of $\mathbb{V}$ that is not part of $\mathbb{W}$.

The other criteria (associativity, commutativity, etc.) are inherited automatically from $\mathbb{V}$, so closure is the one that can fail.

Trivial subspaces. Every vector space $\mathbb{V}$ has two:

$$\{\vec{0}\} \quad (\text{the zero vector alone}) \quad \text{and} \quad \mathbb{V} \subseteq \mathbb{V} \quad (\text{the whole space}).$$

Example 4.1. The set of 2×2 matrices whose first entry is zero (Example 3.6) is a subset of $\mathbb{M}_{22}$, and we showed it is closed under addition and scalar multiplication. So it is a **subspace** of $\mathbb{M}_{22}$. ✓

By contrast, the set of 2×2 matrices with *at least one* zero entry (Example 3.5) is *not* a subspace of $\mathbb{M}_{22}$. ✗

Remark. You can verify on your own that the set of all two-dimensional vectors whose first and second entry are equal is a subset of $\mathbb{R}^2$ and is closed under addition and scalar multiplication, and so that is a valid subspace of $\mathbb{R}^2$.

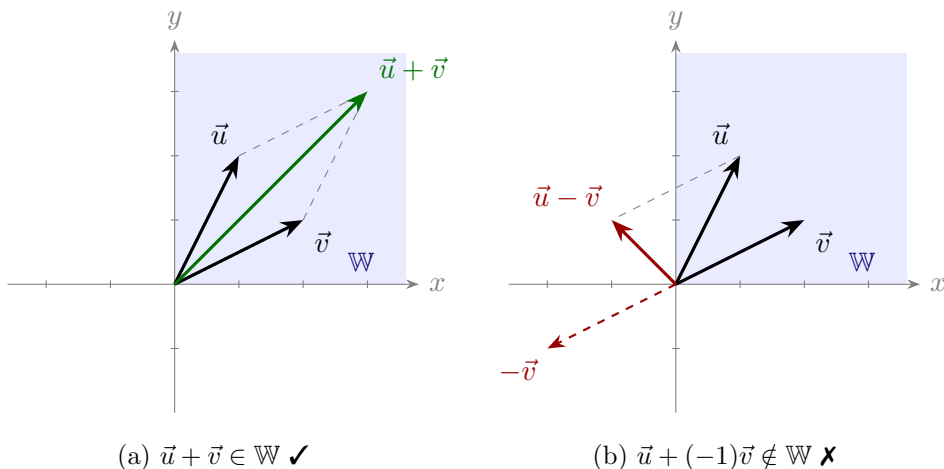
4.1 A non-example: the first quadrant of $\mathbb{R}^2$

Example 4.2. Let $\mathbb{W} = \{[x, y]^T : x \geq 0, y \geq 0\}$, the first quadrant of $\mathbb{R}^2$ (including its edges). Is it a subspace of $\mathbb{R}^2$?

Adding two vectors in $\mathbb{W}$ always stays in $\mathbb{W}$ (panel a). But scaling by a *negative* number does not: with $\vec{u} = [1, 2]^T$, $\vec{v} = [2, 1]^T$ and $c = -1$,

$$\vec{u} + c\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \notin \mathbb{W}$$

(panel b). Even more simply, $(-1)\vec{v} = [-2, -1]^T \notin \mathbb{W}$.



The first quadrant is closed under addition but **not** under scalar multiplication, so it violates closure and is **not a subspace** of $\mathbb{R}^2$.

4.2 Another non-example: the first and third quadrants of $\mathbb{R}^2$

Example 4.3. Now let $\mathbb{W}$ be the first *and* third quadrants of $\mathbb{R}^2$ (including their edges):

$$\mathbb{W} = \{[x, y]^\top : x \geq 0, y \geq 0\} \cup \{[x, y]^\top : x \leq 0, y \leq 0\} = \{[x, y]^\top : xy \geq 0\}.$$

This time, **scalar multiplication** always stays in $\mathbb{W}$: if $xy \geq 0$, then for any $c \in \mathbb{R}$,

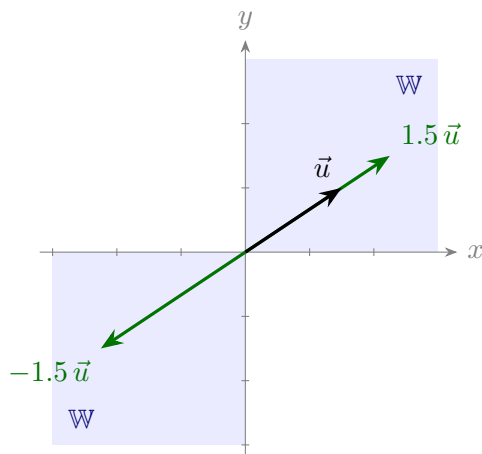
$$(cx)(cy) = c^2 xy \geq 0,$$

so $c\vec{u} \in \mathbb{W}$. Geometrically, scaling by $c \geq 0$ keeps a vector in its own quadrant, and scaling by $c < 0$ flips it into the *other* quadrant of $\mathbb{W}$ (panel a).

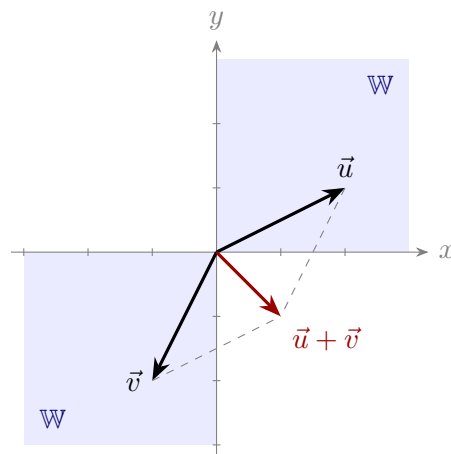
But **addition** does not: with $\vec{u} = [2, 1]^\top$ (first quadrant) and $\vec{v} = [-1, -2]^\top$ (third quadrant),

$$\vec{u} + \vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \notin \mathbb{W},$$

which lands in the fourth quadrant (panel b).



(a) $c\vec{u} \in \mathbb{W}$ for every c ✓



(b) $\vec{u} + \vec{v} \notin \mathbb{W}$ ✗

The first and third quadrants together are closed under scalar multiplication but **not** under addition, so this set also violates closure and is **not a subspace** of $\mathbb{R}^2$. A subspace needs *both* kinds of closure.