Up to now we have primarily worked with the algebraic structure of vector spaces — there are three more structures that prove to be very useful:

- Inner products
- Norms
- Metrics

**Geometry**
- Distances, also angles
- A metric $d$

We will see that inner products induce a norm:

$\mathbf{H}: V \to \mathbb{R}$.

1. $\|v\| \geq 0$ and $\|v\| = 0$ iff $v = 0$
2. $\|av\| = |a| \|v\|$ for any scalar $a$
3. $\|v + w\| \leq \|v\| + \|w\|$
Def: \( \mathcal{V} \) is a vector space over \( \mathbb{R} \).

\[ \langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R} \]

is called an inner product if:

1. \( \text{symmetric} \Rightarrow \langle u, v \rangle = \langle v, u \rangle \) (symmetric)
   
   \( \Rightarrow \) \( F = 1 \) (different if \( F \neq 1 \))

2. \( \text{bilinear} \Rightarrow \langle u, v \rangle \) is \( \text{bilinear} \)
   
   \[ \langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle \]

3. \( \text{positivity} \Rightarrow \langle u, u \rangle \geq 0 \) \( \forall u \in \mathcal{V} \) \( \& \) \( \langle u, u \rangle = 0 \iff u = 0 \).

Some examples: \( \mathcal{V} = \mathbb{R}^n \)

\[ \langle x, y \rangle = x^T y \] (dot product)

\( \mathcal{V} : \text{cont. functions on } [0,1] \):

\[ \langle f, g \rangle = \int_0^1 f(t) g(t) \, dt \]

\[ \Rightarrow \int_0^1 f(t)^2 \, dt \geq \]

\[ \Rightarrow f(t) = 0 \text{ a.e.} \]
for complex-valued case:

\[ \langle \cdot, \cdot \rangle : V \times V \to \mathbb{C} \]

positivity is the same, but we require

sesquilinear

\[ \langle \lambda u, v \rangle = \lambda \langle u, v \rangle \]
\[ \langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle \]

Hermitian

\[ \langle u, v \rangle = \overline{\langle v, u \rangle} \]

complex conjugate.

If \( \lambda = \alpha + i\beta \)
\[ \overline{\lambda} = \alpha - i\beta \]

Claim: \( \sqrt{\langle x, x \rangle} \) is a norm for \( x \in V \).

That is, the inner product induces a norm on \( V \).
positivity: \( \sqrt{\langle x, x \rangle} \geq 0 \) \( \forall x \) & \( \sqrt{\langle x, x \rangle} = 0 \) if \( x = 0 \).

\[ \sqrt{\langle ax, ax \rangle} = \sqrt{a^2 \langle x, x \rangle} = |a| \sqrt{\langle x, x \rangle} \]

modulus of a complex \( \neq 0 \)

triangle inequality:

\[ ||x + y|| \leq ||x|| + ||y|| \]

\[ \frac{1}{2} \langle x+y, x+y \rangle \leq \langle x, x \rangle \frac{1}{2} + \langle y, y \rangle \]

or

\[ \frac{1}{2} \langle x+y, x+y \rangle \leq \langle x, x \rangle + 2 \langle x, x \rangle \langle y, y \rangle + \langle y, y \rangle \]

To show this we need a nice inequality known as

Cauchy - Schwanz: Suppose \( \langle x, x \rangle = 1 \cdot 1 \).

Then

\[ |\langle x, y \rangle| \leq ||x|| ||y|| \]

or

\[ |\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle \]
Proof: Let
\[
g(t) = \| x + ty \|^2
\]
\[
= \| x \|^2 + 2\langle x, y \rangle + t^2 \| y \|^2 \geq 0. \quad \forall t.
\]
\[
\Rightarrow \quad g(t) = \alpha t + \beta t + \gamma \geq 0 \quad \forall t.
\]
\[
\Rightarrow \quad \beta^2 - 4\alpha\gamma \leq 0 \quad \Rightarrow \quad \beta^2 \leq 4\alpha\gamma
\]
\[
\Rightarrow \quad 4\langle x, y \rangle^2 \leq 4\| x \|^2 \| y \|^2
\]
\[
\Rightarrow \quad |\langle x, y \rangle| \leq \| x \| \| y \|.
\]

Or
\[
|\langle x, y \rangle| \leq \langle x, x \rangle \langle y, y \rangle
\]
Let us now show that the norm using
\[ \|x\| = \sqrt{\langle x, x \rangle} \]
is in fact a norm: triangle inequality.

\[ \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \]

\[ \|x + y\|^2 = 1 \langle x, y \rangle \leq 2 \langle x, x \rangle \langle y, y \rangle + \langle y, y \rangle \]

\[ \leq \langle x, x \rangle + 2 \langle x, x \rangle \langle y, y \rangle + \langle y, y \rangle \]

\[ = (\|x\| + \|y\|)^2 \]

\[ \Rightarrow \|x + y\| \leq \|x\| + \|y\| \]
Notice that

\[ \|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2) \rightarrow \text{parallelogram law} \]

Two vectors \( u, v \) are called orthogonal if

\[ \langle u, v \rangle = 0 \rightarrow u \perp v \]

A set of vectors \( u_1, \ldots, u_n \) is orthonormal if

\[ \|u_i\| = 1 \quad \forall i \]

A group of vectors \( v_1, \ldots, v_n \) are orthonormal if

\[ v_i \perp v_j \quad \forall i, j, i \neq j \]

\[ \|v_i\|^2 = 1 \quad \forall i \]
So for the vector space of continuous functions on \([0, 1]\),
two functions are \( \perp \) (with respect to \( \langle f, g \rangle = \int_0^1 f(t)g(t) \, dt \))
if \( \int_0^1 f(t)g(t) \, dt = 0 \).

Note that if \( u_1 \perp u_2 \) then \( u_1 + u_2 \) are linearly indep.

\[ \| u_1 \| = 1 \]

Why?

\[ \langle \alpha_1 u_1 + \alpha_2 u_2, u_1 \rangle = \alpha_1 \| u_1 \|^2 = 0 \Rightarrow \alpha_1 = 0 \]
\[ \langle \alpha_1 u_1 + \alpha_2 u_2, u_2 \rangle = \alpha_2 \| u_2 \|^2 = 0 \Rightarrow \alpha_2 = 0 \]

and more generally for \( u_{11}, \ldots, u_m \).
Gram-Schmidt.

Any (set of) basis can be transformed to a set of orthonormal basis.

\[ e_1, e_2 \]

- orthornormal

\[ u_1, u_2 \]

- not orthonormal

? How can we start from \( u_1, u_2 \) as basis for \( \mathbb{R}^2 \) and end up with a orthonormal basis?
Suppose we start from $\mathbf{u}_1$:

$$\mathbf{\tilde{u}}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||} \rightarrow \text{normalization}.$$

Now consider

1. Span $\{\mathbf{\tilde{u}}_1\}$
2. $\text{span}\{\mathbf{\tilde{u}}_1\} \perp \text{span}\{\mathbf{\tilde{u}}_1\}$

Then

$$\mathbf{u}_2 = \alpha \mathbf{\tilde{u}}_1 + \mathbf{v},$$

where $\mathbf{v}$ is a linear combination of basis for $\text{span}\{\mathbf{\tilde{u}}_2\}$.

So

$$\langle \mathbf{\tilde{u}}_1, \mathbf{u}_2 \rangle \mathbf{\tilde{u}}_1$$

is the projection of $\mathbf{u}_2$ onto $\text{span}\{\mathbf{\tilde{u}}_2\}$.
What is \( \mathbf{u}_2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \mathbf{u}_1 \)?

Well we can check:

Could it be that

\[
\mathbf{u}_2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \mathbf{u}_1 \perp \mathbf{u}_1
\]

\[
\langle \mathbf{u}_1, \mathbf{u}_2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \mathbf{u}_1 \rangle = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle
\]

= 0 ! basis for \( \mathbb{R}^2 \)

So what we have done is: start from \( \mathbf{u}_1, \mathbf{u}_2 \)

\[
\mathbf{\tilde{u}}_1 = \frac{\mathbf{u}_1}{\| \mathbf{u}_1 \|}, \text{ first vector in the orthonormal}
\]

\[
\mathbf{\tilde{u}}_2 = \frac{\mathbf{u}_2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \mathbf{u}_1}{\| \mathbf{u}_2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \mathbf{u}_1 \|}, \text{ second orthonormal}
\]
In general if we have $u_1, \ldots, u_m$ as orthonormal set in a vector space of dimension $n$,

then we can create a new set member for this orthonormal set by normalizing

$$v = \sum_{i=1}^{m} \frac{\langle v, u_i \rangle}{\langle u_i, u_i \rangle} u_i$$

This is the basic idea behind Gram-Schmidt. Suppose we are given $v_1, \ldots, v_n$ a basis in $\mathbb{R}^n$ of dim $n$.

then the algorithm:

$$u_1 = \frac{v_1}{\|v_1\|}$$

$$u_2 = v_2 \text{ normalized } (v_2 - \langle v_2, u_1 \rangle u_1)$$

$$u_k = \text{ normalized } (v_k - \sum_{i=1}^{k-1} \langle v_k, u_i \rangle u_i)$$
We notice that the operation
\[ v \mapsto \sum_{j=1}^{m} \langle v, u_j \rangle u_j \]
is a projection of \( v \) onto \( \text{span} \{ u_1, \ldots, u_m \} \).

We can denote this projection by \( \Pi_{W}(v) \).

We have seen that
\[ v - \Pi_{W}(v) \in W^\perp \]
so any \( v \in V \) can write any \( v \in V \) given a subspace \( W \).