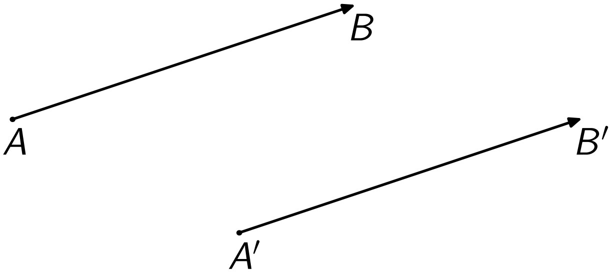
MATH 304

Linear Algebra

Lecture 17:
Euclidean structure in $\mathbb{R}^n$ (continued).
Orthogonal complement.
Orthogonal projection.

Vectors: geometric approach



- A vector is represented by a directed segment.
- Directed segment is drawn as an arrow.
- Different arrows represent the same vector if

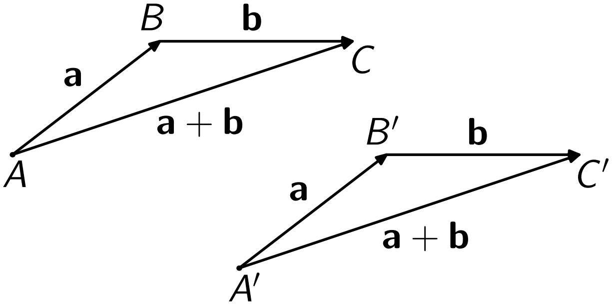
they are of the same length and direction.

Notation: $\overrightarrow{AB}$ ($= \overrightarrow{A'B'}$).

Linear structure: vector addition

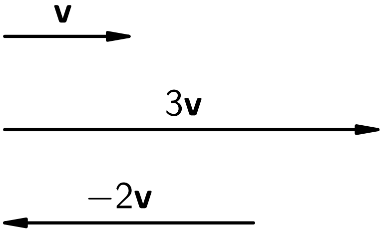
Given vectors $\mathbf{a}$ and $\mathbf{b}$, their *sum* $\mathbf{a}+\mathbf{b}$ is defined by the rule $\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}$.

That is, choose points A, B, C so that $\overrightarrow{AB}=\mathbf{a}$ and $\overrightarrow{BC}=\mathbf{b}$. Then $\mathbf{a}+\mathbf{b}=\overrightarrow{AC}$.



Linear structure: scalar multiplication

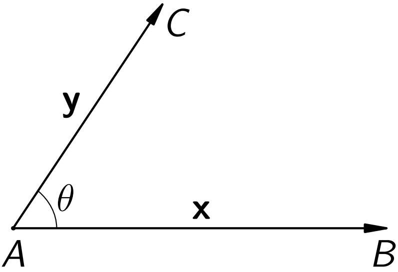
Let $\mathbf{v}$ be a vector and $r \in \mathbb{R}$. By definition, $r\mathbf{v}$ is a vector whose magnitude is $|r|$ times the magnitude of $\mathbf{v}$. The direction of $r\mathbf{v}$ coincides with that of $\mathbf{v}$ if $r > 0$. If $r < 0$ then the directions of $r\mathbf{v}$ and $\mathbf{v}$ are opposite.



Beyond linearity: Euclidean structure

Euclidean structure includes:

- length of a vector: $|\mathbf{x}|$,
- angle between vectors: θ,
- dot product: $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}|\,|\mathbf{y}|\cos\theta$.



Vectors: algebraic approach

An n-dimensional coordinate vector is an element of $\mathbb{R}^n$, i.e., an ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of real numbers.

Let $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ be vectors, and $r \in \mathbb{R}$ be a scalar. Then, by definition,

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n),$$

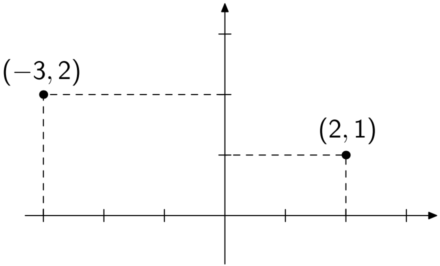
$$r\mathbf{a} = (ra_1, ra_2, \ldots, ra_n),$$

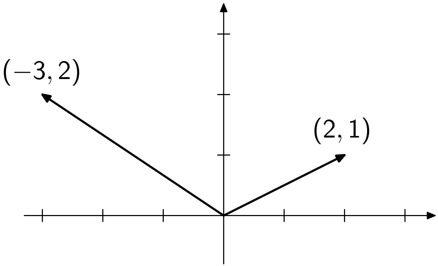
$$\mathbf{0} = (0, 0, \ldots, 0),$$

$$-\mathbf{b} = (-b_1, -b_2, \ldots, -b_n),$$

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}) = (a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n).$$

Cartesian coordinates: geometric meets algebraic





Once we specify an *origin* O, each point A is associated a *position vector* $\overrightarrow{OA}$. Conversely, every vector has a unique representative with tail at O.

Cartesian coordinates allow us to identify a line, a plane, and space with $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$, respectively.

Length and distance

Definition. The **length** of a vector $\mathbf{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

The **distance** between vectors/points $\mathbf{x}$ and $\mathbf{y}$ is $\|\mathbf{y} - \mathbf{x}\|$.

Properties of length:

$\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$\|r\mathbf{x}\| = |r|\,\|\mathbf{x}\|$ (homogeneity)

$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality)

Scalar product

Definition. The **scalar product** of vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ is

$$\boxed{\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.}$$

Properties of scalar product:

$\mathbf{x} \cdot \mathbf{x} \geq 0$, $\mathbf{x} \cdot \mathbf{x} = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ (symmetry)

$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$ (distributive law)

$(r\mathbf{x}) \cdot \mathbf{y} = r(\mathbf{x} \cdot \mathbf{y})$ (homogeneity)

In particular, $\mathbf{x} \cdot \mathbf{y}$ is a **bilinear** function (i.e., it is both a linear function of $\mathbf{x}$ and a linear function of $\mathbf{y}$).

Relations between lengths and scalar products:

$\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}}$

$|\mathbf{x}\cdot\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$ (Cauchy-Schwarz inequality)

$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\,\mathbf{x}\cdot\mathbf{y}$

By the Cauchy-Schwarz inequality, for any nonzero vectors $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ we have

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \quad \text{for some} \quad 0 \leq \theta \leq \pi.$$

θ is called the **angle** between the vectors **x** and **y**. The vectors **x** and **y** are said to be **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\mathbf{x}\cdot\mathbf{y} = 0$ (i.e., if $\theta = 90^{\circ}$).

Problem. Find the angle θ between vectors $\mathbf{x} = (2, -1)$ and $\mathbf{y} = (3, 1)$.

$\mathbf{x} \cdot \mathbf{y} = 5$, $\|\mathbf{x}\| = \sqrt{5}$, $\|\mathbf{y}\| = \sqrt{10}$.

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|} = \frac{5}{\sqrt{5}\,\sqrt{10}} = \frac{1}{\sqrt{2}} \implies \theta = 45^{\circ}$$

Problem. Find the angle ϕ between vectors $\mathbf{v} = (-2, 1, 3)$ and $\mathbf{w} = (4, 5, 1)$.

$$\mathbf{v} \cdot \mathbf{w} = 0 \implies \mathbf{v} \perp \mathbf{w} \implies \phi = 90^{\circ}$$

Orthogonality

Definition 1. Vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\boxed{\mathbf{x} \cdot \mathbf{y} = 0.}$

Definition 2. A vector $\mathbf{x} \in \mathbb{R}^n$ is said to be **orthogonal** to a nonempty set $Y \subset \mathbb{R}^n$ (denoted $\mathbf{x} \perp Y$) if $\mathbf{x} \cdot \mathbf{y} = 0$ for any $\mathbf{y} \in Y$.

Definition 3. Nonempty sets $X, Y \subset \mathbb{R}^n$ are said to be **orthogonal** (denoted $X \perp Y$) if $\mathbf{x} \cdot \mathbf{y} = 0$ for any $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

Examples in $\mathbb{R}^3$. • The line $x = y = 0$ is orthogonal to the line $y = z = 0$.

Indeed, if $\mathbf{v} = (0, 0, z)$ and $\mathbf{w} = (x, 0, 0)$ then $\mathbf{v} \cdot \mathbf{w} = 0$.

• The line $x = y = 0$ is orthogonal to the plane $z = 0$.

Indeed, if $\mathbf{v} = (0, 0, z)$ and $\mathbf{w} = (x, y, 0)$ then $\mathbf{v} \cdot \mathbf{w} = 0$.

• The line $x = y = 0$ is not orthogonal to the plane $z = 1$.

The vector $\mathbf{v} = (0, 0, 1)$ belongs to both the line and the plane, and $\mathbf{v} \cdot \mathbf{v} = 1 \neq 0$.

• The plane $z = 0$ is not orthogonal to the plane $y = 0$.

The vector $\mathbf{v} = (1, 0, 0)$ belongs to both planes and $\mathbf{v} \cdot \mathbf{v} = 1 \neq 0$.

Proposition 1 If $X, Y \in \mathbb{R}^n$ are orthogonal sets then either they are disjoint or $X \cap Y = \{\mathbf{0}\}$.

Proof: $\mathbf{v} \in X \cap Y \implies \mathbf{v} \perp \mathbf{v} \implies \mathbf{v} \cdot \mathbf{v} = 0 \implies \mathbf{v} = \mathbf{0}$.

Proposition 2 Let V be a subspace of $\mathbb{R}^n$ and S be a spanning set for V. Then for any $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} \perp S \implies \mathbf{x} \perp V.$$

Proof: Any $\mathbf{v} \in V$ is represented as $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, where $\mathbf{v}_i \in S$ and $a_i \in \mathbb{R}$. If $\mathbf{x} \perp S$ then

$$\mathbf{x} \cdot \mathbf{v} = a_1(\mathbf{x} \cdot \mathbf{v}_1) + \cdots + a_k(\mathbf{x} \cdot \mathbf{v}_k) = 0 \implies \mathbf{x} \perp \mathbf{v}.$$

Example. The vector $\mathbf{v} = (1, 1, 1)$ is orthogonal to the plane spanned by vectors $\mathbf{w}_1 = (2, -3, 1)$ and $\mathbf{w}_2 = (0, 1, -1)$ (because $\mathbf{v} \cdot \mathbf{w}_1 = \mathbf{v} \cdot \mathbf{w}_2 = 0$).

Orthogonal complement

Definition. Let $S \subset \mathbb{R}^n$. The **orthogonal complement** of S, denoted $S^\perp$, is the set of all vectors $\mathbf{x} \in \mathbb{R}^n$ that are orthogonal to S. That is, $S^\perp$ is the largest subset of $\mathbb{R}^n$ orthogonal to S.

Theorem 1 $S^\perp$ is a subspace of $\mathbb{R}^n$.

Note that $S \subset (S^\perp)^\perp$, hence $\mathrm{Span}(S) \subset (S^\perp)^\perp$.

Theorem 2 $(S^\perp)^\perp = \mathrm{Span}(S)$. In particular, for any subspace V we have $(V^\perp)^\perp = V$.

Example. Consider a line $L = \{(x, 0, 0) \mid x \in \mathbb{R}\}$ and a plane $\Pi = \{(0, y, z) \mid y, z \in \mathbb{R}\}$ in $\mathbb{R}^3$. Then $L^\perp = \Pi$ and $\Pi^\perp = L$.

Fundamental subspaces

Definition. Given an $m \times n$ matrix A, let

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\},$$
$$R(A) = \{\mathbf{b} \in \mathbb{R}^m \mid \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

$R(A)$ is the range of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$, $L(\mathbf{x}) = A\mathbf{x}$. $N(A)$ is the kernel of L.

Also, $N(A)$ is the nullspace of the matrix A while $R(A)$ is the column space of A. The row space of A is $R(A^T)$.

The subspaces $N(A), R(A^T) \subset \mathbb{R}^n$ and $R(A), N(A^T) \subset \mathbb{R}^m$ are **fundamental subspaces** associated to the matrix A.

Theorem $N(A) = R(A^T)^\perp$, $N(A^T) = R(A)^\perp$.
That is, the nullspace of a matrix is the orthogonal complement of its row space.

Proof: The equality $A\mathbf{x} = \mathbf{0}$ means that the vector $\mathbf{x}$ is orthogonal to rows of the matrix A. Therefore $N(A) = S^\perp$, where S is the set of rows of A. It remains to note that $S^\perp = \mathrm{Span}(S)^\perp = R(A^T)^\perp$.

Corollary Let V be a subspace of $\mathbb{R}^n$. Then $\dim V + \dim V^\perp = n$.

Proof: Pick a basis $\mathbf{v}_1, \dots, \mathbf{v}_k$ for V. Let A be the $k \times n$ matrix whose rows are vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. Then $V = R(A^T)$ and $V^\perp = N(A)$. Consequently, $\dim V$ and $\dim V^\perp$ are rank and nullity of A. Therefore $\dim V + \dim V^\perp$ equals the number of columns of A, which is n.

Orthogonal projection

Theorem 1 Let V be a subspace of $\mathbb{R}^n$. Then any vector $\mathbf{x} \in \mathbb{R}^n$ is uniquely represented as $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in V$ and $\mathbf{o} \in V^{\perp}$.

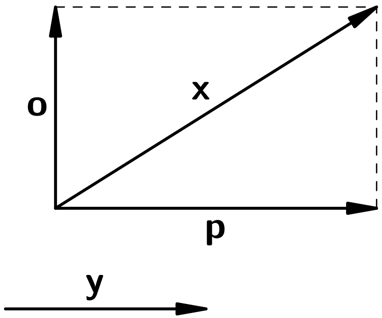
In the above expansion, $\mathbf{p}$ is called the **orthogonal projection** of the vector $\mathbf{x}$ onto the subspace V.

Theorem 2 $\|\mathbf{x} - \mathbf{v}\| > \|\mathbf{x} - \mathbf{p}\|$ for any $\mathbf{v} \neq \mathbf{p}$ in V.

Thus $\|\mathbf{o}\| = \|\mathbf{x} - \mathbf{p}\| = \min_{\mathbf{v} \in V} \|\mathbf{x} - \mathbf{v}\|$ is the **distance** from the vector $\mathbf{x}$ to the subspace V.

Orthogonal projection onto a vector

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, with $\mathbf{y} \neq \mathbf{0}$.
Then there exists a unique decomposition $\mathbf{x} = \mathbf{p} + \mathbf{o}$ such that $\mathbf{p}$ is parallel to $\mathbf{y}$ and $\mathbf{o}$ is orthogonal to $\mathbf{y}$.



$\mathbf{p}$ = orthogonal projection of $\mathbf{x}$ onto $\mathbf{y}$

Orthogonal projection onto a vector

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, with $\mathbf{y} \neq \mathbf{0}$.
Then there exists a unique decomposition $\mathbf{x} = \mathbf{p} + \mathbf{o}$ such that $\mathbf{p}$ is parallel to $\mathbf{y}$ and $\mathbf{o}$ is orthogonal to $\mathbf{y}$.

We have $\mathbf{p} = \alpha\mathbf{y}$ for some $\alpha \in \mathbb{R}$. Then

$$0 = \mathbf{o} \cdot \mathbf{y} = (\mathbf{x} - \alpha\mathbf{y}) \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} - \alpha\mathbf{y} \cdot \mathbf{y}.$$

$$\implies \quad \alpha = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}} \quad \implies \quad \boxed{\mathbf{p} = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}}\,\mathbf{y}}$$

Problem. Find the distance from the point $\mathbf{x} = (3, 1)$ to the line spanned by $\mathbf{y} = (2, -1)$.

Consider the decomposition $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p}$ is parallel to $\mathbf{y}$ while $\mathbf{o} \perp \mathbf{y}$. The required distance is the length of the orthogonal component $\mathbf{o}$.

$$\mathbf{p} = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}} \mathbf{y} = \frac{5}{5} (2, -1) = (2, -1),$$

$$\mathbf{o} = \mathbf{x} - \mathbf{p} = (3, 1) - (2, -1) = (1, 2), \quad \|\mathbf{o}\| = \sqrt{5}.$$

Problem. Find the point on the line $y = -x$ that is closest to the point $(3, 4)$.

The required point is the projection $\mathbf{p}$ of $\mathbf{v} = (3, 4)$ on the vector $\mathbf{w} = (1, -1)$ spanning the line $y = -x$.

$$\mathbf{p} = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w} = \frac{-1}{2} (1, -1) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

Problem. Let Π be the plane spanned by vectors $\mathbf{v}_1 = (1,1,0)$ and $\mathbf{v}_2 = (0,1,1)$.

(i) Find the orthogonal projection of the vector $\mathbf{x} = (4,0,-1)$ onto the plane Π.

(ii) Find the distance from $\mathbf{x}$ to Π.

We have $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in \Pi$ and $\mathbf{o} \perp \Pi$. Then the orthogonal projection of $\mathbf{x}$ onto Π is $\mathbf{p}$ and the distance from $\mathbf{x}$ to Π is $\|\mathbf{o}\|$.

We have $\mathbf{p} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ for some $\alpha, \beta \in \mathbb{R}$. Then $\mathbf{o} = \mathbf{x} - \mathbf{p} = \mathbf{x} - \alpha\mathbf{v}_1 - \beta\mathbf{v}_2$.

$$\begin{cases} \mathbf{o}\cdot\mathbf{v}_1 = 0 \\ \mathbf{o}\cdot\mathbf{v}_2 = 0 \end{cases} \iff \begin{cases} \alpha(\mathbf{v}_1\cdot\mathbf{v}_1) + \beta(\mathbf{v}_2\cdot\mathbf{v}_1) = \mathbf{x}\cdot\mathbf{v}_1 \\ \alpha(\mathbf{v}_1\cdot\mathbf{v}_2) + \beta(\mathbf{v}_2\cdot\mathbf{v}_2) = \mathbf{x}\cdot\mathbf{v}_2 \end{cases}$$

$$\mathbf{x} = (4, 0, -1), \quad \mathbf{v}_1 = (1, 1, 0), \quad \mathbf{v}_2 = (0, 1, 1)$$

$$\begin{cases} \alpha(\mathbf{v}_1 \cdot \mathbf{v}_1) + \beta(\mathbf{v}_2 \cdot \mathbf{v}_1) = \mathbf{x} \cdot \mathbf{v}_1 \\ \alpha(\mathbf{v}_1 \cdot \mathbf{v}_2) + \beta(\mathbf{v}_2 \cdot \mathbf{v}_2) = \mathbf{x} \cdot \mathbf{v}_2 \end{cases}$$

$$\Longleftrightarrow \begin{cases} 2\alpha + \beta = 4 \\ \alpha + 2\beta = -1 \end{cases} \Longleftrightarrow \begin{cases} \alpha = 3 \\ \beta = -2 \end{cases}$$

$$\mathbf{p} = 3\mathbf{v}_1 - 2\mathbf{v}_2 = (3, 1, -2)$$

$$\mathbf{o} = \mathbf{x} - \mathbf{p} = (1, -1, 1)$$

$$\|\mathbf{o}\| = \sqrt{3}$$