Week 6: Inner Product Space / Gram-Schmidt Process (textbook §6.1, 6.2)

Norms

Def$^{\underline{n}}$: Let $(V, \langle , \rangle)$ be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

We define the norm / length of a vector $x \in V$ as

$$\boxed{\|x\| := \sqrt{\langle x, x \rangle}}$$

Remarks: ① Positivity of $\langle , \rangle \Rightarrow \| \cdot \|$ is well-defined

② When $(V, \langle , \rangle)$ = standard inner product on $\mathbb{R}^n$, $\| \cdot \|$ is the usual length of a vector.

③ Using "conjugate bilinearity" & "positivity" of $\langle , \rangle$:

- $\|cx\| = |c| \, \|x\| \quad \forall \, c \in \mathbb{F}, \ x \in V$
- $\|x\| \geq 0$ and "=" holds iff $x = 0$

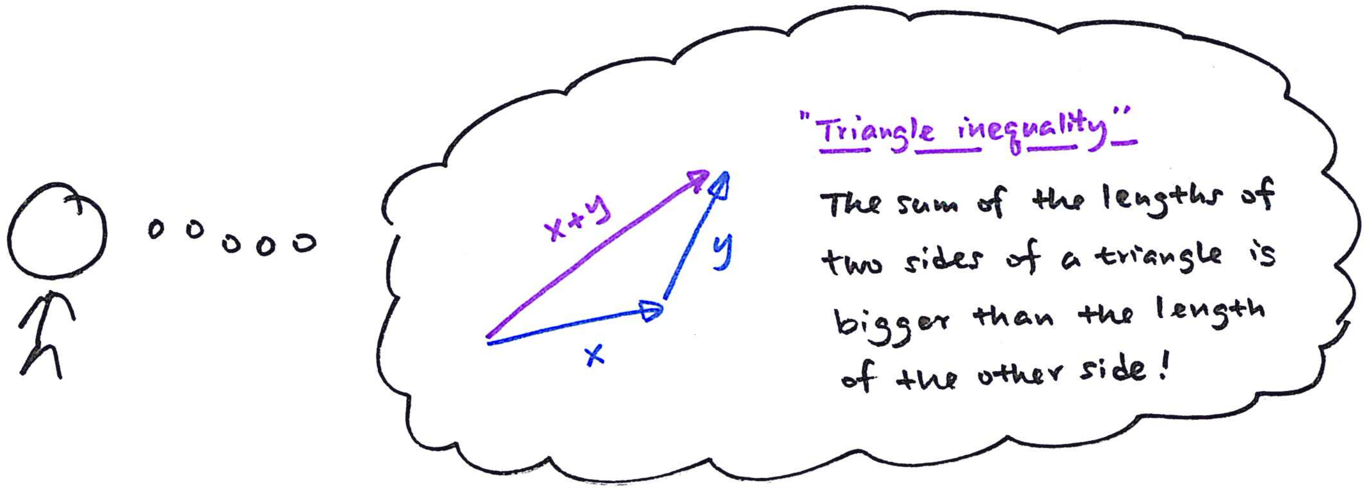
Two Important Inequalities

For any inner product space $(V, \langle , \rangle)$,

Cauchy-Schwarz: $\boxed{|\langle x, y \rangle| \leq \|x\| \cdot \|y\|}$

Triangle Inequality: $\boxed{\|x + y\| \leq \|x\| + \|y\|}$

Q: What do these inequalities say geometrically?

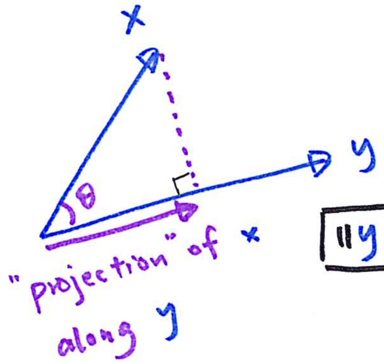


What about Cauchy-Schwarz inequality? $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$

Assume $\|y\| = 1$: (which can be achieved by rescaling unless $y = 0$)

Recall that for the standard dot product in $\mathbb{R}^n$, we have

$$\boxed{\langle x,y\rangle = \|x\|\,\|y\|\cos\theta} \quad \text{cosine law.}$$



$\boxed{\|y\| = 1}$

when $\|y\| = 1$,

$$|\langle x,y\rangle| = \|x\|\,\underbrace{|\cos\theta|}_{\leq 1}$$

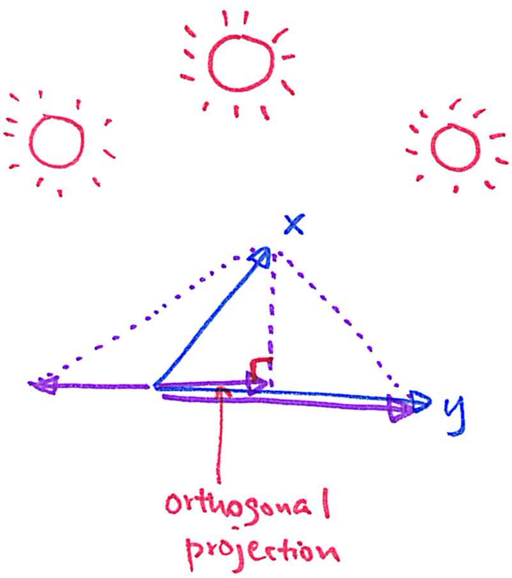
$=$ length of "projection" of x along y

$\leq \|x\|$

Therefore,

Cauchy-Schwarz inequality	$\Leftrightarrow$	The length of the "projection" of a vector is shorter than the length of the original vector.

In fact, the statement on the right above is not yet totally precise since there are many different ways to "project" a vector.



Different position of the sun yields different "projections"

BUT, among all these "projections" there is ONE which is most "efficient" called the

"Orthogonal projection"

Proof of Cauchy-Schwarz : $\boxed{|\langle x, y\rangle| \leq \|x\| \cdot \|y\|}$

When $y = 0$ $\Rightarrow$ trivial.

When $y \neq 0$ $\Rightarrow$ WLOG, can assume $\|y\| = 1$.

For <u>any</u> $c \in \mathbb{F}$, consider

$$0 \leq \|x - cy\|^2 \qquad \text{(positivity)}$$

$$= \langle x - cy, x - cy\rangle \qquad (\text{def}^{\underline{n}} \text{ of norm})$$

$$= \langle x, x\rangle - c\langle y, x\rangle - \bar{c}\langle x, y\rangle + |c|^2 \langle y, y\rangle$$

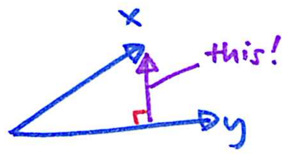
In particular, if we take $\boxed{c = \dfrac{\langle x, y\rangle}{\langle y, y\rangle}}$,

$$0 \leq \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2} - \cancel{\frac{|\langle x,y\rangle|^2}{\|y\|^2}} + \cancel{\frac{|\langle x,y\rangle|^2}{\|y\|^2}}$$

rearranging gives $|\langle x, y\rangle|^2 \leq \|x\|^2 \|y\|^2$.

$\square$

What is the vector $x - cy$ for this particular c chosen?



Proof of Triangle inequality : $\boxed{\|x + y\| \leq \|x\| + \|y\|}$

$$\|x+y\|^2 = \langle x+y, x+y\rangle$$

$$= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

$$= \|x\|^2 + \underbrace{\langle x, y\rangle + \overline{\langle x, y\rangle}}_{2\,\mathrm{Re}\,\langle x, y\rangle} + \|y\|^2$$

$$\leq \|x\|^2 + 2\,|\langle x, y\rangle| + \|y\|^2$$

$\wedge\!\!\!|$ Cauchy-Schwarz ineq.

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

$\square$

Exercise: When does "=" hold in Cauchy-Schwarz and Triangle inequality? (Think geometrically!)

Orthogonality

We have seen that an inner product $\langle \, , \rangle$ can be used to define a norm $\| \cdot \|$ which satisfies expected properties and allows us to define distance between two vectors x and y by

$$\boxed{\text{dist}(x, y) := \| x - y \|}$$

(If you are familiar with point set topology, this distance function $\text{dist}(\cdot, \cdot)$ makes V a metric space.)

Exercise: Although any inner product $\langle \, , \rangle$ gives a norm $\| \cdot \|$, not all the "norms" come from an inner product!! (see #25 in §6.1)

Therefore, $\langle \, , \rangle$ gives not only the concept of length / distance, but also the extra information of angle between two vectors. In particular, we can say when two vectors are "perpendicular".

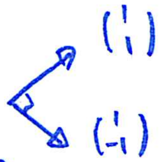
Def$^{\underline{n}}$: Let $(V, \langle \, , \rangle)$ be an inner product space.

$x, y \in V$ are orthogonal $\Leftrightarrow$ $\boxed{\langle x, y \rangle = 0}$.

A vector $x \in V$ is a unit vector if $\boxed{\| x \| = 1}$.

- $S \subset V$ is orthogonal if any $x, y \in S$ are orthogonal.
- $S \subset V$ is orthonormal if all $x \in S$ are unit vectors and S is orthogonal.

Example: In $\mathbb{R}^2$ with standard inner product,



$$S = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \text{ is orthogonal } \left(\because \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 - 1 = 0 \right)$$

but not orthonormal since $\left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\| = \sqrt{2} \neq 1$.

To get an orthonormal subset from an orthogonal subset, we just normalize the vectors by rescaling it to unit length.

$$S' = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \text{ is orthonormal !!}$$

Normalization: $x \neq 0 \Rightarrow \boxed{y = \frac{1}{\|x\|} x}$ is a unit vector.

Def$^{\underline{n}}$: An orthonormal basis β for an inner product space is ① an ordered basis which is ② orthonormal.

Example: (1) $\beta = \{ e_1, e_2, \cdots, e_n \}$ the standard basis is an orthonormal basis for $\mathbb{R}^n$ (or $\mathbb{C}^n$).

(2) S' in the above example is an orthonormal basis for $\mathbb{R}^2$.

(3) $V = C([0, 2\pi])$; $\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} \, dt$ $\boxed{e^{i\theta} = \cos\theta + i \sin\theta}$

complex-valued continuous functions on $[0, 2\pi]$ Euler

$$\beta = \{ f_n(t) := e^{int} \mid n \in \mathbb{Z} \} \text{ is an orthonormal subset !}$$

Check: $\langle f_n, f_m \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{int} \cdot e^{-imt} \, dt \overset{(Ex)}{=} \begin{cases} 1 & \text{if } n = m. \\ 0 & \text{if } n \neq m. \end{cases}$

! This example is of utmost importance in Fourier analysis !.

In fact, β is an orthonormal basis (in suitable sense as V is ∞-dimensional.)

Q: Why do we like orthonormal basis?

Recall that if $\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for a finite-dim'l vector space V, then any $v \in V$ can be expressed uniquely as

$$(*) \quad \boxed{v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n} \qquad \text{where} \quad a_i \in \mathbb{F}$$

To calculate the "coefficients" a_i, we have to solve a system of linear equations (which takes time!).

However, if β is orthonormal, then $\boxed{a_i = \langle v, v_i \rangle}$

just calculate an inner product!!

Why? β is orthonormal $\Leftrightarrow$ $\boxed{\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}}$

"Kronecker delta"

Taking inner product with v_i on both sides of $(*)$, using linearity

$$\langle v, v_i \rangle = a_1 \langle v_1, v_i \rangle + a_2 \langle v_2, v_i \rangle + \cdots + a_n \langle v_n, v_i \rangle$$

Since $\langle v_i, v_j \rangle = \delta_{ij}$, we get $a_i = \langle v, v_i \rangle$.

A: It is easy to find the coefficients by "expanding" a vector in an orthonormal basis!

The idea above actually proves the following:

Thm: Let $S = \{v_1, v_2, \dots, v_k\}$ be an orthogonal subset of non-zero vectors in an inner product space $(V, \langle , \rangle)$.

If $y \in \operatorname{span} S$, then $\boxed{y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\|v_i\|^2} v_i}$

Proof: Exercise!

Corollary: $\boxed{\text{orthogonal} + \text{nonzero vectors} \Rightarrow \text{linearly independent}}$

Gram-Schmidt Orthogonalization Process

Recall that:

Prop: An <u>orthogonal</u> subset of <u>nonzero</u> vectors S is linearly indep.

Pf: Suppose

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \qquad \text{—— } (**)$$

for some $a_i \in \mathbb{F}$, $v_i \in S \subset V$. We claim that all $a_i = 0$.

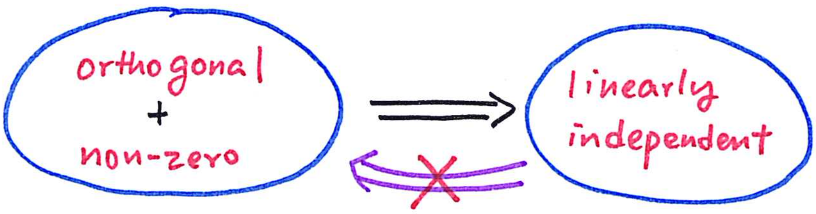
Since S is an orthogonal subset, we have $\langle v_i, v_j \rangle = 0$ whenever $i \neq j$, and that $\langle v_i, v_i \rangle \neq 0$ since $v_i \neq 0$.

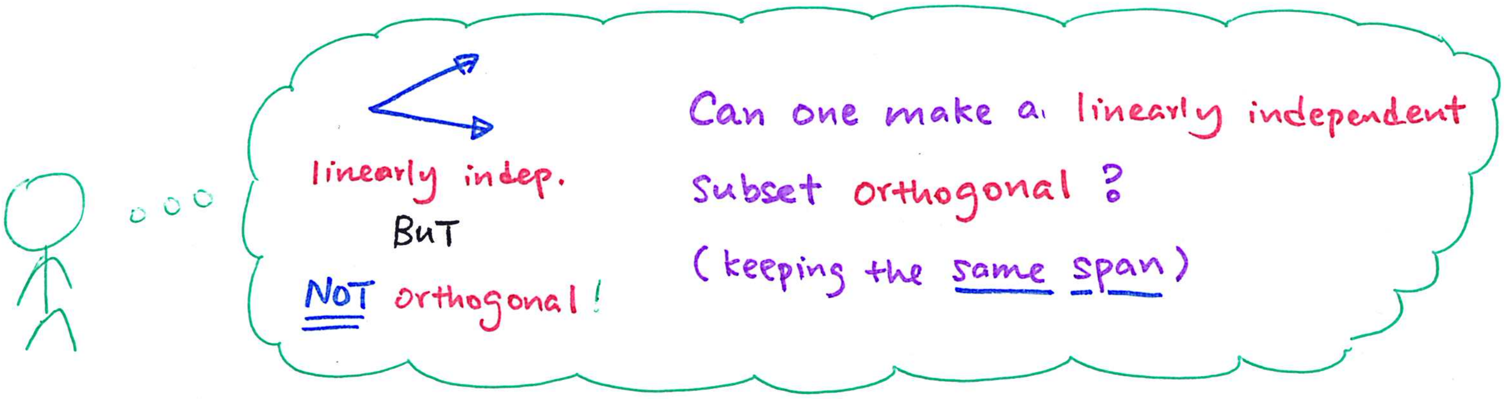
Therefore, $(**)$ implies that

$$a_i \langle v_i, v_i \rangle = \langle a_1 v_1 + a_2 v_2 + \cdots + a_n v_n, v_i \rangle = \langle \underset{\text{vector}}{0}, v_i \rangle = \underset{\text{scalar}}{0}.$$

Since $\langle v_i, v_i \rangle \neq 0$, we must have $a_i = 0$.

—— $\square$

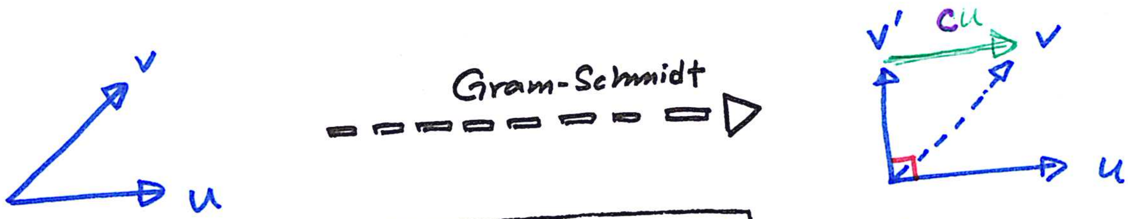




Yes! $\Longrightarrow$ <u>Gram-Schmidt Process</u> !

Let's begin with the simplest case.

Consider two independent vectors u and v



$S = \{u, v\}$ linearly independent

$\boxed{\text{span } S = \text{span } S'}$

$S' = \{u, v'\}$ orthogonal

GOAL: Choose a suitable $c \in \mathbb{F}$ s.t.

$$v' = v - cu \perp u$$

i.e. $\langle v - cu, u \rangle = 0 \Rightarrow \boxed{c = \dfrac{\langle v, u \rangle}{\langle u, u \rangle}}$

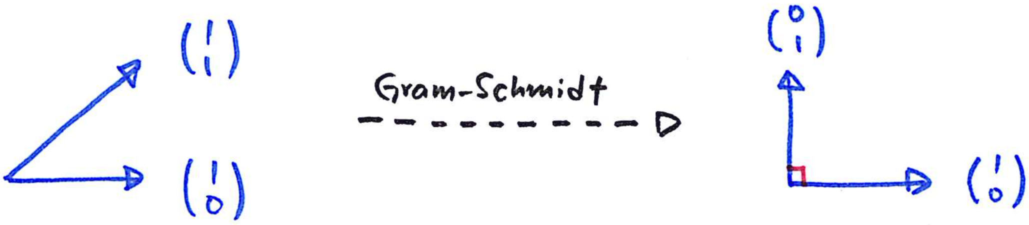
Example: $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $S = \{u, v\} \subset \mathbb{R}^2$ linearly independent

$$\left.\begin{aligned} \langle u, v \rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \\ \langle u, u \rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \end{aligned}\right\} \Rightarrow c = \frac{\langle v, u \rangle}{\langle u, u \rangle} = 1.$$

Define $v' = v - cu = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Then, $S' = \{u, v'\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is orthogonal and

span S = span S'.



What if there are more than two vectors?

Just repeat the same procedure, "correcting" one vector at a time.

Example: Consider the linearly independent subset $S = \{w_1, w_2, w_3\}$ of $\mathbb{R}^3$ where:

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Apply Gram-Schmidt process:

$$\begin{cases} v_1 = w_1 \\ v_2 = w_2 - \dfrac{\langle w_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \quad \Rightarrow \quad \boxed{v_2 \perp v_1 \ \& \ \operatorname{span}\{w_1, w_2\} = \operatorname{span}\{v_1, v_2\}} \\ v_3 = w_3 - \dfrac{\langle w_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \dfrac{\langle w_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 \quad \Rightarrow \quad \boxed{\begin{array}{l} v_3 \perp v_1, \quad v_3 \perp v_2 \\ \operatorname{span}\{w_1, w_2, w_3\} = \operatorname{span}\{v_1, v_2, v_3\} \end{array}} \end{cases}$$

For our example,

$$v_1 = w_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

$$v_2 = w_2 - \frac{3}{5} v_1 = \begin{pmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ 1 \end{pmatrix} \perp v_1$$

$$v_3 = w_3 - \frac{1}{5} v_1 - \frac{1}{3} v_2 = \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} \perp v_1 \ \& \ v_2$$

Therefore, we obtain an orthogonal subset

$$S' = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} \right\}$$

to obtain an orthonormal subset, we just do a normalization

$$S'' = \left\{ \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{30}} \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \right\}.$$

To summarize, we have the following theorem:

Thm (Gram-Schmidt Orthogonalization)

Given any linearly independent subset $\{w_1, w_2, \dots, w_k\}$ in an inner product space $(V, \langle\,,\,\rangle)$, there exists an orthogonal subset $\{v_1, v_2, \dots, v_k\}$ s.t. $\text{span}\{w_1, w_2, \dots, w_i\} = \text{span}\{v_1, v_2, \dots, v_i\}$ for $i = 1, \dots, k$.

More explicitly, one can take (defined inductively)

$$\begin{cases} v_1 = w_1 \\ v_2 = w_2 - \dfrac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \\ \vdots \\ v_k = w_k - \displaystyle\sum_{j=1}^{k-1} \dfrac{\langle w_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j \end{cases}$$

new vectors

Proof: By induction on k. $\square$

In particular, if we start with an ordered basis:

$$\boxed{\text{ANY ordered basis}} \xrightarrow[\text{Schmidt}]{\text{Gram-}} \boxed{\text{orthogonal basis}} \xrightarrow{\text{normalization}} \boxed{\text{orthonormal basis}}$$

Corollary: Any nonzero finite dimensional inner product space has an orthonormal basis.

If $\beta = \{v_1, v_2, \dots, v_n\}$ is an orthonormal basis, then any $v \in V$ has the expansion

$$v = \sum_{i=1}^{n} \underbrace{\langle v, v_i\rangle}_{} v_i$$

Fourier coefficients of v relative to β

These are called Fourier coefficients because in a Fourier series expansion:

$$f(t) = b_0 + a_1 \sin t + b_1 \cos t + a_2 \sin 2t + b_2 \cos 2t$$

$$+ \cdots + a_n \sin(nt) + b_n \cos(nt) + \cdots\cdots$$

where $a_k = \frac{1}{2\pi}\int_0^{2\pi} f(t) \sin(nt)\,dt$

$b_k = \frac{1}{2\pi}\int_0^{2\pi} f(t) \cos(nt)\,dt$

are the coefficients w.r.t. the orthonormal "basis" $\{\sin nt, \cos mt\}_{\substack{n=1,2,\ldots \\ m=0,1,2,\ldots}}$

One more example: Consider the inner product space

$$V = P_2(\mathbb{R}) \subseteq C([-1,1]) \quad ; \quad \langle f, g\rangle := \int_{-1}^{1} f(t)\,g(t)\,dt.$$

polynomials with degree ≤ 2

Let $\beta = \{1, x, x^2\}$ be the standard ordered basis for $P_2(\mathbb{R})$.

Find an orthonormal basis for $P_2(\mathbb{R})$ by Gram-Schmidt process.

Define $v_1 = 1$. Then $\langle x, 1\rangle = \int_{-1}^{1} t \cdot 1\,dt = 0$

$\langle 1, 1\rangle = \int_{-1}^{1} 1 \cdot 1\,dt = 2$

Define $v_2 = x - 0 \cdot 1 = x$. Then.

$\langle x^2, 1\rangle = \int_{-1}^{1} t^2 \cdot 1\,dt = \frac{2}{3}$, $\langle x, x\rangle = \int_{-1}^{1} t \cdot t\,dt = \frac{2}{3}$

$\langle x^2, x\rangle = \int_{-1}^{1} t^2 \cdot t\,dt = 0$

Define $v_3 = x^2 - \frac{1}{3} \cdot 1 - 0 \cdot x = x^2 - \frac{1}{3}$.

Hence, $\{1, x, x^2 - \frac{1}{3}\}$ is an orthogonal basis.

Finally, to get an orthonormal basis, we need to normalize the vectors to unit length.

$$\langle 1, 1 \rangle = 2$$

$$\langle x, x \rangle = \frac{2}{3}$$

$$\langle x^2 - \tfrac{1}{3}, x^2 - \tfrac{1}{3} \rangle = \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 dt = \int_{-1}^{1} (t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9}) dt =$$

$$= \frac{t^5}{5} - \frac{2t^3}{9} + \frac{t}{9} \Big|_{t=-1}^{1} = \frac{8}{45}$$

Therefore, we get an orthonormal basis

$$\beta' = \left\{ \frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}\, x, \sqrt{\frac{5}{8}}\,(3x^2 - 1) \right\} \text{ for } P_2(\mathbb{R})$$

These are called Legendre polynomials (up to degree 2).