MATH 2121 — Lecture #19

Outline:

- Final properties of eigenvalues, diagonalization
- Introduction to inner products, orthogonality, etc.

$\mathrm{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d$ vs $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$

1 Last time: properties of eigenvalues

$\mathrm{tr}(A) = \mathrm{tr}(A^T)$

$\det(A) = \det(A^T)$

The *trace* of a square matrix A is the sum of its diagonal entries.

We denote this by $\mathrm{tr}(A)$. For 2×2 matrices we have $\mathrm{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d$.

Suppose A and B are $n \times n$ matrices. In general $\mathrm{tr}(AB) \neq \mathrm{tr}(A)\mathrm{tr}(B)$, but

$$\boxed{\mathrm{tr}(AB) = \mathrm{tr}(BA)} \quad \text{and} \quad \boxed{\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA).}$$

Let A be an $n \times n$ matrix and write I for the $n \times n$ identity matrix.

The fundamental theorem of algebra says there are numbers $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ such that $\det(A - xI) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$.

Theorem. It holds that $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$ and $\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

$= \det(A)\det(B)\det(C)$

$\det(ABC) = \det(BAC)$

$\mathrm{tr}(ABC) \neq \mathrm{tr}(BAC)$

$\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$

repeat according to exponent of $(\lambda - x)$ in $\det(A - xI)$

The product of the eigenvalues of A, $\boxed{\text{repeated with multiplicity}}$, is the determinant of A, while the sum of the eigenvalues is the trace of A.

For example if $A = \begin{bmatrix} \lambda_1 & a & b \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_3 \end{bmatrix}$ then

$$\det(A - xI) = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x) \quad \text{and} \quad \mathrm{tr}A = \lambda_1 + \lambda_2 + \lambda_3 \quad \text{and} \quad \det A = \lambda_1\lambda_2\lambda_3.$$

A is upper-Δ, so:

① $\det A = \lambda_1 \lambda_2 \lambda_3$ (product of diag)

② $\mathrm{tr}\, A = \lambda_1 + \lambda_2 + \lambda_3$ (sum of diag)

③ $\det(A - xI) = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x)$

as we have $\det(M) = \det(M^T)$ → A and A^T have same characteristic polynomial

Assume A is a square matrix.

A few other properties of eigenvalues and eigenvectors worth noting:

Proposition. A and $A^\top$ have the same eigenvalues.

Proof. Since $\det(A - xI) = \det((A - xI)^\top) = \det(A^\top - xI^\top) = \det(A^\top - xI)$. □

Proposition. A is invertible if and only if 0 is not one of its eigenvalues.

Proof. 0 is an eigenvalue of A if and only if $\det A = 0$. □

$= \lambda_1 \lambda_2 \lambda_3 \cdots \lambda_n$

stronger: eigenspaces of A and A^T have same dimensions (Exercise)

(follows from $\text{rank}(M) = \text{rank}(M^T)$)

if λ is eigval for A

then $\frac{1}{\lambda}$ is eigval for A^{-1}

Proposition. Assume A is invertible.

Then $\boxed{A \text{ and } A^{-1} \text{ have the same eigenvectors,}}$ but reciprocal eigenvalues.

Proof. If $\boxed{Av = \lambda v}$ then $v = A^{-1}Av = A^{-1}\lambda v = \lambda A^{-1}v$ so $\boxed{A^{-1}v = \lambda^{-1}v}$. □

Corollary. If A is invertible and diagonalizable then A^{-1} is diagonalizable.

if $A = PDP^{-1}$

then $A^{-1} = PD^{-1}P^{-1}$

Proof. If A is invertible and diagonalizable, then $\mathbb{R}^n$ has a basis consisting of eigenvectors of A, but this basis is then also made up of eigenvectors of A^{-1}. □

Corollary. If A is diagonalizable then $A^\top$ is diagonalizable.

if $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ then

$D^{-1} = \begin{bmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_n^{-1} \end{bmatrix}$

(also diag)

Proof. Suppose $A = PDP^{-1}$ where D is diagonal. Let $Q = (P^{-1})^\top = (P^\top)^{-1}$.

Then $D^\top = D$ so $A^\top = (PDP^{-1})^\top = \underbrace{(P^{-1})^\top}_{=Q} D^\top \underbrace{P^\top}_{} = QDQ^{-1}$. □

$Q^{-1} = ((P^{-1})^\top)^{-1}$

$= ((P^\top)^{-1})^{-1} = P^\top$

"dot product": a way to get a scalar from two vectors (which measures angles)

Synonym for "perpendicular" (right angle)

2 Inner products and orthogonality

In this lecture, we work with vectors in $\mathbb{R}^n$ and matrices that have all real entries.

Definition. The *inner product* or *dot product* of two vectors (with same size)

$$\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = ac+bd$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$u \cdot v = \underbrace{u^T v}_{\text{matrix mult.}} \in \mathbb{R} \quad \text{(1x1 matrix)}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} -b \\ a \end{bmatrix} = 0$$

in $\mathbb{R}^n$ is the scalar $\boxed{u \bullet v = u_1v_1 + u_2v_2 + \cdots + u_nv_n} = u^\top v = v^\top u = v \bullet u$.

For example, $\begin{bmatrix} a \\ b \end{bmatrix} \bullet \begin{bmatrix} -b \\ a \end{bmatrix} = -ab + ab = 0$ for any $a, b \in \mathbb{R}$.

$v \cdot v = v_1^2 + v_2^2 + \dots + v_n^2 \geq 0$

Definition. The *length* of a vector $v \in \mathbb{R}^n$ is $\boxed{\|v\| = \sqrt{v \bullet v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$.

Note: $u \cdot v = v \cdot u$

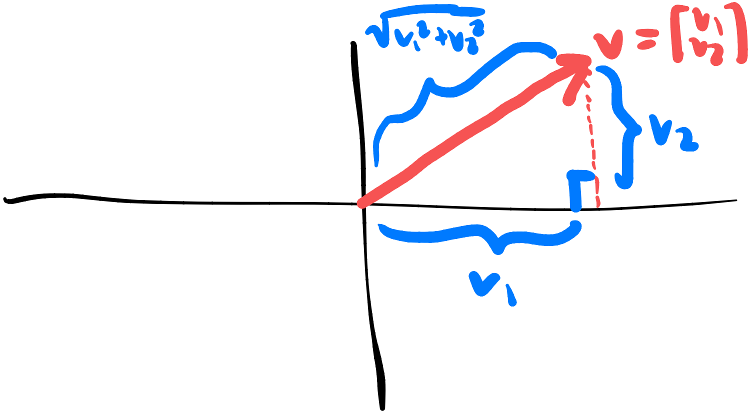
$$u \cdot v = u^T v = (u^T v)^T = v^T u^{T^T} = v^T u = v \cdot u$$

both 1x1

Just defined length of $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$ to be

$$\|v\| \overset{\text{def}}{=} \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \geq 0$$

This definition matches the physical notion of length at least in 2 or 3 dimensions by pythagorean thm:



Notice: can only have $\|v\| = 0$ if $v = 0$

$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \bullet \begin{bmatrix} a \\ b \\ c \end{bmatrix} = ax+by+cz$

$u, v, w \in \mathbb{R}^n$ $c \in \mathbb{R}$

$\bullet$ is a symmetric bilinear operation

<u>Essential properties of length and inner product.</u>

Let $u, v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

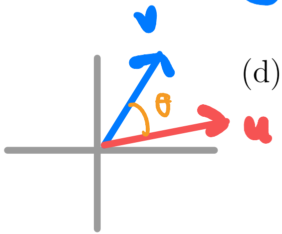
(a) $\boxed{u \bullet v = v \bullet u}$ and $\boxed{(u+v) \bullet w = u \bullet w + v \bullet w}$ and $\boxed{(cv) \bullet w = c(v \bullet w)}$, while

"symmetric"

$\boxed{\|cv\| = |c|\|v\|.}$ multiplying c and v rescales length of v by $|c|$ possibly also reversing orientation of v

(Easy) (b) $v \bullet v = v_1^2 + v_2^2 + \cdots + v_n^2 \geq 0$ and $\|v\| \geq 0$.

(c) $v \bullet v = 0$ if and only if $\|v\| = 0$ if and only if $v = 0 \in \mathbb{R}^n$.



(d) There is a general identity relating $u \bullet v$ to the angle θ between u and v:

$u \bullet v = \|u\| \|v\| \cos\theta$

$$\boxed{u \bullet v = \|u\|\|v\|\cos\theta.}$$ see HW

This holds when $u = 0$ or $v = 0$ (as both sides are 0), although θ is not defined.

Check: $(u+v) \bullet w = (u_1+v_1)w_1 + (u_2+v_2)w_2 + \cdots$

$= u_1w_1 + u_2w_2 + \cdots + v_1w_1 + v_2w_2 + \cdots$

$= u \bullet w + v \bullet w$

Why does $\|cv\| = |c|\,\|v\|$?

$$\|cv\| \overset{\text{def}}{=} \sqrt{(cv)\cdot(cv)} \overset{(1)}{=} \sqrt{c\,(v\cdot(cv))}$$

$$\overset{(2)}{=} \sqrt{c\,((cv)\cdot v)}$$

$$\overset{(3)}{=} \sqrt{c^2(v\cdot v)}$$

$$= \sqrt{c^2}\,\sqrt{v\cdot v}$$

$$= |c|\,\|v\|$$

(1), (3) use $(cv)\cdot w = c(v\cdot w)$

(2) use $u\cdot w = w\cdot u$

matches physical def of distance (between the endpoints of u and v)

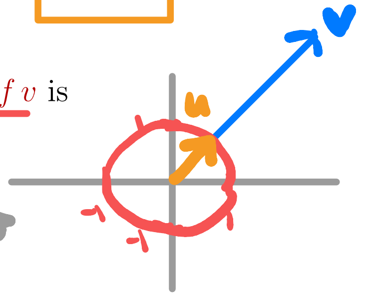
The *distance* between two vectors $u, v \in \mathbb{R}^n$ is the length of the difference $\|u - v\|$.

A *unit vector* is a vector $u \in \mathbb{R}^n$ with $\|u\| = 1$.

If $v \in \mathbb{R}^n$ is any nonzero vector, then the *unit vector in the direction of v* is

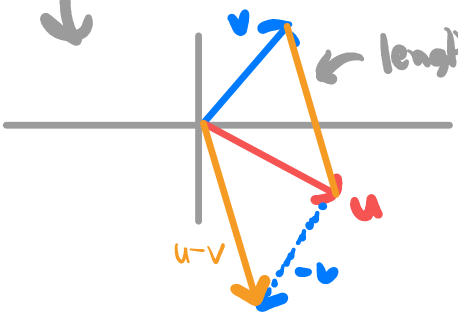
$$\boxed{u = \frac{1}{\|v\|} v \in \mathbb{R}^n.}$$

picture



Note that for this u we have $\|u\| = \|\frac{1}{\|v\|} v\| = |\frac{1}{\|v\|}| \|v\| = \frac{1}{\|v\|} \|v\| = 1$.

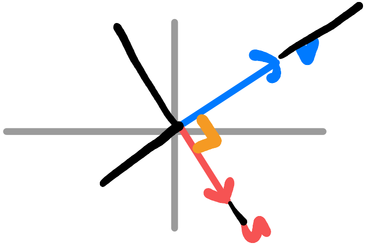
Example. Unit vector in direction of $v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is $u = \frac{1}{\sqrt{1^2+1^2+1^2+1^2}} v = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$.



length of this segment should be the distance between u + v

is this same as $\|u-v\|$?

YES: orange segments have same length, are opposite sides of a parallelogram



1 if they belong to perpendicular lines (form right angle)

Definition. Two vectors $u, v \in \mathbb{R}^n$ are *orthogonal* if $\boxed{u \bullet v = 0.}$ 2

When u and v are orthogonal we also say that "u is orthogonal to v."'

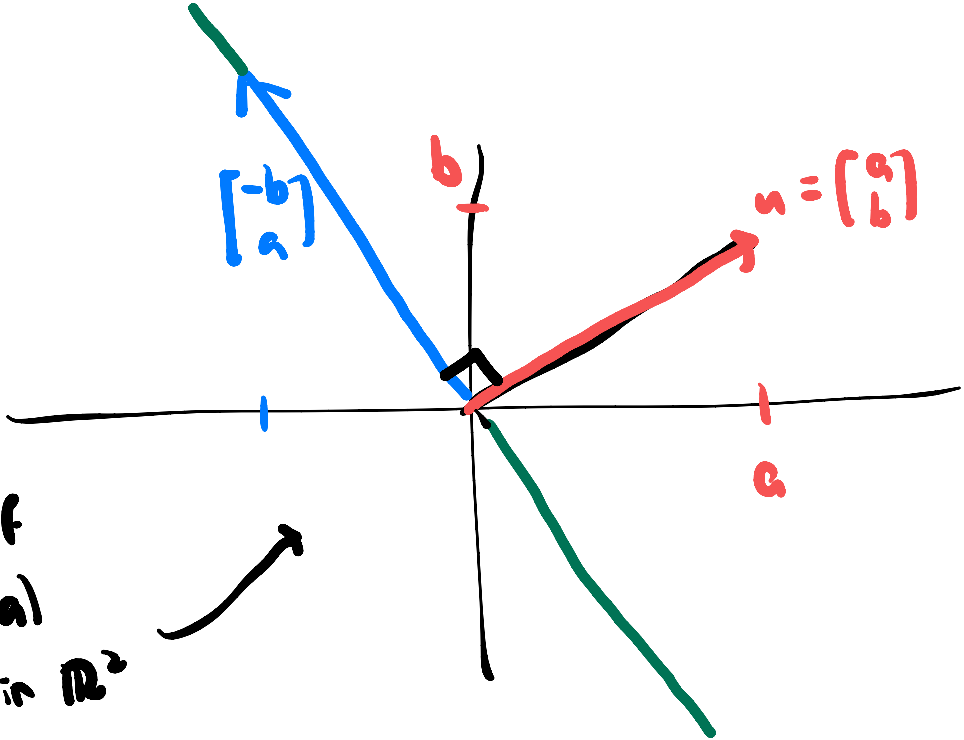
defs 1 and 2 are equiv-alent

Proposition. Suppose $u, v \in \mathbb{R}^2$ are nonzero vectors that are orthogonal to each other, so that $u \bullet v = 0$. Then u and v, drawn as arrows in the xy-plane, belong to perpendicular lines through the origin.

In other words, these vectors are perpendicular in the usual sense of planar geometry.

If $u, v \in \mathbb{R}^2$ are orthogonal and $0 \neq u = \begin{bmatrix} a \\ b \end{bmatrix}$, then v is a scalar multiple $\begin{bmatrix} -b \\ a \end{bmatrix}$, which is the vector obtained by rotating u counterclockwise by 90 degrees.

Proof. This follows directly from the identity $u \bullet v = \|u\|\|v\| \cos\theta$, which implies that $u \bullet v = 0$ if and only if the angle θ between u and v is $\pm\frac{\pi}{2}$. $\square$



picture of orthogonal vectors in $\mathbb{R}^2$

to be orthogonal to u, must be on the green line

(a scalar multiple of $\begin{bmatrix} -b \\ a \end{bmatrix}$)

set of all vectors that are orthogonal to all elements of V

3 Orthogonal complements

Let $V \subseteq \mathbb{R}^n$ be a subspace. The *orthogonal complement* of V is

$$\boxed{V^{\perp} = \{w \in \mathbb{R}^n : v \bullet w = 0 \text{ for all } v \in V\}.}$$

We pronounce "$V^{\perp}$" as "vee perp."

Proposition. If $V \subseteq \mathbb{R}^n$ is a subspace then $V^{\perp} \subseteq \mathbb{R}^n$ is also a subspace.

check rele- cond- ition

Proof. Since $v \bullet 0 = 0$ for all $v \in \mathbb{R}^n$ it holds that $0 \in V^{\perp}$. So $V^{\perp}$ is nonempty. If $x, y \in V^{\perp}$ and $c \in \mathbb{R}$ then

$$v \bullet cx = c(v \bullet x) = 0 \quad \text{and} \quad v \bullet (x + y) = v \bullet x + v \bullet y = 0 + 0 = 0$$

for all $v \in V$ so cx and $x + y$ both belong to $V^{\perp}$. Hence $V^{\perp}$ is a subspace. □

Ex Suppose $n=2$

$$\left\{\begin{bmatrix}0\\0\end{bmatrix}\right\}^\perp \overset{\text{def}}{=} \left\{ v\in\mathbb{R}^2 \;\middle|\; \overbrace{\begin{bmatrix}v_1\\v_2\end{bmatrix}\cdot\begin{bmatrix}0\\0\end{bmatrix}}^{\text{always zero, no matter } v \text{ is}} = 0 \right\} = \mathbb{R}^2$$

$$(\mathbb{R}^2)^\perp \overset{\text{def}}{=} \left\{ v\in\mathbb{R}^2 \;\middle|\; \underbrace{v\cdot w = 0 \text{ for all } w\in\mathbb{R}^2}_{\text{if this holds then (taking } w=v)} \right\} = \left\{\begin{bmatrix}0\\0\end{bmatrix}\right\}$$

we get $v\cdot v=0$ so $v=\begin{bmatrix}0\\0\end{bmatrix}$

Suppose $\begin{bmatrix}a\\b\end{bmatrix}\neq\begin{bmatrix}0\\0\end{bmatrix}$

(the line $\perp$ to first line)

$$\Big(\underbrace{\mathbb{R}\text{-span}\left\{\begin{bmatrix}a\\b\end{bmatrix}\right\}}_{\text{(a line)}}\Big)^\perp \overset{\text{def}}{=} \left\{ v\in\mathbb{R}^2 \;\middle|\; \underbrace{v\cdot c\begin{bmatrix}a\\b\end{bmatrix}=0}_{\text{is true for all } c \text{ if it's true for } c=1} \text{ for all } c\in\mathbb{R} \right\} = \mathbb{R}\text{-span}\left\{\begin{bmatrix}b\\-a\end{bmatrix}\right\}$$

In each case: $(V^\perp)^\perp = V$ which will turn out to be always true

The operation $(\cdot)^\perp$ relates the column space, null space, and transpose of a matrix:

Theorem. Suppose A is an $m \times n$ matrix. $\boxed{\text{Then } (\operatorname{Col} A)^\perp = \operatorname{Nul}(A^\top) \subseteq \mathbb{R}^m.}$

Proof. Write $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ where $a_i \in \mathbb{R}^m$. Let $v \in \mathbb{R}^m$.

If $v \in (\operatorname{Col} A)^\perp$ then we must have $v \bullet a_i = a_i^\top v = 0$ for all i.

Conversely, if $v \bullet a_i = a_i^\top v = 0$ for all i then

$$(c_1 a_1 + c_2 a_2 + \dots + c_n a_n) \bullet v = c_1 \underbrace{(a_1 \bullet v)}_{=0} + c_2 \underbrace{(a_2 \bullet v)}_{=0} + \dots + c_n \underbrace{(a_n \bullet v)}_{=0} = 0$$

for any scalars $c_1, c_2, \dots, c_n \in \mathbb{R}$ so $v \in (\operatorname{Col} A)^\perp$.

Thus $v \in (\operatorname{Col} A)^\perp$ if and only if $v \bullet a_i = a_i^\top v = 0$ for all i. This holds iff $A^\top v = 0$. $\square$

$a_1 \bullet x$

if $A = [a_1\ a_2 \dots]$ then:

$$A^T x = 0 \iff \begin{bmatrix} - a_1^T - \\ - a_2^T - \\ \vdots \end{bmatrix} x = 0 \iff \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \end{bmatrix}$$

$(x \in \operatorname{Nul}(A^T))$

$\iff a_i \bullet x = 0$ for all i $\iff$ x is orthogonal to all columns of A $(x \in (\operatorname{Col} A)^\perp)$

Consequence: how to find a basis for $V^\perp$?

1. express $V = \text{Col}\,A$ for some matrix A
2. find a basis for $\text{Nul}(A^T)$ by usual algorithm.

How to find A in step 1?

find a basis $v_1, v_2, \ldots, v_k$ for V, set $A = [v_1\ v_2 \cdots v_k]$

only vector in both V and $V^\perp$ is zero vector

Lemma. Let $V \subseteq \mathbb{R}^n$ be a subspace. If $w \in V \cap V^\perp$ then $w = 0$.

Proof. If $w \in V$ and $w \in V^\perp$ then $w \bullet w = 0$ so $w = 0$. □

Proposition. Let $V \subseteq \mathbb{R}^n$ be a subspace. If $S \subseteq V$ and $T \subseteq V^\perp$ are two sets of linearly independent vectors, then $S \cup T$ is also linearly independent.

See the lecture notes for a proof. Key idea: if you had a linear dependence among the vectors in $S \cup T$, then some nonzero linear combination of vectors in S would be equal to some nonzero linear combination of vectors in T.

These equal linear combinations would give a nonzero vector in both V and $V^\perp$.

But this is impossible since we just saw that $V \cap V^\perp = \{0\}$.

When S and T are bases for V and $V^\perp$, the previous proposition tells us that $S \cup T$ is a set of linearly independent vectors in $\mathbb{R}^n$ of size $\dim V + \dim V^\perp$. Hence:

Corollary. If $V \subseteq \mathbb{R}^n$ is a subspace then $\boxed{\dim V + \dim V^\perp \leq n.}$

Next time we will see that the inequality $\leq$ here is actually an equality $=$.

4 Orthogonal bases and orthogonal projections

The following proposition is called the *Generalized Pythagorean theorem*.

discuss more next time

Proposition. Two vectors $u, v \in \mathbb{R}^n$ are orthogonal if and only if

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. The proof is just a little algebra:

$$\begin{aligned}
\|u+v\|^2 &= (u+v) \bullet (u+v) \\
&= u \bullet (u+v) + v \bullet (u+v) \\
&= u \bullet u + u \bullet v + v \bullet u + v \bullet v = \|u\|^2 + \|v\|^2 + 2(u \bullet v).
\end{aligned}$$

Then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ if and only if $u \bullet v = 0$.

The equivalence of this proposition to the classical Pythagorean theorem boils down to observation that orthogonal vectors in $\mathbb{R}^2$ form the sides of a right triangle. $\square$

Ex $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$
$e_1 \quad e_2 \quad e_3$

if all pairs in list are orthogonal

Vectors $u_1, u_2, \ldots, u_p \in \mathbb{R}^n$ are *orthogonal* if $u_i \bullet u_j = 0$ whenever $1 \leq i < j \leq p$.

An *orthogonal basis* of $\mathbb{R}^n$ is a basis in which any two vectors are orthogonal.

For example, the standard basis $e_1, e_2, \ldots, e_n$ is an orthogonal basis for $\mathbb{R}^n$.

Theorem. Suppose the vectors $u_1, u_2, \ldots, u_p \in \mathbb{R}^n$ are orthogonal and all nonzero.

Then $u_1, u_2, \ldots, u_p$ are linearly independent.

Proof. Suppose $c_1u_1 + c_2u_2 + \cdots + c_pu_p = 0$ for some coefficients $c_1, c_2, \ldots, c_p \in \mathbb{R}$.

For each $i = 1, 2, \ldots, p$, we then have

$$0 = (c_1u_1 + c_2u_2 + \cdots + c_pu_p) \bullet u_i = c_1(u_1 \bullet u_i) + c_2(u_2 \bullet u_i) + \cdots + c_p(u_p \bullet u_i) = c_i\|u_i\|^2 = c_i(u_i \bullet u_i)$$

since $u_j \bullet u_i = 0$ if $i \neq j$. But since u_i is nonzero, $\|u_i\|^2 \neq 0$, so it must hold that $c_i = 0$. As this applies to each index i, we deduce that $c_1 = c_2 = \cdots = c_p = 0$. $\square$

for orthogonal vectors,

can recover coeffs in any linear comb.

by taking inner products

Corollary. Any nonzero, orthogonal vectors are an orthogonal basis for the subspace they span. (as they are independent by thm)

Corollary. Any n nonzero, orthogonal vectors in $\mathbb{R}^n$ are an orthogonal basis for $\mathbb{R}^n$.

Proposition. Suppose $u_1, u_2, \ldots, u_p$ is an orthogonal basis for a subspace $V \subseteq \mathbb{R}^n$. Let $y \in V$. Then we can write $y = c_1u_1 + c_2u_2 + \cdots + c_pu_p$ where

$$c_i = \frac{y \bullet u_i}{u_i \bullet u_i} = \frac{y \bullet u_i}{\|u_i\|^2}.$$

Proof. A basis must span V, so $y = c_1u_1 + c_2u_2 + \cdots + c_pu_p$ for some $c_1, c_2, \ldots, c_p \in \mathbb{R}$. Since $y \bullet u_i = c_i(u_i \bullet u_i)$ for each $i = 1, 2, \ldots, p$, the result follows. □

used in prev proof

Example. Suppose $u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$ and $u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$ and $u_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$.

You can check that these three vectors are orthogonal.

For example, $u_1 \bullet u_3 = -3/2 - 2 + 7/2 = 0$.

The vectors are therefore linearly independent, so are an orthogonal basis for $\mathbb{R}^3$.

For $y = \begin{bmatrix} 6 \\ 1 \\ 8 \end{bmatrix}$ we have $y \bullet u_1 = 11$ and $y \bullet u_2 = -12$ and $y \bullet u_3 = -33$.

We also have $u_1 \bullet u_1 = 11$ and $u_2 \bullet u_2 = 6$ and $u_3 \bullet u_3 = 33/2$.

Therefore $y = u_1 - 2u_2 - 2u_3$.

Let $u \in \mathbb{R}^n$ be a nonzero vector. Suppose $y \in \mathbb{R}^n$ is any vector.

Definition. The *orthogonal projection* of y onto u is the vector $\boxed{\widehat{y} = \frac{y \bullet u}{u \bullet u} u.}$

The *component of y orthogonal to u* is the vector $\boxed{z = y - \widehat{y} = y - \frac{y \bullet u}{u \bullet u} u.}$

It always holds that $y = \widehat{y} + z$. As its name suggests, we have $z \bullet u = 0$ since

$$z \bullet u = y \bullet u - \frac{y \bullet u}{u \bullet u} u \bullet u = y \bullet u - y \bullet u = 0.$$

Observation. $\widehat{y}$ and z do not change if u is replaced by a nonzero scalar multiple: if we change u to cu for some $0 \neq c \in \mathbb{R}$ then all the factors of c cancel:

$$\frac{y \bullet cu}{cu \bullet cu} cu = \frac{c(y \bullet u)}{c^2(u \bullet u)} cu = \frac{y \bullet u}{u \bullet u} u = \widehat{y}.$$

Let $L = \mathbb{R}\text{-span}\{u\}$. Then $\widehat{y}$ and z may also be called the *orthogonal projection* of y onto L the *component* of y orthogonal to L. We will write $\boxed{\mathrm{proj}_L(y) = \widehat{y} \in L.}$

In $\mathbb{R}^2$, the distance from a point (x, y) to a line $L = \mathbb{R}\text{-span}\{u\}$ is the length

$$\left\| \begin{bmatrix} x \\ y \end{bmatrix} - \operatorname{proj}_L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) \right\|.$$

Example. To find the distance from the point $(x, y) = (7, 6)$ to the line L defined by $y = \frac{1}{2}x$, note that L contains the vector $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Let $w = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$. Then

$$\operatorname{proj}_L\left(\begin{bmatrix} 7 \\ 6 \end{bmatrix}\right) = \frac{w \bullet u}{u \bullet u} u = \frac{28 + 12}{16 + 4} u = \frac{40}{20} u = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

so the distance is

$$\left\| \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\| = \sqrt{1 + 4} = \sqrt{5}.$$