# Inner product spaces

July 18, 2017

Theory of vector spaces gives a generalization of some of the geometry we do in $\mathbb{R}^2$ and $\mathbb{R}^3$: adding vectors, multiplying them by a scalar, etc. But so far, in an abstract vector space we haven't been able to talk about the angle between two vectors, or about distances.

## Introduction - scalar product in $\mathbb{R}^2$

In $\mathbb{R}^2$, when are two vectors $x$ and $y$ orthogonal? When $x-y$ and $x+y$ have the same length (draw picture - see symmetry iff equality in lengths).

Now if $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ we get $x+y = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}$ and $x-y = \begin{bmatrix} x_1-y_1 \\ x_2-y_2 \end{bmatrix}$

$$\|x+y\|^2 = (x_1+y_1)^2 + (x_2+y_2)^2 = x_1^2 + 2x_1y_1 + y_1^2 + x_2^2 + 2x_2y_2 + y_2^2$$
$$\|x-y\|^2 = (x_1-y_1)^2 + (x_2-y_2)^2 = x_1^2 - 2x_1y_1 + y_1^2 + x_2^2 - 2x_2y_2 + y_2^2$$

thus we have equality iff $x_1y_1 + x_2y_2 = 0$.

**Definition 0.1:** *The **scalar product**, or **dot product**, or **inner product** on $\mathbb{R}^2$ is defined by*

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 + x_2y_2$$

Thus two vectors are orthogonal iff their scalar product is zero.

NB: The scalar product takes two **vectors**, and returns a **number**.

Note that scalar products also enable us to talk about the length of a vector, because the length of $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is $\|x\| = \sqrt{x_1^2 + x_2^2} = \sqrt{x \cdot x}$.

**Some properties of the scalar product:**

1. It is symmetric: $x \cdot y = y \cdot x$;
2. It is linear in each variable: $(x + x') \cdot y = x \cdot y + x' \cdot y$ and for any scalar $t \in \mathbb{R}$ we have $(tx) \cdot y = t(x \cdot y)$, and similarly $x \cdot (y + y') = x \cdot y + x \cdot y'$ and for any scalar $t \in \mathbb{R}$ we have $x \cdot (ty) = t(x \cdot y)$.

**Geometric interpretation of the scalar product in $\mathbb{R}^2$?** Project $u$ orthogonally onto the line spanned by $v$ - get a vector $u' = \alpha v$, such that $(u - u')$ is orthogonal to $v$. That is, we have

$$0 = (u - u') \cdot v = (u \cdot v) - \alpha(v \cdot v)$$

Hence $u \cdot v = \alpha\|v\|^2$. Now we can also notice that $\|u'\| = |\alpha|\|v\|$ so we get that $|u \cdot v|$ is the product of the lengths of $u'$ and of $v$, and its sign is positive if $u'$ is a positive multiple of $v$, and negative if $u'$ is a negative multiple of $v$.

Using trigonometry, we can also see that $\|u'\| = |\cos\alpha|\|u'\|$ where $\alpha$ is the angle from $v$ to to $u$. Thus $|u \cdot v| = \|u'\|\|v\| = |\cos\alpha|\|u\|\|v\|$. In particular, $|u \cdot v| \le \|u\|\|v\|$.

**Generalizing the scalar product in more dimensions** This definition can of course be extended to $\mathbb{R}^3$ (and then further to $\mathbb{R}^n$) by

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Again we have that two vectors are orthogonal in $\mathbb{R}^3$ iff their scalar product is zero, and that the length of $u$ is given by $\|u\| = \sqrt{u \cdot u}$.

**Generalizing the scalar product over $\mathbb{C}$** What could we give as a definition of the "dot product" over $\mathbb{C}^2$? If we set

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1 b_1 + a_2 b_2$$

then the dot product of a vector with itself is not a real number in general, for example

$$\begin{bmatrix} 1+i \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1+i \\ 1 \end{bmatrix} = 2i + 1$$

so it is hard to think about it as a length.

To get intuition, think of dimension 1: take $V = \mathbb{C}$. What is the "length" of $z \in \mathbb{C}$? It is its modulus $|z|$, computed by taking $|z|^2 = \bar{z}z$.

Thus the right thing to do over $\mathbb{C}^2$ is to set

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \overline{a_1} b_1 + \overline{a_2} b_2$$

But then note that the symmetry is broken: we have $a \cdot b = \overline{b \cdot a}$.

What we need on a general vector space to be able to talk about orthogonality and length is something similar to the scalar product. We will want to define this for vector spaces which are either over $\mathbb{R}$ or over $\mathbb{C}$.

# 1 Inner product

For the rest of this chapter, we will always have $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Let $V$ be a vector space over $\mathbb{F}$.

**Definition 1.1:** *An inner product on $V$ is a map $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{F}$ such that for any $u, u', v$ in $V$ and scalar $a \in \mathbb{F}$ we have*

1. *(linearity in the second variable) $\langle u|v + v'\rangle = \langle u|v\rangle + \langle u|v'\rangle$ (additivity) and $\langle u|av\rangle = a\langle u|v\rangle$;*
2. *(conjugate symmetry) $\langle v|u\rangle = \overline{\langle u|v\rangle}$ where the bar denotes complex conjugation;*
3. *(positive definiteness) $\langle u|u\rangle$ is real, non-negative and $\langle u|u\rangle = 0$ iff $u = 0$.*

Note that if $\mathbb{F} = \mathbb{R}$, the second condition reduces to $\langle v|u\rangle = \langle u|v\rangle$.

**Remark 1.2:** *Let $\langle\cdot|\cdot\rangle$ be an inner product on $V$. For any $u, u', v$ in $V$ and scalar $a \in \mathbb{F}$ we have*

1. *(sesquilinearity in first variable) $\langle u + au'|v\rangle = \overline{\langle v|u + au'\rangle} = \langle u|v\rangle + \overline{a}\langle u'|v\rangle$, in particular if $\mathbb{F} = \mathbb{R}$ the inner product is linear also in the first variable;*
2. *$\langle u|0\rangle = 0$ (indeed $\langle u|0\rangle = \langle u|0v\rangle = 0\langle u|v\rangle = 0$ by linearity in the second variable).*

**Example 1.3:** 1. $V = \mathbb{F}^n_{col}$ (with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ as usual); and if $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ we let

$$\langle x|y\rangle = \overline{x}^t y = \overline{x_1}y_1 + \ldots + \overline{x_n}y_n$$

Exercise: check that it has all the required properties. Note that in the case where $n = 2$ and $\mathbb{F} = \mathbb{R}$, this is exactly the scalar product that we defined in the introduction. This is called the standard inner product on $\mathbb{F}^n_{col}$.

2. $V = \mathbb{R}^2_{col}$; and $\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} | \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \rangle = (x_1 + x_2)(y_1 + y_2) + x_2y_2 = x_1y_1 + x_2y_1 + x_1y_2 + 2x_2y_2$. This is linear in $x$, symmetric, positive definite $\langle x|x\rangle = x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_2^2$. It is a **different** inner product than the scalar product on $\mathbb{R}^2_{col}$. On one and the same vector space, there are many different possible choices of inner products!

3. $V = C([0, 1])$ the space of all continuous functions $[0, 1] \to \mathbb{R}$. Define

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

This is an inner product on $V$.

4. $V = M_n(\mathbb{C})$ space of complex $n$-by-$n$ matrices. Define

$$\langle A|B\rangle = \mathrm{tr}(B^t\overline{A})$$

5. Building a new inner product out of an old one: suppose $\langle\cdot|\cdot\rangle$ is an inner product on $V$. Let $f : V \to V$ be an invertible operator on $V$. Then $\{\cdot|\cdot\} : V \times V \to \mathbb{F}$ given by $\{u|v\} = \langle f(u)|f(v)\rangle$ is also an inner product on $V$. Exercise: try to see why example 2. above is a special case of this.

**Definition 1.4:** *An **inner product space** is a real vector space $V$ over $\mathbb{R}$ or $\mathbb{C}$ together with an inner product $\langle\cdot|\cdot\rangle$.*

*A finite dimensional real inner product space is also called a **Euclidean space**, while a finite dimensional complex inner product space is called a **Hermitian space**.*

# 2 Norms

Let $(V, \langle\cdot|\cdot\rangle)$ be an inner product space.

**Definition 2.1:** *For any $v \in V$, we define the **norm** of $v$ by $\|v\| = \sqrt{\langle v|v\rangle}$.*

Note that this is well defined, and a real number, since $\langle v|v\rangle$ is real and nonnegative.

**Remark 2.2:** *For any $u, v \in V$ and $a \in \mathbb{F}$ we have*

1. *(positive definiteness) $\|u\| \geq 0$, and $\|u\| = 0$ iff $u = 0$;*
2. *(homogeneity) $\|av\| = \sqrt{\langle av|av\rangle} = \sqrt{\overline{a}a\langle v|v\rangle} = \sqrt{|a|^2 \langle v|v\rangle} = |a| \sqrt{\langle v|v\rangle} = |a|\|v\|$;*

Note that given any nonzero vector $v \in V$, the vector $\frac{v}{\|v\|}$ has norm 1 - we say it is a **unit vector**.

The following example highlights the fact that the norm of a vector depends on the inner product we have put on the space!

**Example 2.3:** $V = \mathbb{R}^2_{col}$ - consider the vector $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

1. if $\langle\cdot|\cdot\rangle$ is the standard inner product, given by $\langle \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} | \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} \rangle = x^1y^1 + x^2y^2$, then $\|e_2\| = 1$.
2. if $\langle \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} | \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} \rangle = (x^1 + x^2)(y^1 + y^2) + x^2y^2$, then $\|e_2\| = \sqrt{1+1} = \sqrt{2}$.

It is possible to recover the vector product from the norm, though the formula is different in the real and the complex cases:

**Proposition 2.4:** *(Polarization formula over $\mathbb{R}$) Let $(V, \langle\cdot|\cdot\rangle)$ be a real inner product space. For any $u, v \in V$ we have*

$$\langle u|v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$$

*Proof.* It suffices to compute

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v|u+v\rangle = \langle u|u\rangle + 2\langle u|v\rangle + \langle v|v\rangle = \|u\|^2 + \|v\|^2 + 2\langle u|v\rangle \\
\|u-v\|^2 &= \langle u-v|u-v\rangle = \langle u|u\rangle - 2\langle u|v\rangle + \langle v|v\rangle = \|u\|^2 + \|v\|^2 - 2\langle u|v\rangle
\end{aligned}$$

So we get $\|u+v\|^2 - \|u-v\|^2 = 4\langle u|v\rangle$. □

Over $\mathbb{C}$, we have

**Proposition 2.5:** *(Polarization formula over $\mathbb{C}$) Let $(V, \langle\cdot|\cdot\rangle)$ be a complex inner product space. For any $u, v \in V$ we have*

$$\langle u|v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2 - i\|u+iv\|^2 + i\|u-iv\|^2)$$

*Proof.* It suffices to compute

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v|u+v\rangle = \langle u|u\rangle + \langle u|v\rangle + \langle v|u\rangle + \langle v|v\rangle = \|u\|^2 + \|v\|^2 + \langle u|v\rangle + \overline{\langle u|v\rangle} \\
\|u-v\|^2 &= \langle u-v|u-v\rangle = \langle u|u\rangle - \langle u|v\rangle - \langle v|u\rangle + \langle v|v\rangle = \|u\|^2 + \|v\|^2 - \langle u|v\rangle - \overline{\langle u|v\rangle} \\
\|u+iv\|^2 &= \langle u+iv|u+iv\rangle = \langle u|u\rangle + \langle u|iv\rangle + \langle iv|u\rangle + \langle iv|iv\rangle = \|u\|^2 + \|v\|^2 + i\langle u|v\rangle - i\overline{\langle u|v\rangle} \\
\|u-iv\|^2 &= \langle u-iv|u-iv\rangle = \langle u|u\rangle - \langle u|iv\rangle - \langle iv|u\rangle + \langle iv|iv\rangle = \|u\|^2 + \|v\|^2 - i\langle u|v\rangle + i\overline{\langle u|v\rangle}
\end{aligned}$$

So we get $\|u+v\|^2 - \|u-v\|^2 - i\|u+iv\|^2 + i\|u-iv\|^2 = 4\langle u|v\rangle$. □

Recall that one of our goal in defining inner products was to generalize the notion of orthogonality. Let $(V, \langle\cdot|\cdot\rangle)$ be an inner product space.

**Definition 2.6:** *We say two vectors $u, v \in V$ are **orthogonal** if $\langle u|v\rangle = 0$. We then write $u \perp v$. Note that this is a symmetric relation.*

**Lemma 2.7:** *(Pythagoras) Let $u, v \in V$. If $u \perp v$ then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.*

*Proof.* We have $\|u+v\|^2 = \langle u+v|u+v\rangle = \langle u|u\rangle + \langle u|v\rangle + \langle v|u\rangle + \langle v|v\rangle = \|u\|^2 + 2\Re(\langle u|v\rangle) + \|v\|^2$. Thus $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ iff $\Re(\langle u|v\rangle) = 0$. □

Note that over $\mathbb{R}$, $\Re\langle u|v\rangle = 0$ iff $\langle u|v\rangle = 0$, hence the converse to Pythagoras also holds:

**Lemma 2.8:** *Let $u, v$ be vectors in a real inner product space $(V, \langle\cdot|\cdot\rangle)$. Then $u \perp v$ iff $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.*

This is false on $\mathbb{C}$: suppose $u = \begin{bmatrix} i \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then $\|u+v\|^2 = 2 = \|u\|^2 + \|v\|^2$ but $\langle u|v\rangle = i$.

Recall that for the dot product, we had that $|u \cdot v| = \|u\|\|v\||\cos(\text{ angle}(u,v))|$, in particular $|\langle u|v\rangle| \leq \|u\|\|v\|$. This remains true for a general inner product:

**Proposition 2.9:** *(Cauchy-Schwarz) For any $u, v \in V$ we have*

$$\left|\langle u|v\rangle\right| \leq \|u\|\|v\|$$

*with equality iff $u$ and $v$ are linearly dependent.*

*Proof.* If one of $u, v$ is zero this is obvious, so we may assume they are not. There exists $\alpha \in \mathbb{F}$ such that $\langle v|u - \alpha v\rangle = 0$, indeed, $\langle v|u - \alpha v\rangle = \langle v|u\rangle - \alpha\langle v|v\rangle$ so it is enough to take $\alpha = \frac{\langle v|u\rangle}{\langle v|v\rangle} = \frac{\langle v|u\rangle}{\|v\|^2}$.

Now by Pythagoras, we have $\|u\|^2 = \|\alpha v\|^2 + \|u - \alpha v\|^2$ thus we get

$$\|u\|^2 \geq \|\alpha v\|^2 = |\alpha|^2\|v\|^2 = \frac{\left|\langle v|u\rangle\right|^2}{\|v\|^4}\|v\|^2 = \frac{\left|\langle v|u\rangle\right|^2}{\|v\|^2}$$

Thus $\left|\langle v|u\rangle\right| \leq \|u\|\|v\|$.

To have equality we must have $\|u - \alpha v\|^2 = 0$, that is, $u = \alpha v$. On the other hand, if $v = au$ for some $a \in \mathbb{F}$ we get $\left|\langle u|v\rangle\right| = \left|\langle u|au\rangle\right| = |a|\left|\langle u|u\rangle\right| = |a|\|u\|^2 = \|u\|\|v\|$. ☐

**Example 2.10:** If $f, g$ are two continuous functions $[0,1] \to \mathbb{R}$ we have by Cauchy-Schwarz

$$\left|\int_0^1 f(t)g(t)\mathrm{d}t\right|^2 \leq \int_0^1 f^2(t)\mathrm{d}t \int_0^1 g^2(t)\mathrm{d}t$$

The following is another important property of the norm.

**Proposition 2.11:** *(triangle inequality) For any $u, v \in V$ we have $\|u + v\| \leq \|u\| + \|v\|$ with equality iff either vector is zero or one vector is a positive real multiple of the other.*

*Proof.* First note that if $z = x + iy \in \mathbb{C}$, we have $z + \bar{z} = (x + iy) + (x - iy) = 2x = 2\mathrm{Re}(z)$ and $|z| = \sqrt{x^2 + y^2} \geq x$ with equality iff $y = 0$ and $x \geq 0$. Thus we get $z + \bar{z} \leq |z|$ with equality iff $z$ is real and positive.

We have

$$\begin{aligned}
\|u + v\|^2 &= \|u\|^2 + \langle u|v\rangle + \overline{\langle u|v\rangle} + \|v\|^2 = \|u\|^2 + 2\mathrm{Re}(\langle u|v\rangle) + \|v\|^2 \\
&\leq \|u\|^2 + 2\left|\langle u|v\rangle\right| + \|v\|^2 \text{ with equality iff } \langle u|v\rangle \text{ is real and positive} \\
&\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \text{ by Cauchy-Schwarz - thus with equality iff u,v linearly dependent} \\
&= (\|u\| + \|v\|)^2.
\end{aligned}$$

We have equality overall iff the inequalities on both lines are equalities, that is, iff $\langle u|v\rangle$ is real and positive and $u, v$ are linearly dependent. But $u, v$ are linearly dependent $\iff$ there exists $a \in \mathbb{F}$ such that $v = au$; and in this case $\langle u|v\rangle = a\langle u|u\rangle = a\|u\|^2$ thus $\langle u|v\rangle$ is real and positive iff $a$ is. ☐

Now that we have a way to define the "length" of a vector, we can use it to define a distance function on $V$.

**Definition 2.12:** *The* ***distance function*** *on $V$ is given by*

$$\begin{aligned}
d : &V \times V \to \mathbb{R}^+ \\
&(v, w) \mapsto \|w - v\|
\end{aligned}$$

It satisfies $d(v, w) = 0$ iff $v = w$, it is symmetric $d(v, w) = d(w, v)$, and it satisfies the triangle inequality $d(u, w) \leq d(u, v) + d(v, w)$ - these are all properties one expects a "distance" to have.

# 3 Orthogonality

Let $(V, \langle\cdot|\cdot\rangle)$ be an inner product space.

**Definition 3.1:** *If $S, T \subseteq V$ we write*

- *$u \perp S$ if for any $v \in S$, we have $u \perp v$;*
- *$S \perp T$ if for any $u \in S$, $v \in T$, we have $u \perp v$;*
- *$S^\perp = \{v \in V \mid v \perp S\}$*

The following lemma shows how orthogonality interacts with the vector space structure

**Lemma 3.2:** *If $S \subseteq V$ we have*

1. *$v \perp S$ iff $v \perp \mathrm{Span}S$;*
2. *$S^\perp$ is a vector subspace of $V$;*
3. *if $S \subseteq T$ then $T^\perp \subseteq S^\perp$.*

*Proof.* Exercise. □

## 3.1 Orthogonal projection of a vector

Because inner products induce a distance, we can now ask the following question: given $W$ a vector subspace of an inner product space $(V, \langle\cdot|\cdot\rangle)$, and $v \in V$, what is the point of $W$ closest to $v$?

Consider what happens in $\mathbb{R}^2$: the point of a line $l$ closest to $v$ is the vector $v'$ in $l$ such that $v - v' \perp l$. This motivates the following definition

**Definition 3.3:** *Let $W$ be a subspace of $V$. Let $v$ be a vector. We say that $v_W$ is an **orthogonal projection** of $v$ on $W$ if it satisfies*

*(i)* $v_W \in W$;

*(ii)* $(v - v_W) \in W^\perp$.

In other words, $v_W$ is an orthogonal projection of $v$ on $W$ if $v = v_W + v'$ with $v' \perp W$.

Let us see that this really solves our minimization problem

**Proposition 3.4:** *Let $W$ be a finite dimensional subspace of an inner product space $(V, \langle\cdot|\cdot\rangle)$. Let $v \in V$, and let $v_W$ be an orthogonal projection of $v$ on $W$. Then for any $w \in W$ we have*

$$\|v - v_W\| \leq \|v - w\|$$

*with equality iff $w = v_W$. In particular, $v$ admits at most one orthogonal projection onto $W$.*

*Proof.* Let $w \in W$. We have $(v - v_W) \perp W$, and $v_W - w \in W$ so $(v - v_W) \perp (v_W - w)$. By Pythagoras

$$\|v - v_W\|^2 + \|v_W - w\|^2 = \left\|(v - v_W) + (v_W - w)\right\|^2 = \|v - w\|^2$$

Thus $\|v - v_W\|^2 \leq \|v - w\|^2$, with equality iff $\|v_W - w\|^2 = 0$, that is iff $w = v_W$. Suppose $v'_W$ is another orthogonal projection of $v$ onto $W$. Since $v'_W \in W$, we have that $\|v - v_W\|^2 \leq \left\|v - v'_W\right\|^2$, with equality iff $v'_W = v_W$. But by symmetry we can swap the roles of $v_W$ and $v'W$, hence we also have that $\left\|v - v'_W\right\|^2 \leq \|v - v_W\|^2$. Therefore $v_W = v'_W$. □

Careful! We haven't proved that there always is such a vector...in fact if $W$ is infinite dimensional, sometimes the infemum of $d(v, w)$ over all $w \in W$ is not reached. But we will prove later that if $W$ is finite dimensional, the orthogonal projection always exists and is moreover unique.

**Remark 3.5:** *Note that if $v_W$ is an orthogonal projection of $v$ on $W$, then $u = (v - v_W) \perp W$ so (i) $u = (v - v_W) \in W^\perp$, and $v - u = v - (v - v_W) = v_W \in W$ so (ii) $v - u \perp W^\perp$. In other words, $u = v - v_W$ is an orthogonal projection of $v$ on $W^\perp$.*

## 3.2 Orthogonal/orthonormal families of vectors

**Definition 3.6:** *Let $v_1, \ldots, v_k$ be a family of vectors. We say that it is an* ***orthogonal*** *family if for any $i \neq j$ we have $\langle v_i | v_j \rangle = 0$.*

*We say that it is an* ***orthonormal*** *family if*

- *for any $i \neq j$ we have $\langle v_i | v_j \rangle = 0$;*
- *for any $i$ we have $\|v_i\| = 1$ (i.e. the $v_i$'s are all unit vectors).*

We can use the Kronecker delta notation $\delta_{ij}$ where $\delta_{ij}$ takes the value 1 if $i = j$ and 0 if $i \neq j$.

Then $v_1, \ldots, v_k$ form an orthonormal family of vectors iff $\langle v_i | v_j \rangle = \delta_{ij}$.

**Proposition 3.7:** *Let $(v_1, \ldots, v_m)$ be an orthogonal family of nonzero vectors. Then it is linearly independent.*

*Proof.* Suppose there exist scalars $a_1, \ldots, a_m \in \mathbb{F}$ such that $a_1 v_1 + \ldots + a_m v_m = 0$. Fix $i$: we have

$$\begin{aligned} \langle v_i | a_1 v_1 + \ldots + a_m v_m \rangle &= a_1 \langle v_i | v_1 \rangle + \ldots + a_i \langle v_i | v_i \rangle + \ldots + a_m \langle v_i | v_m \rangle \\ &= a_i \|v_i\|^2 . \end{aligned}$$

On the other hand, $\langle v_i | a_1 v_1 + \ldots + a_m v_m \rangle = \langle v_i | 0 \rangle = 0$, and $\|v_i\|^2 \neq 0$, hence $a_i = 0$. This holds for every $i$, hence the vectors $v_1, \ldots, v_m$ are linearly independent. ☐

If an orthonormal family spans a subspace $W$, then orthogonal projections on $W$ exist.

**Proposition 3.8:** *Let $W$ be a finite dimensional vector subspace of $V$. Assume that $w_1, \ldots, w_k$ is an orthonormal basis for $W$. For any $v \in V$, let $v_W = \sum_{i=1}^{k} \langle w_i | v \rangle w_i$.*

*Then $v_W$ is the unique orthogonal projection of $v$ on $W$.*

*Proof.* Clearly we have that $(i)$ $v_W \in W$. Let us see that $v' = v - v_W \perp W$: by Lemma 3.2, it is enough to show that $\langle w_j | v' \rangle = 0$ for all $j$. But

$$\langle w_i | v' \rangle = \langle w_j |\ v - \sum_{i=1}^{k} \langle v | w_i \rangle w_i\ \rangle = \langle w_j | v \rangle - \sum_{i=1}^{k} \langle w_i | v \rangle \langle w_j | w_i \rangle = \langle w_j | v \rangle - \langle w_j | v \rangle = 0$$

Thus $v_W$ is an orthogonal projection of $v$ on $W$.

Suppose now $u$ is an orthogonal projection of $v$ on $W$ - then $u \in W$ so $u = a_1 w_1 + \ldots + a_k w_k$. Since $v - u \perp W$ we must have for each $i \leq k$

$$\begin{aligned} 0 = \langle w_i | v - u \rangle = \langle w_i | v \rangle - \langle w_i | u \rangle &= \langle w_i | v \rangle - \langle w_i | a_1 w_1 + \ldots + a_k w_k \rangle \\ &= \langle w_i | v \rangle - (a_1 \langle w_i | w_1 \rangle + \ldots + a_k \langle w_i | w_k \rangle) \\ &= \langle w_i | v \rangle - a_i \langle w_i | w_i \rangle = \langle w_i | v \rangle - a_i . \end{aligned}$$

Hence $u = \sum_{i=1}^{k} \langle w_i | v \rangle w_i = v_W$ - the orthogonal projection is unique. ☐

## 3.3 Gram-Schmidt orthonormalization procedure

We will now see that we can always find an orthonormal family of vectors which spans the space.

The following gives a way to construct an orthonormal family of vectors starting from any family:

**Proposition 3.9:** *(Gram-Schmidt orthonormalization procedure) Let $(V, \langle \cdot | \cdot \rangle)$ be an inner product space. Let $(b_1, \ldots, b_m)$ be an ordered linearly independent family of vectors. Then there exists an ordered orthonormal family $(u_1, \ldots, u_m)$ such that for each $k \leq m$ we have*

$$\mathrm{Span}(b_1, \ldots, b_k) = \mathrm{Span}(u_1, \ldots, u_k)$$

*Proof.* We construct the vectors $u_j$ by induction. We set $u_1 = b_1/\|b_1\|$: the family $(u_1)$ is an orthonormal family.

Suppose that we have constructed an orthonormal family $(u_1, \ldots, u_j)$ such that $\mathrm{Span}(u_1, \ldots, u_k) = \mathrm{Span}(b_1, \ldots, b_k)$ for all $k \leq j$. We let $b'_{j+1}$ be the orthogonal projection of $b_{j+1}$ on $\mathrm{Span}(u_1, \ldots, u_j)$ - it exists and is unique by Lemma 3.8 (and in fact it is given by $b'_{j+1} = \sum_{i=1}^{j} \langle b_{j+1} | b_i \rangle b_i$).

By definition of orthogonal projection, $(b_{j+1} - b'_{j+1}) \perp \mathrm{Span}(u_1, \ldots, u_j)$, so for each $i \leq j$ we have $\langle b_{j+1} - b'_{j+1} | u_i \rangle = 0$.

We set $u_{j+1} = (b_{j+1} - b'_{j+1})/\left\| b_{j+1} - b'_{j+1} \right\|$. The vector $u_{j+1}$ is a rescaling of $b_{j+1} - b'_{j+1}$, so it is also orthogonal to all the vectors $u_1, \ldots, u_j$. Moreover, it is a unit vector. Thus $(u_1, \ldots, u_{j+1})$ is an orthonormal family.

Now $u_{j+1} \in \mathrm{Span}(u_1, \ldots, u_j, b_{j+1}) = \mathrm{Span}(b_1, \ldots, b_{j+1})$ so $\mathrm{Span}(u_1, \ldots, u_{j+1}) \subseteq \mathrm{Span}(b_1, \ldots, b_{j+1})$. Since the $u_i$ are linearly independent, we have in fact equality. □

**Corollary 3.10:** *If $(V, \langle \cdot | \cdot \rangle)$ is a finite dimensional inner product space, then it admits an orthonormal basis.*

## 3.4 Orthonormal bases

Suppose now that $(V, \langle \cdot | \cdot \rangle)$ is a finite dimensional inner product space. Let $(u_1, \ldots, u_n)$ be an orthonormal basis for $V$.

The following proposition shows that we then have a nice way of expressing the coordinates of a vector $v$ in this basis.

**Proposition 3.11:** *Let $(V, \langle \cdot | \cdot \rangle)$ be a Euclidean or a Hermitian space. Let $(u_1, \ldots, u_n)$ be an ordered orthonormal basis. For any vector $v \in V$ we have $v = \sum_{i=1}^{n} \langle u_i | v \rangle u_i$ - in other words*

$$[v]_{\mathcal{B}} = \begin{bmatrix} \langle u_1 | v \rangle \\ \vdots \\ \langle u_n | v \rangle \end{bmatrix}$$

*Proof.* Suppose $v = \sum_{i=1}^{n} a_i u_i$. Then for any $j$ we have

$$\langle u_j | v \rangle = \langle u_j | \sum_{i=1}^{n} a_i u_i \rangle = \sum_{i=1}^{n} \langle u_j | a_i u_i \rangle = \sum_{i=1}^{n} a_i \langle u_j | u_i \rangle = a_j \langle u_j | u_j \rangle = a_j.$$

□

From this we get the following expression for the inner product of two vectors.

**Corollary 3.12:** *Let $(V, \langle \cdot | \cdot \rangle)$ be a Euclidean or a Hermitian space. Let $(u_1, \ldots, u_n)$ be an ordered orthonormal basis. For any vectors $v, w \in V$ we have*

1. *(Parseval)* $\langle v | w \rangle = \sum_{i=1}^{n} \overline{\langle u_i | v \rangle} \langle u_i | w \rangle = \overline{[v]}_{\mathcal{B}}^{t} [w]_{\mathcal{B}}$;
2. *(Bessel)* $\|v\|^2 = \sum_{i=1}^{n} \left| \langle u_i | v \rangle \right|^2$.

**Remark 3.13:** *What this proposition shows is that once we have an orthonormal basis $\mathcal{B}$ on a finite dimensional inner product space $(V, \langle | \rangle)$, and we identify each vector to the column vector of its coordinates in $\mathcal{B}$, the inner product can be thought of as the standard inner product on $\mathbb{F}^n_{col}$.*

*Proof.* 1. Suppose $v = \sum_{i=1}^{n} a_i u_i$ and $w = \sum_{i=1}^{n} b_i u_i$. We get

$$\begin{aligned} \langle v | w \rangle &= \langle \sum_{i=1}^{n} a_i u_i | \sum_{j=1}^{n} b_j u_j \rangle = \sum_{i=1}^{n} \overline{a_i} \langle u_i | \sum_{j=1}^{n} b_j u_j \rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{a_i} b_j \langle u_i | u_j \rangle = \sum_{i=1}^{n} \overline{a_i} b_i \text{ since } \langle u_i | u_j \rangle \text{ is 0 if } i \neq j \text{ and 1 if } i = j \end{aligned}$$

But by Proposition 3.11 above, we have $v = \sum_{i=1}^{n} \langle u_i | v \rangle u_i$ and $w = \sum_{j=1}^{n} \langle u_j | w \rangle u_j$, hence $a_i = \langle u_i | v \rangle$ and $b_i = \langle u_i | w \rangle$.

2. Using Parseval we get $\|v\|^2 = \langle v | v \rangle = \sum_{i=1}^{n} \overline{\langle u_i | v \rangle} \langle u_i | v \rangle = \sum_{i=1}^{n} \left| \langle u_i | v \rangle \right|^2$.

□

Having an orthonormal basis also helps understand the perp of a subspace generated by the first $k$ elements of the basis.

**Corollary 3.14:** *Let $(V, \langle \cdot | \cdot \rangle)$ be a Euclidean or a Hermitian space. Let $(u_1, \ldots, u_n)$ be an ordered orthonormal basis. Let $W = \operatorname{Span}(u_1, \ldots, u_k)$. Then $W^\perp = \operatorname{Span}(u_{k+1}, \ldots, u_n)$.*

*Proof.* If we fix $i > k$, then for each $j \leq k$, we have $u_j \perp u_i$, so $u_i \subseteq \{u_1, \ldots, u_k\}^\perp = \operatorname{Span}(u_1, \ldots, u_k)^\perp = W^\perp$ (here we used Lemma 3.2). On the other hand, suppose $v \in W^\perp$. By Proposition 3.11, $v = \sum_{i=1}^{n} \langle u_i | v \rangle u_i$, but for each $i \leq k$ we have $\langle u_i | v \rangle = 0$, hence $v = \sum_{i=k+1}^{n} \langle u_i | v \rangle u_i \in \operatorname{Span}(u_{k+1}, \ldots, u_n)$. □

## 3.5 Orthogonal complements and projections

Note that Gram-Schmidt gives us more than just the existence of an orthonormal basis for the whole space. Indeed, it means that from any basis we can build an orthonormal one, with the first $k$ vectors of the new basis spanning the same vector subspace as the first $k$ vectors of the old one. Thus in particular, if we are given a vector subspace $W$ of $V$, we can choose a basis for $W$, extend it to a basis for $V$ and then apply Gram-Schmidt procedure. This gives:

**Corollary 3.15:** *If $W$ is a vector subspace of a finite dimensional inner product space $V$, there exists an orthonormal basis $(u_1, \ldots, u_n)$ of $V$ such that $W = \operatorname{Span}(u_1, \ldots, u_k)$.*

Corollary 3.14 now tells us that $W^\perp = \operatorname{Span}(u_{k+1}, \ldots, u_n)$ so we deduce immidiately that

**Corollary 3.16:** *If $W$ is a vector subspace of a finite dimensional inner product space $V$, $\dim W + \dim W^\perp = \dim V$.*

Now recall that if the dimensuion of the span of two subspace is equal to the sum of their dimension, they are in direct sum (see Proposition 4.8 in the "Operators" notes)- thus we get

**Corollary 3.17:** *If $W$ is a vector subspace of a finite dimensional inner product space $V$, then $V = W \oplus W^\perp$*

In other words, the spaces $W$ and $W^\perp$ are complementary.

**Definition 3.18:** *The space $W^\perp$ is called the* ***orthogonal complement*** *of $W$.*

Note that in general, a subspace $W$ has many complements - but it has only one orthogonal complement.

Recall that when we write the space $V$ as a direct sum of two subspace $U_1 \oplus U_2$, we can write any vector $v \in V$ uniquely as $v = u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$. Then the projection $p_{U_1,U_2}$ on $U_1$ parallel to $U_2$ is defined to be the linear map which sends $v$ to $u_1$. Similarly, the projection $p_{U_2,U_1}$ on $U_2$ parallel to $U_1$ is defined to be the linear map which sends $v$ to $u_2$.

**Remark 3.19:** *Here, we write $V$ as $W \oplus W^\perp$, so eachy vector $v$ can be written uniquely as $v = v_1 + v_2$ with $v_1 \in W, v_2 \in W^\perp$. We have $(i) v_1 \in w$ and $(ii) v_2 = v - v_1 \perp W$ so $v_1$ is exactly the orthogonal projection of $v$ on $W$! Thus in the particular case where $U_1, U_2$ are of the form $W, W^\perp$, the projection on $W$ parallel to $W^\perp$ sends a vector $v$ to its orthogonal projection to $W$.*

**Definition 3.20:** *We call the map $p_{W,W^\perp}$ the orthogonal projection onto $W$. It is a linear map $V \to V$ whose image lies in $W$.*

(before we had defined the orthogonal projection *of a given vector*)

**Corollary 3.21:** *Let $W$ be a vector subspace of a finite dimensional inner product space $(V, \langle \cdot | \cdot \rangle)$. Then $(W^\perp)^\perp = W$.*

*Proof.* Let $w \in W$. Then $w \perp W^\perp$ (indeed, for any $u \in W^\perp$, we have $u \perp w$ by definition of $W^\perp$) . Hence $w \in (W^\perp)^\perp$, so we have shown that $W \subseteq (W^\perp)^\perp$.

But $\dim(W^\perp)^\perp = \dim V - \dim W^\perp = \dim V - (\dim V - \dim W) = \dim W$ so we have equality. $\square$

# 4 Orthogonal and unitary operators

Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Let $(V, \langle\cdot|\cdot\rangle)$ and $(W, \{\cdot \mid \cdot\})$ be inner product spaces over $\mathbb{F}$.

**Definition 4.1:** *An isometry between $(V, \langle\cdot|\cdot\rangle)$ and $(W, \{\cdot \mid \cdot\})$ is a map $f : V \to W$ which preserves the inner product, that is, which is such that for any $u, v \in V$ we have $\{f(u) \mid f(v)\} = \langle u|v\rangle$.*

**Remark 4.2:** *Note also that $f$ must preserve norms, i.e., for all $v \in V$ we have $\|f(v)\|_W = \|v\|_V$. Indeed, $\|f(v)\|^2 = \{f(v) \mid f(v)\} = \langle v|v\rangle = \|v\|^2$.*

In an inner product space we think of $\|v - w\|$ as the distance between the vectors $v, w$, hence isometries are maps which preserve the distance.

**Example 4.3:** If $(V, \langle\cdot|\cdot\rangle)$ is an inner product space of dimension $n$, there exists an isometry $f : V \to \mathbb{F}^n$ between $V$ and $\mathbb{F}^n$ endowed with the scalar product. Indeed, pick an orthonormal basis $\mathcal{B}$ for $V$, and let $f(v) = [v]_\mathcal{B}$. If $u, v \in V$ we have

$$f(u) \cdot f(v) = [u]_\mathcal{B} \cdot [v]_\mathcal{B} = \overline{[u]_\mathcal{B}}^t [v]_\mathcal{B} = \langle u|v\rangle$$

where the last equality comes from Parseval equality.

**Remark 4.4:** *An isometry is linear. Indeed, for any $u, v \in V$ we have*

$$\begin{aligned}
\{f(\lambda u) - \lambda f(u) \mid f(\lambda u) - \lambda f(u)\} &= \{f(\lambda u) \mid f(\lambda u)\} - \bar\lambda\{f(u) \mid f(\lambda u)\} - \lambda\{f(\lambda u) \mid f(u)\} + \lambda\bar\lambda\{f(u) \mid f(u)\} \\
&= \langle \lambda u|\lambda u\rangle - \bar\lambda\langle u|\lambda u\rangle - \lambda\langle \lambda u|u\rangle + \lambda\bar\lambda\langle u|u\rangle \\
&= \lambda\bar\lambda\langle u|u\rangle - \bar\lambda\lambda\langle u|u\rangle - \lambda\bar\lambda\langle u|u\rangle + \lambda\bar\lambda\langle u|u\rangle \\
&= 0
\end{aligned}$$

*Hence $f(\lambda u) = \lambda f(u)$. Similarly, one can check that*

$$\{f(u+v) - f(u) - f(v) \mid f(u+v) - f(u) - f(v)\} = 0$$

*to conclude.*

We will be especially interested in isometries from a space to itself, that is, operators which are also isometries.

**Definition 4.5:** *Let $(V, \langle\cdot|\cdot\rangle)$ be an inner product space. Let $f : V \to V$ be an isometry, that is, for any $v, w \in V$ we have*

$$\langle f(v)|f(w)\rangle = \langle v|w\rangle$$

*In the case $\mathbb{F} = \mathbb{R}$ we say that $f$ is* ***orthogonal***. *In the case $\mathbb{F} = \mathbb{C}$ we say $f$ is* ***unitary***.

Sometimes we also say "real unitary" instead of "orthogonal".

**Remark 4.6:** *As the name suggests in the real case, the operator $f$ preserves orthogonality, that is, if $v \perp w$ then $f(v) \perp f(w)$. But not all operators preserving orthogonality are orthogonal!*

**Example 4.7:** Let $f : V \to V$ be defined by $v \mapsto 2v$: if $\langle v|w\rangle = 0$ then $\langle 2v|2w\rangle = 4\langle v|w\rangle = 0$. However if $v \neq 0$, we have $\langle f(v)|f(v)\rangle = \langle 2v|2v\rangle = 4\langle v|v\rangle \neq \langle v|v\rangle$ so $f$ is not orthogonal.

**Example 4.8:** Let $V = \mathbb{R}^2_{col}$ be endowed with the standard inner product. Let $f : V \to V; v \mapsto A_f v$ be orthogonal. It must send the standard basis $(e_1, e_2)$ to vectors whose norm is also 1, and $f(e_1) \perp f(e_2)$. It must be either a rotation, or a rotation composed with a reflection.

We now give another characterization of orthogonal/unitary operators. First, a definition:

**Definition 4.9:** *Let $(V, \langle\cdot|\cdot\rangle)$ be an inner product space. A unit vector is a vector $u$ of norm 1, that is $u \in V$ such that $\|u\| = 1$.*

Note that for any $v \in V - \{0\}$, the vector $u = \frac{v}{\|v\|}$ is a unit vector which lies in $\mathrm{Span}(v)$. Indeed $u = \alpha v$ with $\alpha = 1/\|v\| \in \mathbb{R}^+$, so $\|u\| = \|\alpha v\| = |\alpha|\|v\| = (1/\|v\|)\|v\| = 1$.

**Proposition 4.10:** *Let $f : V \to V$. The following are equivalent*

1. *$f$ is orthogonal/unitary;*
2. *$f$ preserves the norm, i.e. for any $v \in V$ we have $\|f(v)\| = \|v\|$;*
3. *$f$ sends every unit vector to a unit vector.*

The third characterization explains the use of the term "unitary" for orthogonal operators.

*Proof.* $(1 \to 2)$ See remark above.

$(2 \to 1)$ Suppose first that $\mathbb{F} = \mathbb{R}$. In this case the polarization formula gives us $\langle u|v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$, therefore we get that for any $u, v \in V$

$$\begin{aligned}\langle f(u)|f(v)\rangle &= \frac{1}{4}(\|f(u)+f(v)\|^2 - \|f(u)-f(v)\|^2) \\ &= \frac{1}{4}(\|f(u+v)\|^2 - \|f(u-v)\|^2) \text{ since } f \text{ is linear.} \\ &= \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2) \text{ since } f \text{ preserves norms.} \\ &= \langle u|v\rangle\end{aligned}$$

What matters here is not the precise formula, but the fact that the inner product can be defined purely in terms of the norm. Therefore, if $\mathbb{F} = \mathbb{C}$ we can use the complex polarization formula to deduce the result.

$(2 \to 3)$ Obvious.

$(3 \to 2)$ Let $v \in V$. If $v = 0$, then $f(v) = 0$ so $\|v\| = \|f(v)\| = 0$. If $v \neq 0$, the vector $v' = \frac{1}{\|v\|}v$ is a unit vector, so $\|f(v')\| = \|v'\| = \left\|\frac{1}{\|v\|}v\right\| = \frac{1}{\|v\|}\|v\|$. On the other hand, $\|f(v')\| = \left\|f(\frac{1}{\|v\|}v)\right\| = \frac{1}{\|v\|}\|f(v)\|$. Thus we get $\frac{1}{\|v\|}\|f(v)\| = \frac{1}{\|v\|}\|v\|$, so $f$ also preserves the norm of $v$. □

**Corollary 4.11:** *An orthogonal/unitary operator $f$ on a finite dimensional inner product space is invertible.*

*Proof.* It is enough to show that $f$ is injective. Suppose $f(v) = 0$ - since $f$ preserves the norm we have $\|v\| = \|f(v)\| = 0$ so $v = 0$. □

**Remark 4.12:** *Let $f$ be an operator on a finite dimensional inner product space. If $f$ is orthogonal, then $f^{-1}$ is also orthogonal. Indeed, for any $v \in V$, $\|f^{-1}(v)\| = \|f(f^{-1}(v))\| = \|v\|$.*

**Remark 4.13:** *What eigenvalues can an orthogonal/unitary operator have? If $f(v) = \lambda v$, we must have $\|v\| = \|f(v)\| = |\lambda|\|v\|$, hence $|\lambda| = 1$.*

*If $V$ is a vector space over the reals this means $\lambda = \pm 1$. If $V$ is a complex vector space, $\lambda$ can be any complex number of modulus 1.*

**Example 4.14:** Let $f$ be an orthogonal operator on $\mathbb{R}^3$. Its characteristic polynomial is a real valued polynomial of degree 3, hence it must have a root in $\mathbb{R}$ (think of its decomposition into irreducible polynomials over $\mathbb{R}$). By the remark above, this root can only be $+1$ or $-1$.

Hence there is a line $l$ which is either fixed or reversed. Denote by $W$ the plane orthogonal to $l$. Because $f$ preserves orthogonality, it preserves $W$. We can think of $f|_W$ as an orthogonal operator on $\mathbb{R}^2$ - it is a rotation or a rotation + reflection.

**Proposition 4.15:** *Let $f : V \to V$ be an operator on a finite-dimensional inner product space $V$. The following conditions are equivalent:*

1. *$f$ is orthogonal/unitary;*
2. *$f$ sends any orthonormal basis to an orthonormal basis;*
3. *$f$ sends some orthonormal basis to an orthonormal basis.*

Recall that $\mathcal{B} = (b_1, \dots, b_n)$ is an orthonormal basis $\iff$ $\langle b_i | b_j \rangle = \delta_{ij}$.

Recall that we saw in Remark 3.13 that for any vectors $v, w \in V$ with $v = \sum_{i=1}^n v_i b_i$ and $w = \sum_{j=1}^n w_j b_j$, or equivalently $[v]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $[w]_{\mathcal{B}} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, we have

$$\langle v | w \rangle = \sum_{i=1}^n \overline{v}_i w_i = \begin{bmatrix} \overline{v}_1 \dots \overline{v}_n \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \overline{[v]}_{\mathcal{B}}^t [w]_{\mathcal{B}}$$

*Proof.*

(1 $\to$ 2) If $f$ is orthogonal/unitary we get $\langle f(b_i) | f(b_j) \rangle = \langle b_i | b_j \rangle = \delta_{ij}$, hence $(f(b_1), \dots, f(b_n))$ is an orthonormal basis.

(2 $\to$ 3) Obvious.

(3 $\to$ 1) Suppose that the image $f(b_1), \dots, f(b_n))$ of the orthonormal basis $(b_1, \dots, b_n)$ is orthonormal. Let $v, w \in V$, suppose $v = \sum_{i=1}^n v_i b_i$ and $w = \sum_{j=1}^n w_j b_j$. We have

$$\begin{aligned} \langle f(v) | f(w) \rangle &= \langle f(\sum_{i=1}^n v_i b_i) | f(\sum_{j=1}^n w_j b_j) \rangle = \langle \sum_{i=1}^n v_i f(b_i) | \sum_{j=1}^n w_j f(b_j) \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \overline{v}_i w_j \langle f(b_i) | f(b_j) \rangle = \sum_{i=1}^n \overline{v}_i w_i = \langle v | w \rangle \end{aligned}$$

□

**Proposition 4.16:** *Let $f : V \to V$ be an operator on a finite-dimensional inner product space $V$. Let $\mathcal{B}$ be an orthonormal basis for $V$, and let $A = [f]_{\mathcal{B}}$. Then $f$ is orthogonal/unitary if and only if $\overline{A}^t A = I$.*

*Proof.* $f$ is orthogonal iff $(f(b_1), \dots, f(b_n))$ is an orthonormal basis iff $\langle f(b_i) | f(b_j) \rangle = \delta_{ij}$ for all $i, j$. But

$$\langle f(b_i) | f(b_j) \rangle = \overline{[f(b_i)]}_{\mathcal{B}}^t [f(b_j)]_{\mathcal{B}} = \overline{(A[b_i]_{\mathcal{B}})}^t (A[b_j]_{\mathcal{B}}) = \overline{[b_i]}_{\mathcal{B}}^t \overline{A}^t A [b_j]_{\mathcal{B}} = [\overline{A}^t A]_j^i$$

Thus $f$ is orthogonal iff $\overline{A}^t A = I$. □

**Remark 4.17:** *In particular, we see that the matrix of $f^{-1}$ with respect to $\mathcal{B}$ is $\overline{A}^t$.*

**Definition 4.18:** *A matrix $A \in M_n(\mathbb{R})$ is said to be **orthogonal** if it satisfies $A^t A = I$. A matrix $A \in M_n(\mathbb{C})$ is said to be **unitary** if it satisfies $\overline{A}^t A = I$.*

**Exercise 4.19:** Show that the columns of an orthogonal/unitary matrix of $M_n(\mathbb{F})$ form an orthonormal basis of $\mathbb{F}^n_{col}$ endowed with the standard inner product. Show that its rows form an orthonormal basis of $\mathbb{F}^n_{row}$ endowed with the standard inner product.

# 5 Adjoint operator

## 5.1 Riesz representation theorem

Let $(V, \langle | \rangle)$ be an inner product space over a field $\mathbb{F}$. Fix a vector $u \in V$, and consider the map

$$\begin{aligned} \langle u| \quad : \quad & V \longrightarrow \mathbb{F} \\ & v \longmapsto \langle u|v\rangle \end{aligned}$$

**Example 5.1:** If $(b_1, \dots, b_n)$ is an orthonormal basis for $V$, we saw that for any vector $v \in V$, the first coordinate of $v$ in the basis $\mathcal{B}$ is given by $\langle b_i|v\rangle$. Hence the map $\langle b_1|$ takes a vector and outputs its first coordinate in the basis $\mathcal{B}$.

Note that $\langle u|$ is a linear map, since the inner product is linear in the second variable. It takes a vector as an input, and outputs a scalar.

**Definition 5.2:** *Let $V$ be a vector space over a field $\mathbb{F}$. A linear form (or linear functional) on $V$ is a linear map $V \to \mathbb{F}$.*

Now if $V$ is finite dimensional in fact all linear forms are of this type.

**Proposition 5.3:** *Suppose $V$ is finite dimensional, and let $l$ be a linear functional $l : V \to \mathbb{F}$. Then there is a unique vector $u \in V$ such that $l = \langle u|$.*

*Proof.* Existence of $u$: Let $(b_1, \dots, b_n)$ be an orthonormal basis for $V$. Set $u = \sum_{i=1}^n \overline{l(b_i)} b_i$. For each $j$ we have

$$\langle u|(b_j) = \langle u|b_j\rangle = \langle \sum_{i=1}^n \overline{l(b_i)} b_i | b_j\rangle = \sum_{i=1}^n l(b_i)\langle b_i|b_j\rangle = l(b_j)$$

hence $l = \langle u|$.

Uniqueness of $u$: Suppose $w$ is such that $\langle w| = l = \langle u|$ - then $\langle u|v\rangle = \langle w|v\rangle$ for all $v$, that is, $\langle u - w|v\rangle = 0$ for all $v$. Setting $v = u - w$ we get that $\langle u - w|u - w\rangle = 0$, so $u - w = 0$, meaning $u = w$. □

Geometry behind it: suppose $\dim V = n$. Note that if $l = \langle u|$, then $v \in \mathrm{Ker}(l)$ iff $\langle u|v\rangle = 0$ iff $u \perp v$, so $u \in (\mathrm{Ker} l)^\perp$. Now $l : V \to \mathbb{F}$ so $\dim(\mathrm{Ker} l) = n - 1$ - but this means that $\dim(\mathrm{Ker}\ l)^\perp = \dim V - \dim(\mathrm{Ker} l) = 1$ - we don't have so much choice for $u$!

## 5.2 Adjoint operator - existence and first properties

Suppose now that we have an operator $f : V \to V$. Given a vector $u \in V$, instead of looking at the map $v \mapsto \langle u|v\rangle$, we look at $l : v \mapsto \langle u|f(v)\rangle$. Because $f$ is linear, and $\langle \cdot|\cdot\rangle$ is linear in the second argument, it is itself linear as a composition of two linear maps.

But then by Proposition 5.3, this means that there exists a unique vector $u'$ such that $l = \langle u'|$, that is, such that $l(v) = \langle u|f(v)\rangle = \langle u'|v\rangle$ for all $v \in V$.

Let us fix an orthonormal basis $\mathcal{B} = (b_1, \dots, b_n)$ for $V$ and consider the matrix $A$ representing $f$ with respect to $\mathcal{B}$. We have for any two vectors $u, v \in V$

$$\langle u|f(v)\rangle \quad = \overline{[u]}_{\mathcal{B}}^t [f(v)]_{\mathcal{B}} = \overline{[u]}_{\mathcal{B}}^t A[v]_{\mathcal{B}}$$

so if we set $u'$ to be the vector such that $\overline{[u']}_{\mathcal{B}}^t = \overline{[u]}_{\mathcal{B}}^t A$, we get

$$\langle u|f(v)\rangle = \overline{[u']_{\mathcal{B}}}^t [v]_{\mathcal{B}} = \langle u'|v\rangle$$

Now $[u']_{\mathcal{B}} = \overline{\overline{[u]}_{\mathcal{B}}^t A}^t = \overline{A}^t [u]_{\mathcal{B}}$.

Thus if we denote by $f^*$ the operator represented by the matrix $A^* = \overline{A}^t$ with respect to $\mathcal{B}$ we get $u' = f^*(u)$ and

$$\langle u|f(v)\rangle = \langle f^*(u)|v\rangle$$

We have proved

**Proposition 5.4:** *Let $f : V \to V$ be an operator on a finite dimensional inner product space $V$. There exists a unique operator $f^* : V \to V$ such that for any $u, v \in V$ we have*

$$\langle u|f(v)\rangle = \langle f^*(u)|v\rangle$$

*Moreover, if $A$ is the matrix of $f$ with respect to some orthonormal basis $\mathcal{B}$, then the matrix of $f^*$ with respect to $\mathcal{B}$ is $A^* = \overline{A}^t$.*

**Definition 5.5:** *The map $f^*$ is called the **adjoint** of $f$. The matrix $A^*$ is called the adjoint of $A$.*

**Remark 5.6:**
1. *If $\mathbb{F} = \mathbb{R}$, $A^* = A^t$.*
2. *If $\dim V = 1$, the matrix $A$ is just a scalar $a$, and $a^* = \overline{a}$. Thus the adjoint of a matrix can be thought of as a higher dimensional analogue of the complex conjugate of a number.*

Suppose $f : V \to V$ is an orthogonal/unitary operator on a finite dimensional inner product space $V$. We have for all $u, v \in V$ that

$$\langle f^{-1}(u)|v\rangle = \langle f(f^{-1}(u))|f(v)\rangle = \langle u|f(v)\rangle$$

Thus we see that $f^{-1}$ satisfies the defining property of $f^*$. By uniqueness we must have $f^* = f^{-1}$. We have proved

**Lemma 5.7:** *Let $f : V \to V$ be an orthogonal/unitary operator on a finite dimensional inner product space $(V, \langle|\rangle)$. Then $f^* = f^{-1}$.*

The following properties of the adjoint are left as an exercise.

**Proposition 5.8:** *For any operators $f : V \to V$, $g : V \to V$ and $a \in \mathbb{F}$ we have*

1. $(f + g)^* = f^* + g^*$
2. $(af)^* = \bar{a}f^*$
3. $[f^*]^* = f$;

**Proposition 5.9:** *If $W$ is an $f$-invariant subspace, then $W^\perp$ is $f^*$-invariant.*

*Proof.* Assume that $f(w) \in W$ for all $w \in W$. Let $u \in W^\perp$: for any $w \in W$, we have $f(w) \in W$ so $\langle u|f(w)\rangle = 0$. But $\langle u|f(w)\rangle = \langle f * (u)|w\rangle$, thus for any $w \in W$ we have $\langle f^*(u)|w\rangle = 0$. In other words, $f^*(u) \in W^\perp$. This proves the claim. □

## 5.3 Self-adjoint operators

**Definition 5.10:** *Let $f : V \to V$ be an operator over a finite dimensional inner product space $V$ (over $\mathbb{R}$ or $\mathbb{C}$). $f$ is called a **self-adjoint operator** if $f^* = f$.*

**Remark 5.11:** *An operator $f$ on a Euclidean or Hermitian space is self-adjoint iff its matrix $A = [f]_\mathcal{B}$ in some orthonormal basis satisfies $A = A^*$ iff its matrix $A = [f]_\mathcal{B}$ in any orthonormal basis satisfies $A = A^*$.*

**Definition 5.12:** *Let $A$ be an $n$-by-$n$ matrix over $\mathbb{F}$, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. If $A = A^*$, we say $A$ is a **self-adjoint matrix**. In the case $\mathbb{F} = \mathbb{C}$, it is also called a **Hermitian matrix**. If $\mathbb{F} = \mathbb{R}$, the self-adjointness condition simply means that $A = A^t$, i.e. that the matrix $A$ is symmetric with respect to the diagonal. Such matrices are called **symmetric matrices**.*

**Example 5.13:** Of the 3 matrices

$$\begin{bmatrix} -1 & 3 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 5 & 0 \\ 0 & 2i \end{bmatrix}, \begin{bmatrix} 0 & 2+3i \\ 2-3i & -4 \end{bmatrix}$$

the first and the third are self-adjoint, while the second is not.

**Remark 5.14:** *The elements on the diagonal of a self-adjoint matrix are real.*

The three propositions that follow concern eigenvalues and eigenvectors of self-adjoint operators.

**Proposition 5.15:** *Every eigenvalue of a self-adjoint operator is real.*

*Proof.* Let $f : V \to V$ satisfy $f = f^*$ and $f(u) = \lambda u$ for some $u \neq 0$ and $\lambda \in \mathbb{C}$. Then

$$\begin{aligned}\lambda\langle u|u\rangle &= \langle u|\lambda u\rangle = \langle u|f(u)\rangle \\ &= \langle f^*(u)|u\rangle = \langle f(u)|u\rangle = \langle \lambda u|u\rangle \\ &= \bar{\lambda}\langle u|u\rangle\end{aligned}$$

Since $\langle u|u\rangle \neq 0$, we must have $\lambda = \bar{\lambda}$. □

**Proposition 5.16:** *Eigenvectors corresponding to distinct eigenvalues of a self-adjoint operator are orthogonal.*

*Proof.* The proof is based on a similar calculation. Let $f : V \to V$ be self-adjoint and $f(u) = \lambda u, f(v) = \mu v$ for some $u \neq 0, v \neq 0$ and $\lambda, \mu \in \mathbb{R}$ (the eigenvalues are real by the previous proposition), $\lambda \neq \mu$. Then

$$\begin{aligned}\mu\langle u|v\rangle &= \langle u|\mu v\rangle = \langle u|f(v)\rangle = \langle f^*(u)|v\rangle \\ &= \langle f(u)|v\rangle = \langle \lambda u|v\rangle \\ &= \bar{\lambda}\langle u|v\rangle = \lambda\langle u|v\rangle\end{aligned}$$

Since $\lambda \neq \mu$, we must have $\langle u|v\rangle = 0$. □

Note that for now the two propositions above are in conditional mode only. We do not know yet if and when self-adjoint operators have eigenvalues.

**Proposition 5.17:** *Every self-adjoint operator on a finite-dimensional inner product space admits an eigenvalue.*

*Proof.* Case $\mathbb{F} = \mathbb{C}$.
We have already seen that in this case every operator (not necessarily self-adjoint) admits an eigenvalue.
Case $\mathbb{F} = \mathbb{R}$.
The argument is slightly trickier: Let $A = [f]_{\mathcal{B}}$ for some **orthonormal** basis $\mathcal{B}$ of $V$.

First consider the operator $g : \mathbb{C}^n_{col} \to \mathbb{C}^n_{col}$ defined by $g\bar{x} = A\bar{x}$, i.e. the operator $\mathbb{C}^n_{col} \to \mathbb{C}^n_{col}$ whose matrix relative to the standard basis $\mathcal{E}$ of $\mathbb{C}^n_{col}$ is exactly $A$. Note that the characteristic polynomials of $f$ and $g$ are equal, indeed $\chi_g = \chi_A = \chi_f$. Eigenvalues of $g$ are exactly the roots of $\chi_A$ in $\mathbb{C}$, while eigenvalues of $f$ are the **real** roots of $\chi_A$. By the fundamental theorem of algebra, $\chi_A$ admits a root $\lambda \in \mathbb{C}$, which is therefore an eigenvalue of $g$.

Note now that $[g^*]_{\mathcal{E}} = A^* = A = [g]_{\mathcal{E}}$ so $g$ is self-adjoint. Thus, by Proposition 5.15, $\lambda$ is a real number!

We have just shown that $\chi_f$ has a (real) root, meaning that $f$ has an eigenvalue. □

# 6 Spectral theorems

The goal of this section is to obtain necessary and sufficient conditions for an operator $f : V \to V$ to be diagonalizable in an orthonormal basis.

There are no simple necessary and sufficient conditions for diagonalizability of an operator or a matrix (even over $\mathbb{C}$). However, there are such simple conditions for diagonalizability in an orthonormal basis. In this section we will find these conditions.

## 6.1 Diagonalizing an operator in an orthonormal basis

Let $(V, \langle | \rangle)$ be an inner product space of finite dimension. Let $f : V \to V$ be an operator on $V$.

**Definition 6.1:** *We say $f$ is* ***diagonalizable in an orthonormal basis*** *if there exist an orthonormal basis $\mathcal{B}$ such that $[f]_\mathcal{B}$ is diagonal.*

**Remark 6.2:** *This is equivalent to saying that there is an orthonormal basis of $V$ which consists of eigenvectors of $f$.*

If $\mathcal{C}$ is another orthonormal basis, then $A = [f]_\mathcal{C}$ is not necessarily diagonal. However, denote by $M = M^\mathcal{C}_\mathcal{D}$ the change of basis matrix: we have $[f]_\mathcal{B} = M^{-1}AM$.

Now what can we say about the matrix $M$?

**Lemma 6.3:** *Let $(V, \langle | \rangle)$ be an inner product space, let $\mathcal{B}$ be an orthonormal basis for $V$, and let $\mathcal{C}$ be any basis of $V$. Then $\mathcal{C}$ is orthonormal as well iff the change of basis matrix $M^\mathcal{B}_\mathcal{C}$ is orthogonal iff $M^\mathcal{C}_\mathcal{B}$ is orthogonal.*

*Proof.* First recall that $M^\mathcal{B}_\mathcal{C} = (M^\mathcal{C}_\mathcal{B})^{-1}$, and that if a matrix is orthogonal its inverse also is, so the two last statements are equivalent.

Suppose $\mathcal{B} = (b_1, \dots, b_n)$ and $\mathcal{C} = (c_1, \dots, c_n)$. The columns of $M = M^\mathcal{C}_\mathcal{B}$ are the coordinate vectors $[c_1]_\mathcal{B}, \dots, [c_n]_\mathcal{B}$. Thus the matrix $\overline{M}^t$ has rows given by $\overline{[c_1]}^t_\mathcal{B}, \dots, \overline{[c_n]}^t_\mathcal{B}$, so the element on the $i$-th row and the $j$-th column of the matrix $\overline{M}^t M$ is exactly $[\overline{M}^t M]_{ij} = \overline{[c_i]}^t_\mathcal{B}[c_j]_\mathcal{B}$.

Since $\mathcal{B}$ is an orthonormal basis, we have by Parseval that

$$[\overline{M}^t M]_{ij} = \overline{[c_i]}^t_\mathcal{B}[c_j]_\mathcal{B} = \langle c_i | c_j \rangle$$

Thus

$$\begin{aligned}\overline{M}^t M = I \quad &\iff \quad [\overline{M}^t M]_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \iff \langle c_i | c_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \\ &\iff \quad \mathcal{C} \text{ is an orthonormal basis .}\end{aligned}$$

□

## 6.2 Diagonalizing a matrix orthogonally/unitarily

Let $A \in M_n(\mathbb{F})$: it is diagonalizable if there exists an invertible $M \in M_n(\mathbb{F})$ such that $MAM^{-1}$ is a diagonal matrix.

In light of the discussion above we put an additional constraint on $M$ – it should be an orthogonal/unitary matrix in the real/complex case. So in the real case we are looking for $M$ satisfying $MM^t = I$ such that $MAM^{-1}$ is a diagonal matrix, and in the complex case we are looking for $M$ satisfying $M\overline{M^t} = I$ such that $MAM^{-1}$ is a diagonal matrix.

**Definition 6.4:** *Let $A \in M_n(\mathbb{R})$. We say $A$ is* ***orthogonally diagonalizable*** *if there exists an orthogonal matrix $M$ such that $MAM^{-1}$ (equivalently $MAM^t$) is diagonal.*

*Let $A \in M_n(\mathbb{C})$. We say $A$ is* ***unitarily diagonalizable*** *if there exists a unitary matrix $M$ such that $MAM^{-1}$ (equivalently $MAM^*$) is diagonal.*

Thus we have

**Proposition 6.5:** *Let $(V, \langle | \rangle)$ be an inner product space of finite dimension. Let $f : V \to V$ be an operator on $V$, and let $A = [f]_\mathcal{C}$ be the matrix for $f$ in an orthonormal basis $\mathcal{C}$ for $V$.*

*The operator $f$ is diagonalizable in an orthonormal basis iff $A$ is orthogonally (resp. unitarily) diagonalizable in the real case (resp. the complex case).*

## 6.3 A necessary condition

Note that if $f$ is represented by the diagonal matrix $A$ in the orthonormal basis $\mathcal{B}$, then we have that $f^*$ is represented by the matrix $A^* = \overline{A}^t$ which is also diagonal.

If $\mathbb{F} = \mathbb{R}$ moreover, we have in fact $A = A^*$, so $f$ must be self-adjoint. If $\mathbb{F} = \mathbb{C}$, we cannot deduce that $f$ is self-adjoint, but we can note that $AA^* = A^*A$ since both matrices are diagonal. Thus we have shown a necessary condition for each of the real and complex cases for an operator to be diagonalizable.

**Proposition 6.6:** *Let $f$ be an operator on a finite dimensional inner product space $V$. Suppose $f$ is diagonalizable in an orthonormal basis.*

*If $\mathbb{F} = \mathbb{R}$ then $f$ is self-adjoint. If $\mathbb{F} = \mathbb{C}$ then $f$ commutes to its adjoint $f^*$.*

It turns out these necessary conditions are in fact sufficient.

## 6.4 The real case

**Theorem 6.7:** *Let $f : V \to V$ be an operator in a Euclidean space. Then $f$ is diagonalizable in an orthonormal basis iff $f$ is self-adjoint.*

*Proof.* We prove this by induction on $n = \dim V$. For $n = 1$ the claim holds, just take the basis $(v)$, where $v$ is any unit vector.
For $n \geq 2$ assume the claim true for all $k < n$. Then:
1. By Proposition 5.17 $f$ has an eigenvector $v \in V$. Normalizing, we may assume $v$ to be a unit vector.
2. $W = \mathrm{Span}(v)$ is $f$-invariant, hence by Proposition 5.9 $f^*$ preserves $W^\perp$. However $f = f^*$, so $f$ preserves $W^\perp$. Note that the operator $f|_{W^\perp}$ is self-adjoint (meaning that $\langle u|f(v)\rangle = \langle v|f(u)\rangle$ for all $u, v \in W^\perp$, but this holds even for all $u, v \in V$). So by the induction assumption $W^\perp$ has an orthonormal basis $(b_1, \ldots, b_{n-1})$ consisting of eigenvectors of $f$. Hence $(v, b_1, \ldots, b_{n-1})$ is an orthonormal basis of $V$ consisting of eigenvectors of $f$.

□

The matrix version of this theorem is given by:

**Theorem 6.8:** *Let $A \in M_n(\mathbb{R})$. Then $A$ is orthogonally diagonalizable if and only if $A$ is symmetric, i.e $A = A^t$.*

## 6.5 The complex case

For the technique used in the real case to work, we need two things: a) existence of an eigenvector $v$ for $f$, and b) $f$-invariance of its orthogonal complement.

The first condition holds for every operator $f$ in a complex inner product space. The second one holds for many important classes of complex operators, including self-adjoint and isometries. To characterize <u>all</u> unitarily diagonalizable operators, we give the following definition.

**Definition 6.9:** *An operator $f : V \to V$ on an inner product space is called* ***normal*** *if $ff^* = f^*f$. A matrix $A \in M_n(\mathbb{F})$ is called normal if $A^*A = AA^*$.*

**Remark 6.10:**
- *Self-adjoint operators are normal: if $f = f^*, then ff^* = ff = f^*f$.*
- *Orthogonal/unitary operators are normal: if $f^* = f^{-1}$ then $ff^* = ff^{-1} = \mathrm{Id} = f^*f$.*

**Theorem 6.11:** *Let $f : V \to V$ be an operator on a finite dimensional complex inner product space. Then $f$ is diagonalizable in an orthonormal basis if and only if it is normal.*

We will need the following lemma

**Lemma 6.12:** *Let $f : V \to V$ be an operator over a vector space over $\mathbb{C}$. Let $W$ be a subspace which is both $f$-invariant and $f^*$-invariant.*

*Then $(f\,|_W)^* = (f^*)\,|_W$, that is, the adjoint operator of the restriction of $f$ to $W$ is the restriction of the adjoint operator.*

*Proof.* Denote by $g : W \to W$ the restriction of $f$ to $W$ (it is simply the map defined by $g(w) = f(w)$ for any $w \in W$). Now for any $w, w' \in W$ we have $f^*(w), f^*(w') \in W$ and $\langle w|g(w')\rangle = \langle w|f(w')\rangle = \langle f^*(w)|w'\rangle = \langle (f^*)\mid_W (w)|w'\rangle$.

On the other hand, by Proposition 5.4, the adjoint operator $g^*$ of $g$ is the unique operator $W \to W$ which satisfies $\langle w|g(w')\rangle = \langle g^*(w)|w'\rangle$ for all $w, w' \in W$ so we must have $g^* = (f^*)\mid_W$. □

We can now prove 6.11

*Proof.* We have seen that if $f$ is diagonalizable in an orthonormal basis, then it is normal. We now prove the other direction.

Let $f$ denote a normal operator. We work by induction on $n = \dim V$. For $n = 0$ the claim holds.

Since we are over $\mathbb{C}$, $f$ admits an eigenvalue $\lambda$, denote by $V_\lambda$ the associated eigenspace - it is not reduced to $\{0\}$. We have $V = V_\lambda \oplus V_\lambda^\perp$. If $V = V_\lambda$, the matrix of $f$ in any basis, and thus in any orthonormal basis, is $\lambda I$ which is diagonal, so we may assume $\dim(V_\lambda)$ is strictly smaller than $\dim V$.

We note the following

1. $V_\lambda$ is $f$-invariant;
2. $V_\lambda^\perp$ is $f^*$-invariant;
3. $V_\lambda$ is $f^*$-invariant: indeed, if $v \in V_\lambda$ we have $f(f^*(v)) = f^*(f(v)) = f^*(\lambda v) = \lambda f^*(v)$ so $f^*(v) \in V_\lambda$;
4. $V_\lambda^\perp$ is $f$-invariant: indeed, applying Proposition 5.9 to the $f^*$-invariant subspace $V_\lambda$, we get that $V_\lambda^\perp$ is $(f^*)^*$-invariant, but $(f^*)^* = f$, which proves the claim.

Thus each of the subspaces in the direct sum decomposition $V = V_\lambda \oplus V_\lambda^\perp$ is $f$-invariant.

Conditions 1. and 3. enable us to apply Lemma 6.12 to $W = V_\lambda$, hence we have $(f\mid_{V_\lambda})^* = f^*\mid_{V_\lambda}$. We deduce that $f\mid_{V_\lambda}$ is normal (if two maps commute then their restriction to a subspace also commute). Similarly from 2. and 4. and Lemma 6.12 we get that $f\mid_{V_\lambda^\perp}$ is normal.

Therefore by induction hypothesis there exist orthonormal bases $b_1, \dots, b_k$ and $(b_{k+1}, \dots, b_n)$ respectively of $V_\lambda$, $V_\lambda^\perp$, whose vectors are eigenvectors of $f\mid_{V_\lambda}$ and $f\mid_{V_\lambda^\perp}$ respectively - but this implies that $(b_1, \dots, b_n)$ is an orthonormal basis for $V$ of eigenvectors of $f$. □

And the matrix version of this result is:

**Theorem 6.13:** *Let $A \in M_n(\mathbb{C})$. Then $A$ is unitarily diagonalizable if and only if $A$ is normal.*

# 7 Families of operators

| | Definition | Inner product | Matrix | dim = 1 | Eigenvalues |
|---|---|---|---|---|---|
| Orthogonal ( $\mathbb{F} = \mathbb{R}$) | for all $u, v \in V$ $\langle f(u)\|f(v)\rangle = \langle u\|v\rangle$ | Preserves $\langle\|\rangle$ | $A^t A = I$ | $a^2 = 1$, i.e $a = \pm 1$ | $\lambda = \pm 1$ |
| Unitary ($\mathbb{F} = \mathbb{C}$) | for all $u, v \in V$ $\langle f(u)\|f(v)\rangle = \langle u\|v\rangle$ | Preserves $\langle\|\rangle$ | $A^* A = I$ | Unit circle $\{a \in \mathbb{C} \mid \|a\| = 1\}$ | $\|\lambda\| = 1$ |
| Self-adjoint | $f^* = f$ | for all $u, v \in V$ $\langle u)\|f(v)\rangle = \langle f(u)\|v\rangle$ | $A^* = A$ | $\overline{a} = a$, i.e. $a \in \mathbb{R}$ | $\lambda \in \mathbb{R}$ |
| Normal | $f^* f = f f^*$ | for all $u, v \in V$ $\langle f(u)\|f(v)\rangle = \langle f^*(u)\|f^*(v)\rangle$ | $AA^* = A^* A$ | everyone | if $f(v) = \lambda v$ then $f^*(v) = \overline{\lambda} v$ |