Lecture - 17(a) [Inner product, Orthonormal set]

Inner product

Motivation: $\mathbb{R}^2$

$x = (x_1, x_2)$ $\quad$ $y = (y_1, y_2)$

dot product $\quad x \cdot y = x_1 y_1 + x_2 y_2$

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$x \cdot y = x^T y = x_1 y_1 + x_2 y_2$

Information:

(i) $\quad x \cdot x = x_1^2 + x_2^2$

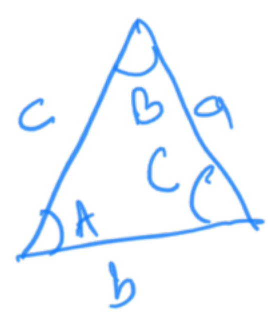
length of $x = \sqrt{x_1^2 + x_2^2}$

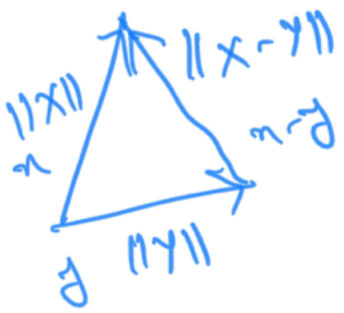
$\|x\|$

(ii) $\quad x \cdot y = \|x\| \, \|y\| \cos\theta$, $\quad \theta \leftarrow$ angle between x and y.

Cosine law:

$c^2 = a^2 + b^2 - 2ab\cos C$.





Question: Can we extend the notion of length and angle to arbitrary vector space?

Observation: Properties satisfied by the dot product in $\mathbb{R}^2$.

i) $x \cdot x = x_1^2 + x_2^2 > 0$ if $x \neq 0$

ii) $x \cdot x = 0 \Leftrightarrow x = 0$

iii) $x \cdot y = y \cdot x$

iv) $(x+y) \cdot z = x \cdot z + y \cdot z$

v) $(\alpha x) \cdot y = \alpha (x \cdot y) \quad \forall \alpha \in \mathbb{R}$.

length of x, $\|x\| = \sqrt{x \cdot x}$

$$\text{angle} = \frac{x \cdot y}{\|x\| \cdot \|y\|}$$

Inner product:

Let $V \leftarrow$ vector space over $\mathbb{F}$
$\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$

A map $V \times V \to \mathbb{F}$ denoted by
$(u, v) \to \langle u | v \rangle$ is called an

inner product on V if

i) $\langle u | u \rangle \geq 0 \quad \forall u \in V$

ii) $\langle u, u \rangle = 0 \Leftrightarrow u = 0$

iii) $\langle u, v\rangle = \langle v, u\rangle \quad \forall\, u, v \in V$

iv) $\langle \alpha u + \beta v | w\rangle = \alpha \langle u|w\rangle + \beta \langle v|w\rangle$

$\forall\, \alpha, \beta \in \mathbb{F}$

and $u, v, w \in V$.

- $\langle u | \alpha v + \beta w\rangle = \bar{\alpha}\langle u|v\rangle + \bar{\beta}\langle u|w\rangle$ (verify).

- A vector space together with an inner product is called an inner product space (i.p.s.)

- $\|v\| = \sqrt{\langle v|v\rangle}$ (length of v)

- angle $= \dfrac{\langle u|v\rangle}{\|u\|\cdot\|v\|}$.

Examples:

(i) $V = \mathbb{R}^n$.

$x = (x_1, x_2, \ldots, x_n)$

$y = (y_1, y_2, \ldots, y_n)$

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \bar{y}_i$$ (standard i.p. on $\mathbb{C}^n$)

(ii) $V = \mathbb{C}^n$ $\quad x = (x_1, x_2, \ldots, x_n)$

$y = (y_1, y_2, \ldots, y_n)$

$$\langle x|y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$ (standard i.p. on $\mathbb{C}^n$).

(iii) $V = C[0,1]$ over $\mathbb{C}$.

Complex valued continuous functions on $[0,1]$.

$$\langle f|g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt.$$

<u>Hint:</u> (i) $\langle f|f\rangle = \int_0^1 |f(t)|^2\,dt \geq 0 \quad \forall f \in V.$

(iii) ✓

(iv) ✓

(ii) $\langle f|f\rangle = 0 \iff f = 0$

$\Leftarrow$ (✓)

$\Rightarrow$

$$\langle f|f\rangle = 0$$
$$\int_0^1 |f(t)|^2\,dt$$

$g(t) = \begin{cases} 1 & \text{if } t = 1/2 \\ 0 & \text{otherwise.} \end{cases}$

$\int_0^1 |g(t)|^2\,dt = 0$ but $g \neq 0$.

<u>Claim:</u> $\langle f|f\rangle = 0 \Rightarrow f = 0$

Indeed, let $\langle f|f\rangle = 0$, suppose that $\exists\ t_0 \in [0,1]$ such that $f(t_0) \neq 0$

$\because$ f is cont^n $\Rightarrow \exists\ \delta > 0$ s.t. $f(t) \neq 0$ $\forall t \in [t_0-\delta, t_0+\delta]$

$$\Rightarrow \int_0^1 |f(t)|^2 dt \geq \int_{t_0-\delta}^{t_0+\delta} |f(t)|^2 dt > 0$$

Contradiction //.

(ii) $V = P_n(\mathbb{R})$ $\leftarrow$ polynomials with real coefficients of degree $\leq n$.

$$\langle p | q \rangle = \int_a^b p(t) \cdot q(t)\, dt$$ is an i.p. on V.

(Verify).

Exercise: (i) characterize all inner products on $\mathbb{R}$.

Hint: Let $\langle \cdot | \cdot \rangle$ on $\mathbb{R}$

for any $a, b \in \mathbb{R}$

$$\langle a | b \rangle = a \langle 1 | b \rangle = ab \underbrace{\langle 1 | 1 \rangle}_{\alpha}$$

Therefore all i.p.'s on $\mathbb{R}$ are of the form $\langle a | b \rangle = \alpha ab$ for some $\alpha > 0$.

(ii) Characterize all i.p.'s on $\mathbb{R}^2$.
[Do it yourself]

Orthogonal / orthonormal set:

Let V be i.p.s. over $\mathbb{F}$

Let $S \subseteq V$, $u, v \in V$

- u, v are said to be orthogonal if
  $\langle u | v \rangle = 0$
- S is called an orthogonal set if for all distinct pairs $u, v \in S$, we have
  $$\langle u | v \rangle = 0$$
- S is called orthonormal if
  i) S is orthogonal
  ii) $\|v\| = 1 \quad \forall \ v \in S$.
- when u, v are orthogonal, we sometimes also write $u \perp v$.

Examples: (i) $V = \mathbb{F}^n$ with standard i.p.

$S = \{e_1, \dots, e_K\}$, $K \leq n$, is an o.n. set

$e_i = (0, \dots, 0, \underset{\uparrow}{1}, 0, \dots 0)$.

i^{th} _ Position

(ii) $V = \mathbb{R}^2$, $x_1 = (1, 0)$

$x_2 = (0, 2)$

Then $\{x_1, x_2\}$ is orthogonal but not orthonormal.

<u>Observation</u>: If $\{v_1, \dots, v_K\}$ is orthogonal in an i.p.s. V. Then $\left\{\frac{v_1}{\|v_1\|}, \dots, \frac{v_K}{\|v_K\|}\right\}$ is orthonormal in V.

<u>Result</u>: Every orthogonal set of nonzero vectors is linearly independent.

<u>Proof</u>: Let S be orthogonal.

Let $\{v_1, \dots v_K\} \subseteq S$.

$$\underbrace{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_K v_K}_{u} = 0$$

$$\Rightarrow \langle u | v_i \rangle = 0$$

$$\langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \mid v_i \rangle$$

$$\|$$

$$\alpha_i$$

$$\Rightarrow \quad \alpha_i = 0 \quad \forall\, i.$$

$$\Rightarrow \quad S \text{ is L.I.}$$

//.

<u>Observation:</u> If $\{v_1, v_2, \dots v_k\}$ is orthogonal in an i.p.s. V, then $\left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \dots, \frac{v_k}{\|v_k\|} \right\}$ is orthonormal.

<u>Result:</u> Every orthogonal set of nonzero vectors is L.I.

<u>Proof:</u> Let S be orthogonal.

<u>claim</u>: S is L.I.

Let $\{v_1, \dots v_k\} \subseteq S$.

$$\underbrace{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k}_{u} = 0$$

for each i $\quad \langle u \mid v_i \rangle = 0$

$$\|$$

$$\langle \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k \mid v_i \rangle$$
$$\|$$
$$\alpha_i \langle v_i \mid v_i \rangle$$

$$\Rightarrow \quad \alpha_i \underbrace{\langle v_i \mid v_i \rangle} = 0$$

$$\Rightarrow \quad \alpha_i = 0 \qquad \forall\, i$$

$\Rightarrow$ S is L.I. //

Lecture-17(b) [Gram-Schmidt orthogonalization

(Inner product continuing)

<u>Exercise:</u> Characterize all i.p.'s in $\mathbb{R}^2$.

<u>Hints:</u> Let $\langle \cdot | \cdot \rangle$ be an i.p. in $\mathbb{R}^2$.

$x = (x_1, x_2)$, $y = (y_1, y_2)$

$$\langle x | y \rangle = \langle (x_1, x_2) \mid (y_1, y_2) \rangle$$
$$= \langle x_1 e_1 + x_2 e_2 \mid y_1 e_1 + y_2 e_2 \rangle$$

$\left\{ \begin{array}{l} e_1 = (1,0) \\ e_2 = (0,1). \end{array} \right.$

$$= x_1 y_1 \underbrace{\langle e_1 | e_1 \rangle}_{a} + x_1 y_2 \underbrace{\langle e_1 | e_2 \rangle}_{b} + x_2 y_1 \underbrace{\langle e_2 | e_1 \rangle}_{c} + x_2 y_2 \underbrace{\langle e_2 | e_2 \rangle}_{d} \quad (*)$$

$a > 0$, $b = c$, $d > 0$

Also,

$$\langle x | x \rangle = a x_1^2 + 2b x_1 x_2 + d x_2^2 > 0 \quad \forall\, x = (x_1, x_2) \in \mathbb{R}^2,\ x \neq 0$$

$$x = (x_1, 1)$$

$$\langle x | x \rangle = \underbrace{a x_1^2 + 2b x_1 + d}_{} > 0 \quad \forall\, x_1 \in \mathbb{R}.$$

$$\Rightarrow \quad ad - b^2 > 0$$

any i.p. in $\mathbb{R}^2$ satisfies $(*)$ with $a, b, d \in \mathbb{R}$ such that $a > 0$, $ad - b^2 > 0$

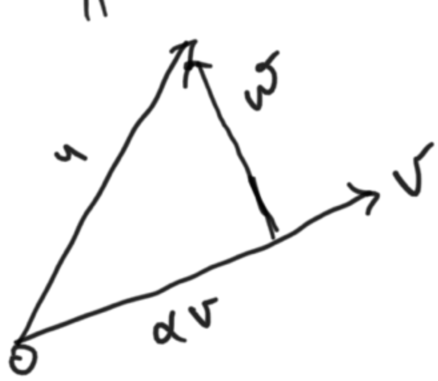
Conversely, for any $a, b, d \in \mathbb{R}$ such that $a > 0$, $ad - b^2 > 0$, then the function $\langle \cdot | \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\boxed{\langle x | y \rangle = a x_1 y_1 + b (x_1 y_2 + x_2 y_1) + d (x_2 y_2)}, \quad x = (x_1, x_2),\ y = (y_1, y_2)$$

defines an i.p. in $\mathbb{R}^2$ (Verify).

Observation: $V \leftarrow$ i.p.s over $\mathbb{F}$

$u, v \in V$



Suppose we want to write

u as a scalar multiple of v plus a vector w orthogonal to v, as suggested in the picture, that means

we want $\quad u = \alpha v + w \quad , \ \alpha \in \mathbb{F}$

$\langle w | v \rangle = 0$

$$w = u - \alpha v$$

to find α, $\quad \langle w | v \rangle = 0$

$$\Rightarrow \langle u - \alpha v | v \rangle = 0$$

$$\Rightarrow \langle u | v \rangle - \alpha \langle v | v \rangle = 0$$

$$\Rightarrow \alpha = \frac{\langle u | v \rangle}{\langle v | v \rangle}$$

$$\Rightarrow u = \underbrace{\frac{\langle u|v\rangle}{\langle v|v\rangle} \cdot v}_{\alpha v} + \underbrace{u - \frac{\langle u|v\rangle}{\langle v|v\rangle} \cdot v}_{w}$$

. Thus if $v \neq 0$, then u can always be written as a scalar multiple of v plus a vector orthogonal to v.

Cauchy - Schwarz inequality:

Let $V \leftarrow$ i.p.s over $\mathbb{F}$

$u, v \in V$

Then $$|\langle u|v\rangle| \leq \|u\|\cdot\|v\| \quad \text{——} (*) .$$

and equality holds in $(*)$ if and only if u and v are linearly dependent.

Proof: If $v=0$, then $(*)$ holds trivially.

Assume that $v \neq 0$.

Then $u = \alpha v + w$, where

$$\alpha = \frac{\langle u|v\rangle}{\langle v|v\rangle} \quad \text{and} \quad w = u - \alpha v.$$

$$\begin{aligned}
0 \leq \langle w|w\rangle &= \langle u-\alpha v \,|\, u-\alpha v\rangle \\
&= \langle u|u\rangle - \alpha\langle v|u\rangle - \bar{\alpha}\langle u|v\rangle \\
&\quad + |\alpha|^2 \langle v|v\rangle \\
&= \langle u|u\rangle - \frac{|\langle v|u\rangle|^2}{\langle v|v\rangle} - \cancel{\frac{|\langle u|v\rangle|^2}{\langle v|v\rangle}} \\
&\quad + \cancel{\frac{|\langle v|u\rangle|^2}{\langle v|v\rangle^2}} \cdot \cancel{\langle v|v\rangle} \\
&= \langle u|u\rangle - \frac{|\langle v|u\rangle|^2}{\langle v|v\rangle} \quad \text{——} (i)
\end{aligned}$$

$$\Rightarrow \quad \frac{|\langle v|u\rangle|^2}{\langle v|v\rangle} \leq \langle u|u\rangle$$

$\Rightarrow \quad |\langle v|u\rangle|^2 \le \langle u,u\rangle \cdot \langle v,v\rangle$

$\Rightarrow \quad \boxed{|\langle v|u\rangle| \le \|u\| \cdot \|v\|}$

from (i) $\quad w \neq 0 \quad \Rightarrow \quad |\langle v|u\rangle| < \|u\| \cdot \|v\|$

Thus equality holds in $(*)$ iff $w = 0$

iff $u - \alpha v = 0$

iff $u = \alpha v$

i.e. u and v are L.D. //

Result: [Gram-Schmidth orthogonalization Process].

Let $V \leftarrow$ i.p.s.,

$\{v_1, \ldots, v_n\} \subseteq V$ be L.I. set

Then one may construct an orthogonal set

$\{q_1, q_2, \ldots q_m\} \subseteq V$ such that

$$\operatorname{Span}(\{v_1, \ldots, v_k\}) = \operatorname{Span}(\{q_1, \ldots q_k\}) \quad \forall\ k = 1, 2, \ldots n.$$

Proof: $q_1 = v_1$, clearly $\operatorname{Span}(\{v_1\}) = \operatorname{Span}(\{q_1\})$

Set $q_2 = v_2 - \frac{\langle v_2 | q_1 \rangle}{\langle q_1 | q_1 \rangle} q_1$

Clearly. $q_2 \perp q_1$

and $\mathrm{Span}(\{q_1, q_2\}) = \mathrm{Span}(\{v_1, v_2\})$

Set $q_3 = v_3 - \frac{\langle v_3 | q_1 \rangle}{\langle q_1 | q_1 \rangle} q_1 - \frac{\langle v_3 | q_2 \rangle}{\langle q_2 | q_2 \rangle} \cdot q_2$

Clearly. $q_3 \perp q_1$ & $q_3 \perp q_2$

and

$\mathrm{Span}(\{q_1, q_2, q_3\}) = \mathrm{Span}(\{v_1, v_2, v_3\})$

(verify)

By induction:

Suppose for $1 \le m < n$ an orthogonal set has been chosen such that for every k, $1 \le k \le m$

$\mathrm{Span}(\{q_1, \ldots, q_k\}) = \mathrm{Span}(\{v_1, \ldots, v_k\})$.

Recall,

q_1, v_2

$v_2 = \alpha q_1 + q_2$,

$q_2 \perp q_1$

$q_2 = v_2 - \frac{\langle v_2 | q_1 \rangle}{\langle q_1 | q_1 \rangle} q_1$

Clearly, $q_2 \perp q_1$

and $\mathrm{Span}\{v_1, v_2\} = \mathrm{Span}\{q_1, q_2\}$

Generalize this idea

v_3, q_1, q_2

$v_3 = \alpha q_1 + \beta q_2 + q_3$,

where

$q_3 \perp q_1$ & $q_3 \perp q_2$.

$\Downarrow$

$\alpha = \frac{\langle v_3 | q_1 \rangle}{\langle q_1 | q_1 \rangle}$

$\beta = \frac{\langle v_3 | q_2 \rangle}{\langle q_2 | q_2 \rangle}$

$q_3 = v_3 - \alpha q_1 - \beta q_2$

To construct the next vector, let

$$q_{m+1} = v_{m+1} - \sum_{j=1}^{m} \frac{\langle v_{m+1} | q_j \rangle}{\langle q_j | q_j \rangle} \cdot q_j \quad (*)$$

Then (i) $q_{m+1} \neq 0$ (✓)

ii) $\langle q_{m+1} | q_j \rangle = 0 \quad \forall\, j = 1, \ldots m.$ (✓)

iii) $\mathrm{Span}(\{q_1, \ldots, q_{m+1}\}) = \mathrm{Span}(\{v_1, \ldots v_{m+1}\})$

(Verify).

<u>Note</u>: One can first construct $\{q_1, \ldots q_n\}$ an orthogonal set and at the end

$$\left\{ \frac{q_1}{\|q_1\|}, \ldots \frac{q_n}{\|q_n\|} \right\} \leftarrow \text{o.n. set.}$$

<u>Example</u>: Consider the vectors

$v_1 = (3,0,4)$, $v_2 = (-1,0,7)$, $v_3 = (2,9,11)$.

Apply Gram-Schmidth on $\{v_1, v_2, v_3\}$ to construct an orthonormal set.

<u>Sol</u>:

$q_1 = v_1 = (3,0,4)$

$$q_2 = v_2 - \frac{\langle v_2 | q_1 \rangle}{\langle q_1 | q_1 \rangle} \cdot q_1$$

$$= (-1,0,7) - \frac{\langle (-1,0,7) \mid (3,0,4) \rangle}{\langle (3,0,4) \mid (3,0,4) \rangle} \cdot (3,0,4)$$

$$= (-1,0,7) - \frac{(-3+28)}{(9+16)} \cdot (3,0,4)$$

$$= (-1,0,7) - (3,0,4)$$

$$= (-4,0,3)$$, clearly $q_1 \perp q_2$.

$$q_3 = v_3 - \frac{\langle v_3 | q_1 \rangle}{\langle q_1 | q_1 \rangle} q_1 - \frac{\langle v_3 | q_2 \rangle}{\langle q_2 | q_2 \rangle} q_2$$

$$= (0,9,0)$$

$\{q_1, q_2, q_3\}$ is an orthogonal set.

$$\left\{ \frac{1}{5}(3,0,4),\ \frac{1}{5}(-4,0,3),\ \frac{1}{9}(0,9,0) \right\}$$

← is an orthonormal set of vectors.