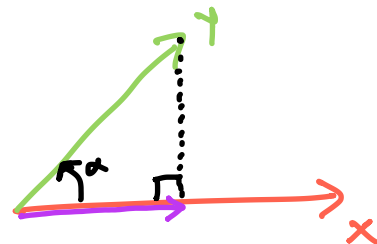


# Inner Product and Hilbert Spaces

- metric  $\rightarrow$  measures distances
- norm  $\rightarrow$  measures distances, lengths
- inner product  $\rightarrow$  measures distances, lengths, angles.

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cos(\alpha)$$



In ML we use

$$\text{cosine similarity: } \cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

Inner Product  $\leftrightarrow$  scalar product  $\leftrightarrow$  dot product

Def: Consider a vector space  $V$ . A mapping  $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$  is called an inner product if

linearity { (P1)  $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$

(P2)  $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$  ( $\lambda \in F$ )

symmetry { (P3)  $\langle x, y \rangle = \langle y, x \rangle$  (if  $F$  is  $\mathbb{R}$ )

$\langle x, y \rangle = \overline{\langle y, x \rangle}$  (if  $F$  is  $\mathbb{C}$ )  
complex conjugate

positive definite { (P4)  $\langle x, x \rangle \geq 0$

(P5)  $\langle x, x \rangle = 0 \Leftrightarrow x = 0$

Examples: · Euclidean inner product on  $\mathbb{R}^n$ :

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

· On  $\mathbb{C}$ ,  $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$

·  $\mathcal{C}([a, b])$ :  $\langle f, g \rangle = \int_a^b f(t)g(t) dt$

is an inner product (but space would not be complete)

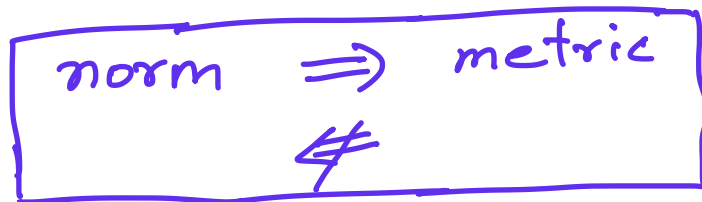
Def: A vector space with a norm is called a normed space. If a normed space is complete (all Cauchy sequences converge), then  $V$  is called a Banach Space. A vector space with an inner product is called a pre-Hilbert-space. If it is additionally complete, then  $V$  is called a Hilbert space.

inner product $\Rightarrow$ norm
inner product $\not\Leftarrow$ norm

Consider a vector space with an inner product  $\langle \cdot, \cdot \rangle$ . Define  $\|\cdot\|: V \rightarrow \mathbb{R}$  as  $\|x\| := \sqrt{\langle x, x \rangle}$ .

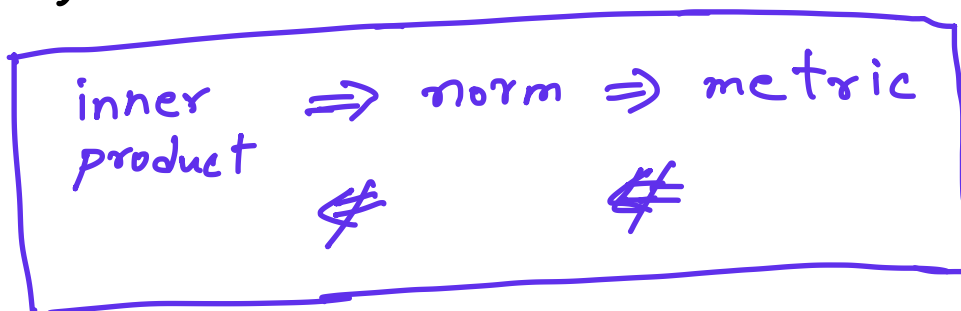
Then  $\|\cdot\|$  is a norm on  $V$ , the norm is induced by  $\langle \cdot, \cdot \rangle$ .

In general, the other way does not work.



Consider a vector space  $V$  with norm  $\|\cdot\|$ . Then  $d: V \times V \rightarrow \mathbb{R}$ ,  $d(x, y) := \|x - y\|$  is a metric on  $V$ , the metric is induced by the norm.

In general, the other direction does not work.



# Orthogonal Basis and Projections

Def: Consider a pre-Hilbert-space  $V$ . Two vectors  $v_1, v_2 \in V$  are called orthogonal if  $\langle v_1, v_2 \rangle = 0$ .

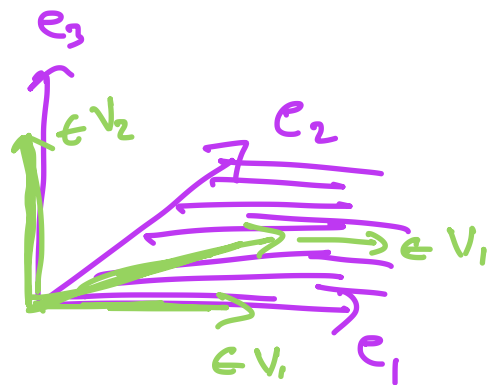
Notation:  $v_1 \perp v_2$

Two sets  $V_1, V_2 \subset V$  are called orthogonal if  $\forall v_1 \in V_1, v_2 \in V_2 : \langle v_1, v_2 \rangle = 0$

$$V_1 = \text{span}\{e_1, e_2\}$$

$$V_3 = \text{span}\{e_3\}$$

$$V_1 \perp V_3$$



Vectors are called orthonormal if additionally the two vectors have norm of 1:

$$\bullet \langle v_1, v_2 \rangle = 0$$

$$\bullet \|v_1\| = 1, \quad \|v_2\| = 1$$

A set of vectors  $v_1, v_2, \dots, v_n$  is called orthonormal if any two vectors are orthonormal.

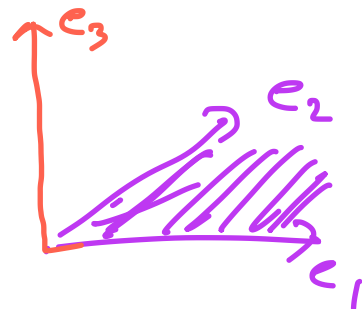
For a set  $S \subseteq V$  we define its orthogonal complement  $S^\perp$  as follows:

$$S^\perp := \{ v \in V \mid v \perp s, \forall s \in S \}$$

$$V_1 = \text{span} \{ e_1, e_2 \}$$

$$V_2 = \text{span} \{ e_3 \}$$

$$V_1^\perp = V_2, \quad V_2^\perp = V_1$$



## Orthogonal Projections

Def:  $A \in$

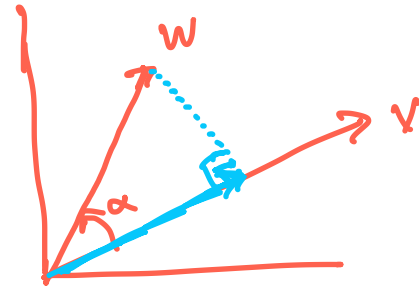
Construction: Let  $v_1, \dots, v_n$  be an orthogonal basis of  $U$ . Define  $P_U: V \rightarrow U$  by

$$P_U(w) = \sum_{i=1}^n \frac{\langle w, v_i \rangle}{\|v_i\|_2^2} v_i$$

Intuition:

If  $\|w\| = 1$   
If  $\|v\| = 1$ , then

$$\|P_v(w)\| = |\langle v, w \rangle| \\ = |\cos \alpha|$$



In particular.  $\langle v, w \rangle = \cos \alpha$

Remark: In an orthonormal basis  $u_1, \dots, u_n$  the representation of a vector  $v$  is given by

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

### Gram-Schmidt orthogonalization

It is a procedure that takes any basis  $v_1, \dots, v_n$  of a finite-dim vector space and transforms it into another basis  $u_1, \dots, u_n$  that is orthonormal.

Intuition: iterative procedure

Step 1:  $u_1 = \frac{v_1}{\|v_1\|}$

$$U_1 = \text{span}\{u_1\}$$

Step k: Assume that we already have

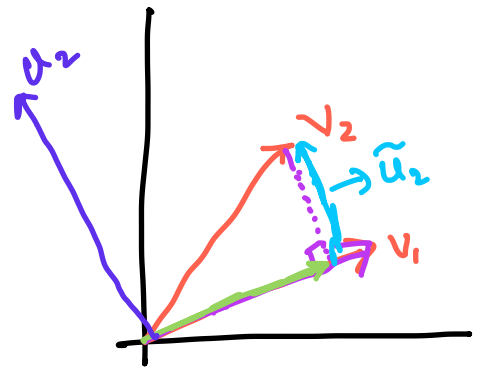
$u_1, u_2 \dots u_{k-1}$ .

- Project  $v_k$  on  $U_{k-1}$  and keep "the rest"

$$\tilde{u}_k = \underline{v_k} - \underline{P_{U_{k-1}}(v_k)}$$

- Renormalize :

$$\underline{u_k} = \frac{\tilde{u}_k}{\|\tilde{u}_k\|}$$



In practice use Householder reflections for a numerically stable orthogonalization.

# Orthogonal Matrices

Def: Let  $Q \in \mathbb{R}^{n \times n}$  be a matrix with orthonormal column vectors (w.r.t. Euclidean inner product). Then  $Q$  is called an orthogonal matrix.

If  $Q \in \mathbb{C}^{n \times n}$  and the columns are orthonormal (w.r.t. the standard inner product on  $\mathbb{C}$ ), then it is called unitary.

## Examples:

• Identity:  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

• Reflection:  $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  reflection about x-axis

• Permutation:  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

• Rotation:  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$



• Rotation in  $\mathbb{R}^3$

• rotate about one of the axes:

$$R_{\theta, 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

• General rotation: can be written as a product of "elementary" rotation

Properties of orthogonal matrix  $Q$ :

• columns are orthogonal  $\Leftrightarrow$  rows are orthogonal

•  $Q$  is always invertible, and  $Q^{-1} = Q^T$

•  $Q$  realizes an isometry:  $\forall v \in V : \|Qv\| = \|v\|$

$\hookrightarrow$  keeps lengths intact

•  $Q$  preserves angles:  $\langle Qu, Qv \rangle = \langle u, v \rangle$   
 $\forall u, v \in V$

•  $|\det Q| = 1$

The respective properties also hold for unitary matrices  $U$ . ( $U^{-1} = \bar{U}^T$ )

Theorem: Let  $S \in L(V)$  for a real vector space  $V$ . Then the following are equivalent:

(a)  $S$  is an isometry:  $\|Sv\| = \|v\|, v \in V$ .

(b) There exists an orthonormal basis of  $V$  such that the matrix of  $S$  has the following form:

$$M = \begin{pmatrix} \square & & & 0 \\ & \square & & \\ & & \square & \\ 0 & & & \square \cdots \square \end{pmatrix}$$

where each of the little blocks

- either a  $1 \times 1$  matrix (one real number) with a value  $+1$  or  $-1$ .
- or a  $2 \times 2$  rotation matrix.

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$