Week 1

Dot-free notation:
Sequence (of vectors):

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = (\mathbf{v}_j)_{j=1}^n$$

${}_{j=1}^{n}$: "all j such that $1 \leq j \leq n$, in increasing order"
$n = 2 : (\mathbf{v}_1, \mathbf{v}_2)$
$n = 1 : (\mathbf{v}_1)$
$n = 0 : ()$ (empty sequence)

Linear combination:

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j$$

$n = 2 : \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2$
$n = 1 : \lambda_1 \mathbf{v}_1$
$n = 0$: $\mathbf{0}$ (without moving, we're stuck at $\mathbf{0}$)

Set (of vectors):

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \{\mathbf{v}_j : j \in [n]\}, \quad [n] = \{1, 2, \ldots, n\}$$

Vectors:

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} = [v_i]_{i=1}^m \qquad \Bigg| \qquad [0]_{i=1}^6 = \mathbf{0} \in \mathbb{R}^6, \qquad [i^2]_{i=1}^5 = \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \\ 25 \end{bmatrix}, \qquad [v_i]_{i=1}^0 = () \in \mathbb{R}^0$$

Scalar products, lengths and angles (Section 1.2)

Scalar product: multiply two vectors!

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1 \cdot 3 + 2 \cdot 4 = 11.$$

Definition 1.9: Let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \in \mathbb{R}^m.$$

The scalar product of $\mathbf{v}$ and $\mathbf{w}$ is the number

$$\mathbf{v} \cdot \mathbf{w} := v_1 w_1 + v_2 w_2 + \cdots + v_m w_m = \sum_{i=1}^{m} v_i w_i.$$

Observation 1.10: Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ be vectors and $\lambda \in \mathbb{R}$ a scalar. Then

(i) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

(ii) $(\lambda \mathbf{v}) \cdot \mathbf{w} = \lambda(\mathbf{v} \cdot \mathbf{w}) = \mathbf{v} \cdot (\lambda \mathbf{w})$

(iii) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ and $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(iv) $\mathbf{v} \cdot \mathbf{v} \geq 0$, with equality exactly if $\mathbf{v} = \mathbf{0}$

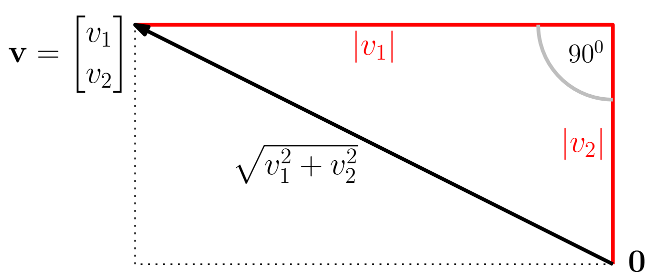
Euclidean norm: defines length of a vector

Definition 1.11: Let $\mathbf{v} \in \mathbb{R}^m$. The Euclidean norm of $\mathbf{v}$ is the number

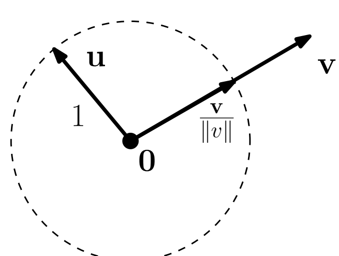
$$\|\mathbf{v}\| := \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

$$\left\| \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} \right\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_m^2} = \sqrt{\sum_{i=1}^{m} v_i^2} \qquad \left\| \begin{bmatrix} -4 \\ 2 \end{bmatrix} \right\| = \sqrt{(-4)^2 + 2^2} = \sqrt{20}$$

In $\mathbb{R}^2$: arrow length (Pythagoras!)

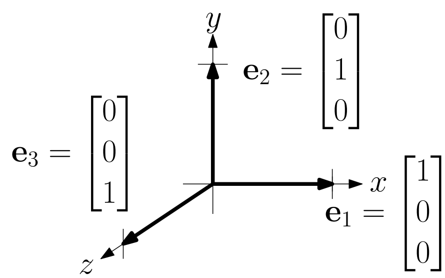


Unit vector: $\|\mathbf{u}\| = 1$.



For $\mathbf{v} \neq \mathbf{0}$, $\frac{\mathbf{v}}{\|\mathbf{v}\|} := \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ is a unit vector.
Standard unit vectors:

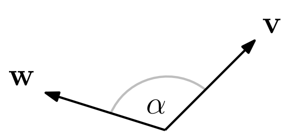
$$\mathbb{R}^3 : \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \Bigg| \qquad \mathbb{R}^m : \mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{coordinate } i$$



Cauchy-Schwarz inequality (Proof and application in lecture notes):
Lemma 1.12: For any two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$,

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\|\|\mathbf{w}\|.$$

Equality holds exactly if one vector is a scalar multiple of the other.
Angle between two vectors:

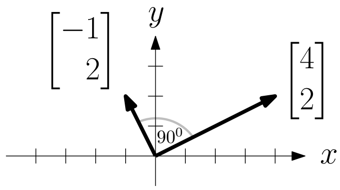


Definition 1.14: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ be two nonzero vectors. The angle between them is the unique α between 0 and π (180 degrees) such that

$$\cos(\alpha) = \underbrace{\frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}}_{\substack{\uparrow \\ \text{between } -1 \text{ and } 1 \text{ by Cauchy-Schwarz}}}, \quad \text{or} \quad \alpha = \arccos\left(\frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}\right).$$

In $\mathbb{R}^2$: the usual angle

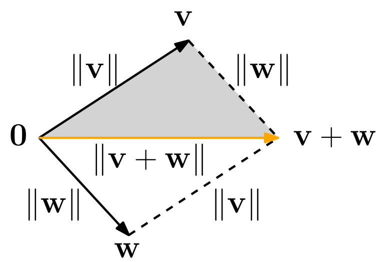
Perpendicular vectors:
Definition 1.15: Vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are *perpendicular* (or *orthogonal*) if $\mathbf{v} \cdot \mathbf{w} = 0$ (same as $\cos(\alpha) = 0$, or 90 degrees).



$$\begin{bmatrix}4\\2\end{bmatrix} \cdot \begin{bmatrix}-1\\2\end{bmatrix} = -4 \cdot 1 + 2 \cdot 2 = 0$$

Triangle inequality (proof from Cauchy-Schwarz):
Lemma 1.16: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$. Then

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$



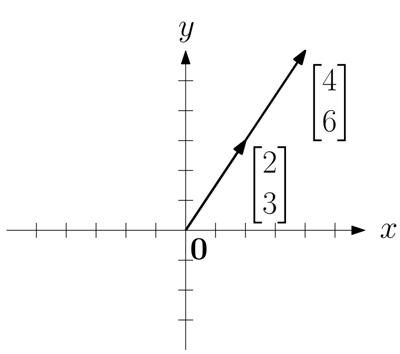
In $\mathbb{R}^2$: From $\mathbf{0}$ directly to $\mathbf{v} + \mathbf{w}$ is shorter than via $\mathbf{v}$ or $\mathbf{w}$.

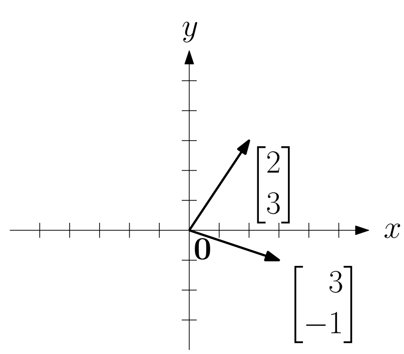
Linear independence (Section 1.3)

Linear (in)dependence:
Definition 1.18: Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are *linearly dependent* if at least one of them is a linear combination of the others, i.e. there exists an index $k \in [n]$ and scalars λ_j such that

$$\mathbf{v}_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} \lambda_j \mathbf{v}_j.$$

Otherwise, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are *linearly independent*.
“$j < k$”: an additional condition on j (all j except k).
$\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}4\\6\end{bmatrix}$ are linearly dependent: $\begin{bmatrix}4\\6\end{bmatrix} = 2\begin{bmatrix}2\\3\end{bmatrix}$.
$\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}3\\-1\end{bmatrix}$ are linearly independent.





collinear linearly independent

Three vectors in $\mathbb{R}^2$ are linearly dependent: either two are collinear, or each is a linear combination of the other two (Challenge 1.6).

linearly independent	linearly dependent
$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}$	
	$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \end{bmatrix}$
	$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^2$
$\mathbf{v} \neq \mathbf{0}$	
	$\mathbf{v} = \mathbf{0}$
	$\dots, \mathbf{0}, \dots$
	$\dots, \mathbf{v}, \dots, \mathbf{v}, \dots$
empty sequence	

Alternative definitions:
Lemma 1.19: Let $\mathbf{v}_1, \mathbf{v}_2 \dots, \mathbf{v}_n \in \mathbb{R}^m$. The following statements are *equivalent* (all true, or all false).

(i) At least one of the vectors is a linear combination of the other ones (linearly dependent by Definition 1.18).

(ii) There are scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ besides $0, 0, \dots, 0$ such that $\sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \mathbf{0}$. Math jargon: $\mathbf{0}$ is a *nontrivial linear combination* of the vectors.

(iii) At least one of the vectors is a linear combination of the previous ones.

Proof idea:
(i) *implies* (ii): *if* (i) is true, *then* also (ii) is true. (i)⇒(ii)
(ii) implies (iii). (ii)⇒(iii)
(iii) implies (i). (iii)⇒(i)
Each statement implies the other ones! (i)⇔(ii)⇔(iii)
Math prose for (i)⇔(ii): (i) *if and only if* (ii)

Proof.
(i) ⇒(ii): Let

$$\mathbf{v}_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} \lambda_j \mathbf{v}_j.$$

Define $\lambda_k = -1$. We get (ii):

$$\mathbf{0} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j.$$

(ii)⇒(iii): Let k be the largest index such that $\lambda_k \neq 0$. Then

$$\mathbf{0} = \sum_{j=1}^{k} \lambda_j \mathbf{v}_j$$

and we get (iii):

$$\mathbf{v}_k = \sum_{j=1}^{k-1} \left(-\frac{\lambda_j}{\lambda_k}\right) \mathbf{v}_j.$$

(iii)⇒(i): a linear combination of the previous ones is also a linear combination of the other ones. □

For linear independence, simply take the opposite statements.

Corollary 1.20: Let $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n \in \mathbb{R}^m$. The following statements are equivalent (all true, or all false).

(i) None of the vectors is a linear combination of the other ones (linearly independent by Definition 1.18.)

(ii) There are no scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ besides $0, 0, \ldots, 0$ such that $\sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \mathbf{0}$. Math jargon: $\mathbf{0}$ can only be written as a *trivial linear combination* of the vectors.

(iii) None of the vectors is a linear combination of the previous ones.

Uniqueness of linear combination:

Lemma 1.21: Let $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n \in \mathbb{R}^m$ be linearly independent, and let $\mathbf{w} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j = \sum_{j=1}^{n} \mu_j \mathbf{v}_j$ be two ways of writing $\mathbf{w}$ as a linear combination. Then $\lambda_j = \mu_j$ for all $j \in [n]$.

Proof. Subtraction:

$$\mathbf{0} = \sum_{j=1}^{n} (\lambda_j - \mu_j) \mathbf{v}_j.$$

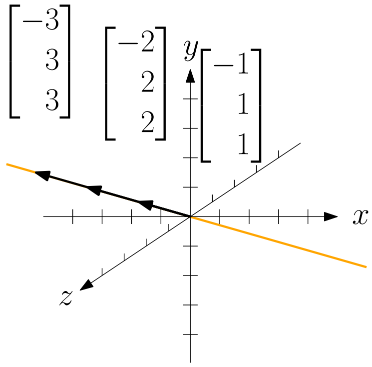
Since $\mathbf{0}$ can only be written as a trivial linear combination, we get $\lambda_j - \mu_j = 0$ for all j. □

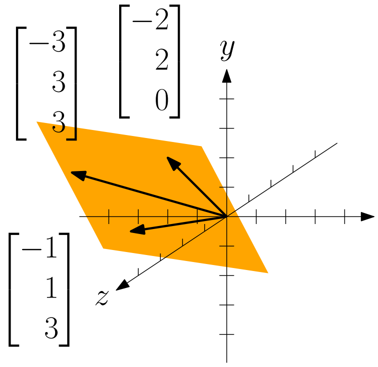
Span of vectors: set of all linear combinations

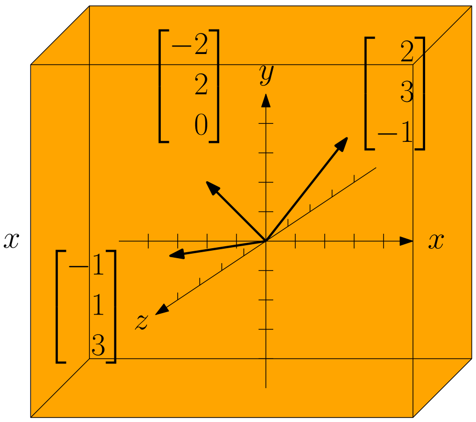
Definition 1.22: Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^m$. Their *span* is

$$\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) := \left\{ \sum_{j=1}^{n} \lambda_j \mathbf{v}_j : \lambda_j \in \mathbb{R} \text{ for all } j \in [n] \right\}.$$

Span of three vectors in $\mathbb{R}^3$:







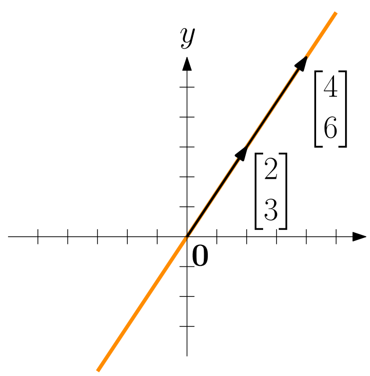
a line | a plane | the whole space

. . . or a point (if all vectors are $\mathbf{0}$)

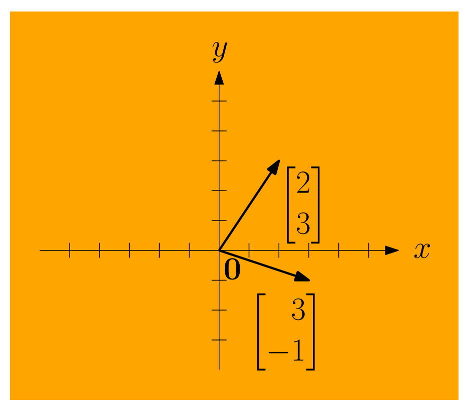
Always: $\mathbf{0} \in \mathbf{Span}(\ldots)$

Fact 1.5:

$$\mathbf{Span}\left(\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}3\\-1\end{bmatrix}\right) = \mathbb{R}^2.$$



collinear



linearly independent

Lemma 1.23: Let $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n \in \mathbb{R}^m$, and let $\mathbf{v} \in \mathbb{R}^m$ be a linear combination of $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n$. Then

$$\underbrace{\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)}_{S} = \underbrace{\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v})}_{T}.$$

Proof idea:

Each element of S is contained in T (S is *subset* of T). $S \subseteq T$

T is subset of S. $T \subseteq S$

The two sets are equal! $S = T$

Proof. $S \subseteq T$: Each $\mathbf{w} \in S$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ and therefore of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}$ (add scalar multiple $0\mathbf{v}$). So $\mathbf{w} \in T$.

$T \subseteq S$: each $\mathbf{w} \in T$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}$,

$$\mathbf{w} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j + \lambda \mathbf{v}.$$

We know: $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n$,

$$\mathbf{v} = \sum_{j=1}^{n} \mu_j \mathbf{v}_j.$$

Together:

$$\mathbf{w} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j + \lambda \mathbf{v} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j + \lambda \left(\sum_{j=1}^{n} \mu_j \mathbf{v}_j \right) = \sum_{j=1}^{n} (\lambda_j + \lambda \mu_j) \mathbf{v}_j.$$

So $\mathbf{w}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, $\mathbf{w} \in S$. □