# Orthogonality

## Inner product

$$\langle x, y \rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = x^\mathsf{T} y$$

important properties:

- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle \geq 0$
- $\langle x, x \rangle = 0 \iff x = 0$

$f(y) = \langle x, y \rangle$ is linear function : $\mathbb{R}^n \to \mathbb{R}$, with linear map defined by row vector $x^\mathsf{T}$

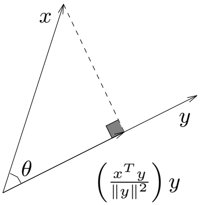
## Cauchy-Schwarz inequality and angle between vectors

for any $x, y \in \mathbb{R}^n$

$$\boxed{|x^{\mathsf{T}}y| \leq \|x\|\|y\|}$$

- (unsigned) angle between vectors in $\mathbb{R}^n$ defined as

$$\theta = \angle(x, y) = \cos^{-1} \frac{x^{\mathsf{T}}y}{\|x\|\|y\|}$$




- thus $x^{\mathsf{T}}y = \|x\|\|y\| \cos\theta$

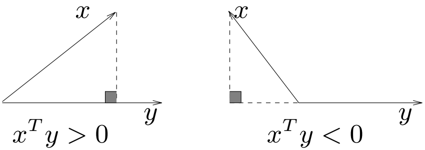
## Special cases

- $x$ and $y$ are *aligned*: $\theta = 0$; $x^T y = \|x\|\|y\|$;
  (if $x \neq 0$) $y = \alpha x$ for some $\alpha \geq 0$
- $x$ and $y$ are *opposed*: $\theta = \pi$; $x^T y = -\|x\|\|y\|$
  (if $x \neq 0$) $y = -\alpha x$ for some $\alpha \geq 0$
- $x$ and $y$ are *orthogonal*: $\theta = \pi/2$ or $-\pi/2$; $x^T y = 0$
  denoted $x \perp y$

## Angles

interpretation of $x^\mathsf{T} y > 0$ and $x^\mathsf{T} y < 0$

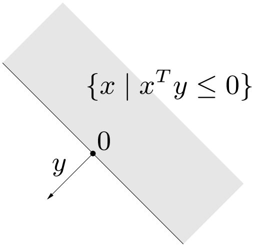
- $x^\mathsf{T} y > 0$ means $\angle(x, y)$ is acute
- $x^\mathsf{T} y < 0$ means $\angle(x, y)$ is obtuse



## Halfspaces

a *halfspace* with outward normal vector $y$, and boundary passing through $0$

$$H = \{x \mid x^{\mathsf{T}} y \leq 0\}$$



## Orthogonal complements

for any set $S \subset \mathbb{R}^n$, the *orthogonal complement* is

$$S^\perp = \{x \mid x^\mathsf{T} y = 0 \text{ for all } y \in S\}$$

- $S^\perp$ is always a subspace
- $S^\perp$ is the set of all vectors $x$, each of which is orthogonal to every vector in $S$

## Orthonormal set of vectors

set of vectors $\{u_1, \ldots, u_k\} \subset \mathbb{R}^n$ is

- *normalized* if $\|u_i\| = 1$, $i = 1, \ldots, k$
  ($u_i$ are called *unit vectors* or *direction vectors*)
- *orthogonal* if $u_i \perp u_j$ for $i \neq j$
- *orthonormal* if both

**slang:** we say '$u_1, \ldots, u_k$ are orthonormal vectors' but orthonormality (like independence) is a property of a *set* of vectors, not vectors individually

in terms of $U = [u_1 \ \cdots \ u_k]$, orthonormal means

$$\boxed{U^\mathsf{T} U = I_k}$$

## Orthonormality

an orthonormal set of vectors is independent

- to see this, multiply $Ux = 0$ by $U^\mathsf{T}$
- hence $\{u_1, \ldots, u_k\}$ is an *orthonormal basis* for

$$\textbf{span}(u_1, \ldots, u_k) = \textbf{range}(U)$$

- **warning**: if $k < n$ then $UU^\mathsf{T} \neq I$ (since its rank is at most $k$)

  (more on this matrix later ...)

## Orthonormal basis for $\mathbb{R}^n$

A matrix $U$ is called *orthogonal* if

$$\boxed{U \text{ is square and } U^\mathsf{T} U = I}$$

- the set of columns $u_1, \ldots, u_n$ is an orthonormal *basis* for $\mathbb{R}^n$
- (you'd think such matrices would be called *orthonormal*, not *orthogonal*)
- it follows that $U^{-1} = U^\mathsf{T}$, and hence also $UU^\mathsf{T} = I$, *i.e.*,

$$\sum_{i=1}^{n} u_i u_i^\mathsf{T} = I$$

## Expansion in orthonormal basis

suppose $U$ is orthogonal, so $x = UU^\mathsf{T} x$, *i.e.*,

$$x = \sum_{i=1}^{n} (u_i^\mathsf{T} x) u_i$$

- $u_i^\mathsf{T} x$ is called the *component* of $x$ in the direction $u_i$
- $a = U^\mathsf{T} x$ *resolves* $x$ into the vector of its $u_i$ components
- $x = Ua$ *reconstitutes* $x$ from its $u_i$ components
- $x = Ua = \sum_{i=1}^{n} a_i u_i$ is called the ($u_i$-) *expansion* of $x$

## Geometric interpretation

if $U$ has orthonormal columns then transformation $w = Uz$

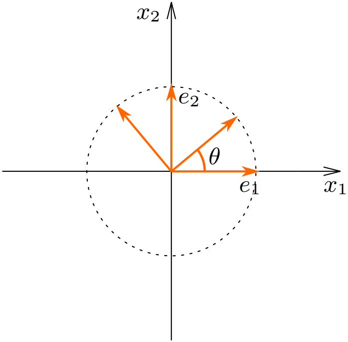
- preserves *norm* of vectors, $i.e.$, $\|Uz\| = \|z\|$
- preserves *angles* between vectors, $i.e.$, $\angle(Uz, U\tilde{z}) = \angle(z, \tilde{z})$
- we say $U$ is *isometric*, it preserves distances

## Example: Rotation

rotation by $\theta$ in $\mathbb{R}^2$ is given by

$$y = U_\theta x, \quad U_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

since $e_1 \mapsto (\cos\theta, \sin\theta)$, $e_2 \mapsto (-\sin\theta, \cos\theta)$



## Example: Reflection

reflection across line $x_2 = x_1 \tan(\theta/2)$ is given by

$$y = R_\theta x, \quad R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

since $e_1 \to (\cos\theta, \sin\theta)$, $e_2 \to (\sin\theta, -\cos\theta)$