Recall that so far we have let scalars multiply vectors and also let operators/matrices "multiply" vectors (or rather act on). Heck we know that scalars can multiply scalars, and we can combine operators, i.e. $ABx = Cx$. So maybe we should let vectors multiply other vectors. Hence the inner product.

An inner-product in a real or complex vector space is a scalar valued function of the ordered pair of vectors x and y s.t.

1. $(x, y) = (y, x)^*$ [If they are real then order doesn't matter]
2. $(\alpha x + \beta y, z) = \alpha^*(x, z) + \beta^*(y, z)$ w/ α, β scalars
3. $(x, x) \geq 0$ for any x; $(x, x) = 0 \Rightarrow x = 0$ "positive-definite, which can be relaxed"

If real $\Rightarrow$ Euclidean, complex $\Rightarrow$ unitary spaces.

With this in hand, we also get a way to define the "length" of a vector (which was hitherto missing):

$\|x\| = (x, x)^{\frac{1}{2}}$ Which is part of the reason for $*$ in above $\Rightarrow$ $\|x\|$ is always real!

We can use it to define the angle between two vectors (in any number of dimensions):

$$\theta = \cos^{-1}\left\{\frac{(x, y)}{\|x\|\,\|y\|}\right\}$$

And furthermore we can use it to define orthogonality:

$(x, y) = 0 \Rightarrow x$ and y are orthogonal

And orthonormality:

The set $\{x_1, x_2, \dots, x_n\}$ is orthonormal if $(x_i, x_j) = \delta_{ij}$ for any i and j

Examples:

1. For $\mathbb{R}^n$ this is the good old dot product: $x = \begin{pmatrix} a \\ b \end{pmatrix}$, $y = \begin{pmatrix} c \\ d \end{pmatrix} \Rightarrow (x, y) = \overbrace{ac + bd}^{\text{matrix language } (a\ b)\begin{pmatrix} c \\ d \end{pmatrix}} = (y, x)$
2. For $\mathbb{C}^n$ this is almost the same: $x = \begin{pmatrix} a \\ b \end{pmatrix}$, $y = \begin{pmatrix} c \\ d \end{pmatrix} \Rightarrow (x, y) = \underbrace{a^*c - b^*d}_{(a\ b)^*\begin{pmatrix} c \\ d \end{pmatrix}} = (y, x)^*$

In both (1) and (2) we are using that the length of basis vectors is 1 and they are orthogonal. That is $x = a\hat{i} + b\hat{j}$ $y = c\hat{i} + d\hat{j} \Rightarrow (x, y) = ac\,\hat{i}\cdot\hat{i} + ad\,\hat{i}\cdot\hat{j} + bc\,\hat{j}\cdot\hat{i} + bd\,\hat{j}\cdot\hat{j}$

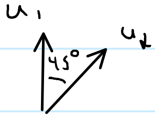
$= ac + bd$

3. Sometimes we might make it a bit more complicated:

So for $x = \begin{pmatrix} a \\ b \end{pmatrix}$, $y = \begin{pmatrix} c \\ d \end{pmatrix}$ over $\mathbb{R}$, $(x, y) = (a\ b)\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}\begin{pmatrix} c \\ d \end{pmatrix} = ac + 2bd = (y, x)$

This plays the role of the metric and says that the length of the second basis is $\sqrt{2}$

4. Or we could be working in terms of a non-orthogonal basis.

Consider in $\mathbb{R}^2$ the basis 

which "works" since the basis is linearly independent.

if $x = a u_1 + b u_2$ and $y = c u_1 + d u_2$ then:

$$(x,y) = ac\underbrace{(u_1,u_1)}_{\neq 0} + ad\underbrace{(u_1,u_2)}_{\neq 0} + bc\underbrace{(u_2,u_1)}_{\neq 0} + bd\underbrace{(u_2,u_2)}_{\neq 0}$$

5. And other times we have to be a bit more careful.

Consider P_1 with $t \in [0,1]$ and $x = \alpha_0 + \alpha_1 t$, $y = \beta_0 + \beta_1 t$

We might consider $(x,y) = \alpha_0\beta_0 + \alpha_1\beta_1$ but what about the basis lengths?

And what about orthogonality?

We will use the following definition: $(x,y) = \int_0^1 x^*(t)\, y(t)\, dt$

Clearly, applying this gives: $(x,y) = \int_0^1 (\alpha_0 + \alpha_1 t)(\beta_0 + \beta_1 t)\,dt$

$$= \alpha_0\beta_0 + \tfrac{1}{2}\alpha_0\beta_1 + \tfrac{1}{2}\alpha_1\beta_0 + \tfrac{1}{3}\alpha_1\beta_1$$

There are two ways to get at this.

a) In the $(1,t)$ basis we note that $(1,t) = \int_0^1 t\,dt = \frac{1}{2}$ so 1 and t are not orthogonal. This means that (x,y) in that basis will have $\neq 0$ cross-terms.

Note: $(1,t)$ is a linearly independent basis, but not orthogonal.

b) In the $(1+\sqrt{3}t,\ 1-\sqrt{3}t)$ basis we have $(1+\sqrt{3}t,\ 1-\sqrt{3}t) = 0$, so this is an orthogonal basis. Better still $(1+\sqrt{3}t,\ 1+\sqrt{3}t) = 2+\sqrt{3}$, so if we instead use $\left(\underbrace{\frac{1+\sqrt{3}t}{\sqrt{2+\sqrt{3}}}}_{u(t)},\ \underbrace{\frac{1-\sqrt{3}t}{\sqrt{2+\sqrt{3}}}}_{v(t)}\right)$, we have an orthonormal basis.

If we express the vector in this basis (for which some work is required) then $x = \alpha_u u + \alpha_v v$, $y = \beta_u u + \beta_v v \Rightarrow (x,y) = \alpha_u\beta_u + \alpha_v\beta_v$ cross-terms vanish

6. We could imagine other definitions of the inner-product for functions including changing integration bounds $\int_a^b$ and also adding in a weighting function, e.g. $(x,y) = \int_0^1 g(t)\, x^*(t)\, y(t)\, dt$

The inner product defined on functions and the convenience of working in an orthonormal basis is one of the ideas that we will be developing in detail. Of course you have already been using them, e.g. Fourier modes, Legendre polynomials, Bessel functions, etc.
We are just going to dive a bit deeper into their construction, hopefully to provide an underlying theme which "unifies" the various results.
Also, clearly we will want to extend these finite dimensional results to infinite ones so that we can cover more functions.

So let's get back to developing inner-products in general.

[In a finite-dimensional vector space, an orthonormal set is complete if it is not contained in any larger orthonormal set.

Now that we have definitions, let's look at some results.

[An orthonormal set is linearly independent (though not the reverse), hence provides a basis (and a complete one if it is complete).

Recall that if $\sum_i \alpha_i x_i = 0 \Rightarrow \alpha_i = 0$ then x_i is linearly independent.

Well consider $(x_j, 0) = 0 = (x_j, \sum_i \alpha_i x_i) = \sum_i \alpha_i \underbrace{(x_j, x_i) = \sum_i \alpha_i \delta_{ij}}_{\text{orthonormality}} = \alpha_j = 0$

So if a set is orthonormal then $\sum_i \alpha_i x_i = 0 \Rightarrow \alpha_i = 0$, hence linearly independent.

[Bessel's inequality: If x_i is a finite orthormal set in an inner-product space and x is any vector, then $\|x\|^2 \geq \sum_i |\alpha_i|^2$ where $\alpha_i = (x_i, x)$.
Furthermore the vector $x' = x - \sum_j \alpha_j x_j$ is orthogonal to each x_i.

Now wait, why is there an inequality? Isn't $\|x\|^2 = \sum_i |\alpha_i|^2$ for an orthonormal basis? $x = \sum_i \alpha_i x_i$ w/ $(x_i, x_j) = \delta_{ij} \Rightarrow \|x\|^2 = \sum_i |\alpha_i|^2$

The answer is that in this x_i is an orthonormal set, not neccessarily a complete basis. So it might be that $x = \sum_i \alpha_i x_i + \underbrace{\sum_\nu \beta_\nu y_\nu}_{\text{notice this is } x'}$ which leads to $\|x\|^2 > \sum_i |\alpha_i|^2$.

Then of course $(x_i, x') = (x_i, x) - (x_i, \sum_j \alpha_j x_j) = (x_i, x) - \sum_j \alpha_j (x_i, x_j)$
$= \alpha_i - \alpha_i = 0$ [book has $\alpha_i^* - \alpha_i^* = 0$ but they do (x', x_i)]

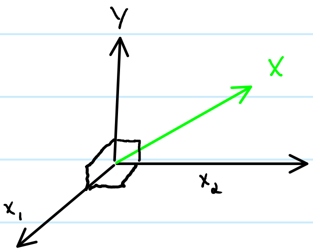
So the question is "when do we know if our orthonormal set X is a complete basis of V?"

Any of these say yes:

1. X is complete (if not contained in larger set)
2. if $(x_i, x) = 0$ for all i $\Rightarrow$ $x = 0$
3. X spans V (every vector in V is a linear combination of elements of X)
4. if $x \in V$, then $x = \sum_i (x_i, x) x_i$
5. if $x, y \in V$ then $(y, x) = \sum_i (y, x_i)(x_i, x)$ "Parseval's equation"
6. if $x \in V$ then $\|x\|^2 = \sum_i |\alpha_i|^2$

All of these provide different tests and definitions which identify a complete orthonormal set. I won't go through the proofs of their connections, but will point out that a couple of them replace $a \Rightarrow b$ w/ $\text{not}(b) \Rightarrow \text{not}(a)$. But this what you should expect. E.g. If $x = \text{blue} \Rightarrow x \in \text{color}$, but if $y \in \text{color} \not\Rightarrow y = \text{blue}$, rather $y \notin \text{color} \Rightarrow y \neq \text{blue}$

Some examples of failure:



$X = \{x_1, x_2\}$ in $\mathbb{R}^3$ is not a complete orthonormal set

(1) $\{x_1, x_2\} \in \{x_1, x_2, x_3 = y\}$

(2) $(y, x_1) = 0$, $(y, x_2) = 0$, but $y \neq 0$

(3) $y \neq \alpha_1 x_1 + \alpha_2 x_2$

(4) $y \neq (x_1, y) x_1 + (x_2, y) x_2 = 0$

(5) $(x, y) = (x, x_1)\cancelto{0}{(x_1, y)} + (x, x_2)\cancelto{0}{(x_2, y)} = 0$ but $(x, y) \neq 0$

(6) Remember Bessel?

Schwarz's inequality: For vectors x and y $\Rightarrow$ $|(x, y)| \leq \|x\|\,\|y\|$

needed since $(\ ,\ )$ can be negative

To prove this we just say if $y = 0$ both sides vanish, but if $y \neq 0$ we form $\frac{y}{\|y\|}$ which by itself is orthonormal, then use Bessel's inequality $\|x\| \geq \left|\left(\frac{y}{\|y\|}, x\right)\right|$ $\Rightarrow$ $\|x\|\,\|y\| \geq |(y, x)|$

no sum since $\frac{y}{\|y\|}$ is the only one in the set