Reminder from last Time

Triangle inequality in an inner product space $(V, \langle\cdot,\cdot\rangle)$: for every $x, y, z \in V$

$$\underset{\|x-z\|}{\overset{\| \text{ by def.}}{\mathrm{dist}(x,z)}} \leq \underset{\|x-y\|}{\overset{\|}{\mathrm{dist}(x,y)}} + \underset{\|y-z\|}{\overset{\|}{\mathrm{dist}(y,z)}}$$

Most commonly, we write this as

Triangle inequality for the norm:

for every $u, w \in V$,

$$\|u+w\| \leq \|u\| + \|w\|$$

$$\Updownarrow$$

$$\|u+w\|^2 \leq \left(\|u\| + \|w\|\right)^2 \quad (*)$$

|| ← recall the proof

$$\langle u+w, u+w\rangle = \|u\|^2 + \|w\|^2 + \langle u,w\rangle + \overline{\langle u,w\rangle}$$

$$\leq \|u\|^2 + \|w\|^2 + 2|\langle u,w\rangle|$$

at which point we can use the Cauchy-Schwarz inequality to upper-bound this by the RHS of $(*)$

Most convenient case to work with: when $\langle u,w\rangle$ (and hence also $\langle w,u\rangle$) is equal to zero, that is, when u, w are orthogonal.

Recall

Definition Let $(V, \langle\cdot,\cdot\rangle)$ be a (real or complex) inner product space, and let $u, w \in V$.
We say that u, w are orthogonal if
$$\langle u, w\rangle = 0.$$
We then write $u \perp w$.

Note clearly the zero vector 0_V is orthogonal to any $u \in V$.

Thm (known as the Pythagorean Theorem) Let $(V, \langle\cdot,\cdot\rangle)$ be a (real or complex) inner product space, and let u, w be orthogonal vectors in V.
Then
$$\|u+w\|^2 = \|u\|^2 + \|w\|^2$$

Useful Remark If $u \perp w$, then for every $\lambda, \mu \in \mathbb{F} (= \mathbb{R} \text{ or } \mathbb{C})$ we have that
$$\lambda u, \mu w \text{ are also orthogonal.}$$
Indeed, $\langle \lambda u, \mu w\rangle = \lambda \langle u, \mu w\rangle = \lambda\bar{\mu} \cdot \langle u, w\rangle = \lambda\bar{\mu} \cdot 0 = 0$.
Thus by the Pythagorean Theorem
$$\|\lambda u + \mu w\|^2 = \|\lambda u\|^2 + \|\mu w\|^2 \underset{\substack{\uparrow \\ \text{remark} \\ \text{from last time} \\ \text{on positive homogeneity} \\ \text{of norm}}}{=} |\lambda|^2 \cdot \|u\|^2 + |\mu|^2 \|w\|^2.$$

This suggests that, if u, w are orthogonal, then we can easily find the norm of any linear combination of u and w.

This motivates the following notion.

Definition Let $(V, \langle \cdot, \cdot \rangle)$ be a (real or complex) inner product space. A subset $\{u_1, u_2, \ldots u_m\}$ of V is called an orthogonal set if, whenever $1 \le i, j \le m$ and $i \neq j$, we have $u_i \perp u_j$.

Remark Clearly we can extend this definition to infinite sets as well: a subset $\{u_i : i \in I\}$ of V is called an orthogonal set if, whenever $i_1, i_2 \in I$ and $i_1 \neq i_2$, we have $u_{i_1} \perp u_{i_2}$.

Very Important: $\boxed{\text{Proposition 1}}$ If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, and if $T \subseteq V$ is an orthogonal set that contains only non-zero vectors, then T is a linearly independent subset of V.

Note This shows that the size of an orthogonal set of non-zero vectors cannot exceed the dimension of the space.

Proposition 2 If $(V, \langle \cdot, \cdot \rangle)$ is an inner product space, and if $\{u_1, u_2, \ldots, u_m\}$ is an orthogonal subset of V, then for every $\lambda_1, \lambda_2, \ldots \lambda_m \in \mathbb{F}$ we have

$$\|\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_m u_m\|^2 = |\lambda_1|^2 \|u_1\|^2 + |\lambda_2|^2 \|u_2\|^2 + \cdots + |\lambda_m|^2 \|u_m\|^2$$

Proof It suffices to use the Pythagorean Theorem and induction.

By linearity of the inner product, we have

$$\langle \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_{m-1} u_{m-1}, \lambda_m u_m \rangle$$

$$= \langle \lambda_1 u_1, \lambda_m u_m \rangle + \langle \lambda_2 u_2, \lambda_m u_m \rangle + \dots + \langle \lambda_{m-1} u_{m-1}, \lambda_m u_m \rangle$$

$$= \lambda_1 \cdot \bar{\lambda}_m \overset{0}{\cancel{\langle u_1, u_m \rangle}} + \lambda_2 \bar{\lambda}_m \overset{0}{\cancel{\langle u_2, u_m \rangle}} + \dots + \lambda_{m-1} \bar{\lambda}_m \overset{0}{\cancel{\langle u_{m-1}, u_m \rangle}}$$

$$= 0.$$

Thus $\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_{m-1} u_{m-1} \perp \lambda_m u_m$ and by the Pythagorean Theorem we get

$$\|(\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_{m-1} u_{m-1}) + \lambda_m u_m\|^2$$
$$= \|\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_{m-1} u_{m-1}\|^2 + \|\lambda_m u_m\|^2$$
$$= \|\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_{m-1} u_{m-1}\|^2 + |\lambda_m|^2 \cdot \|u_m\|^2.$$

Continuing similarly with this summand now...

$$= \dots = |\lambda_1|^2 \|u_1\|^2 + |\lambda_2|^2 \|u_2\|^2 + \dots + |\lambda_m|^2 \|u_m\|^2.$$

<u>Proof of Proposition 1 based on Proposition 2</u>

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $T = \{u_1, u_2, \dots u_m\}$ be an orthogonal subset of V with only non-zero vectors.

Consider $\lambda_1, \lambda_2, \dots \lambda_m \in \mathbb{F}$ $(= \mathbb{R}$ or $\mathbb{C})$ such that

$$0_V = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_m u_m.$$

But then, by Proposition 2 and by properties of the norm,

$$0 = \|0_V\|^2 = \|\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_m u_m\|^2$$
$$= |\lambda_1|^2 \|u_1\|^2 + |\lambda_2|^2 \cdot \|u_2\|^2 + \dots + |\lambda_m|^2 \|u_m\|^2$$

We now also use the assumption that the u_i are non-zero vectors: recall that this gives us that $\|u_i\| \neq 0$ for every $1 \leq i \leq m$.

Thus $0 = |\lambda_1|^2 \cdot \|u_1\|^2 + |\lambda_2|^2 \|u_2\|^2 + \dots + |\lambda_m|^2 \|u_m\|^2$

$\Rightarrow$ for every $1 \leq i \leq m$ $|\lambda_i|^2 \cdot \|u_i\|^2 = 0$

$\Rightarrow$ for every $1 \leq i \leq m$ $|\lambda_i|^2 = 0 \Rightarrow$ for every $1 \leq i \leq m$

since $\|u_i\|^2 \neq 0$ $\lambda_i = 0$.

This shows that $T = \{u_1, u_2, \dots, u_m\}$ is linearly independent

---

Suppose now that such a set is also a spanning set of the space. Then for every $x \in V$ we can find $\mu_1, \mu_2, \dots \mu_m \in \mathbb{F}$ such that

$$x = \mu_1 u_1 + \mu_2 u_2 + \dots + \mu_m u_m.$$

$$\Rightarrow \|x\|^2 = |\mu_1|^2 \|u_1\|^2 + |\mu_2|^2 \|u_2\|^2 + \dots + |\mu_m|^2 \|u_m\|^2$$

and thus the norm of x only depends on the coefficients μ_i and the norms of the u_i.

If moreover we had that $\|u_i\| = 1$ for all $1 \leq i \leq m$, then finding the norm of x would become even simpler.

This motivates the following

Definition Let $(V, <\cdot,\cdot>)$ be an inner product space. A subset $T = \{u_i : i \in I\}$ of V is called

orthonormal

if T is orthogonal and if $\|u_i\| = 1$ for every $i \in I$.

Remark If $T = \{u_1, u_2, \ldots u_m\}$ is an orthogonal set of non-zero vectors, then we can immediately construct a related orthonormal set:
the set $\tilde{T} = \left\{ \frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, \ldots, \frac{u_m}{\|u_m\|} \right\}$ is orthonormal.

Main Theorem A' Let $(V, <\cdot,\cdot>)$ be a (real or complex) inner product space. Then V has an orthogonal (or orthonormal) basis B (that is, there exists a basis B of V that is also an orthogonal (or orthonormal respectively) set).

In fact there is also an "algorithm"/constructive way of finding such nice bases of V once we have found a basis C of V.

Gram-Schmidt Orthogonalisation Process

Let $(V, <\cdot,\cdot>)$ be an inner product space, and suppose $C = \{x_1, x_2, \ldots, x_k\}$ is a basis of V.
Then we can find an orthonormal basis $B = \{u_1, u_2, \ldots, u_k\}$ of V as follows:

Recursive Definition

set $u_1 = \dfrac{x_1}{\|x_1\|}$

set $\tilde{u}_2 = x_2 - \langle x_2, u_1\rangle \cdot u_1$ and $u_2 = \dfrac{\tilde{u}_2}{\|\tilde{u}_2\|}$.

set $\tilde{u}_3 = x_3 - \langle x_3, u_2\rangle \cdot u_2 - \langle x_3, u_1\rangle \cdot u_1$

and $u_3 = \dfrac{\tilde{u}_3}{\|\tilde{u}_3\|}$

$\vdots$

If $u_1, u_2, \dots u_s$ have already been found so that $\{u_1, u_2, \dots, u_s\}$ is an orthonormal set and so that $\text{span}(\{u_1, u_2, \dots, u_s\}) = \text{span}(\{x_1, x_2, \dots x_s\})$, where $1 \le s < k$, then

set $\tilde{u}_{s+1} = x_{s+1} - \langle x_{s+1}, u_s\rangle \cdot u_s - \langle x_{s+1}, u_{s-1}\rangle \cdot u_{s-1} - \dots - \langle x_{s+1}, u_2\rangle \cdot u_2 - \langle x_{s+1}, u_1\rangle \cdot u_1$

and $u_{s+1} = \dfrac{\tilde{u}_{s+1}}{\|\tilde{u}_{s+1}\|}$.

Practice: Verify using induction that, at each step of the "algorithm" / process, the set $\{u_1, \dots u_s\}$ that has already been constructed is orthonormal and satisfies $\text{span}(\{u_1, u_2, \dots u_s\}) = \text{span}(\{x_1, x_2, \dots x_s\})$. Thus at the final step we will have that

$\text{span}(\{u_1, u_2, \dots, u_k\}) = \text{Span}(\{x_1, x_2, \dots, x_k\}) = V$,

and so $B = \{u_1, u_2, \dots u_k\}$ will be an orthonormal basis of V