- $\boxed{\text{Set } V \text{ with } \begin{cases} \text{sum} \\ \text{scalar prod} \end{cases}}$

vector space over $\mathbb{R}$.

- $\boxed{\text{Basis}}$ ← (finite)

- Algebra ← Geometry

dot prod. in $\mathbb{R}^n$

① length $\|\vec{u}\|$

② angle. $\angle(\vec{u}, \vec{v})$

- $\vec{u} \cdot \vec{v} := u_1 v_1 + \dots + u_n v_n$

↓ norm

- $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + \dots + u_n^2}$

inner product

$\angle(\vec{u}, \vec{v}) = \arccos \dfrac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$ ← Ex: dot prod in $\mathbb{R}^n$

Northeastern University, Department of Mathematics

MATH 4570 - Matrix Methods in Data Analysis and Machine Learning

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§5 Inner product spaces

Contents

1. Inner Product Spaces

$\underline{\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}}$

Recall that for vectors $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, the **dot product** of $\vec{u}$ and $\vec{v}$ is $\vec{u}\cdot\vec{v} = u_1v_1 + \cdots + u_nv_n$ $\xrightarrow{\text{want}}$ $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$

Theorem 1 (Properties of the dot Product)**.** *For vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and a scalar $c \in \mathbb{R}$, the following hold:*

(1.) $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$

(2.) $(\vec{u}+\vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$

(3.) $(c\vec{u})\cdot\vec{v} = c(\vec{u}\cdot\vec{v})$

(4.) $\vec{u}\cdot\vec{u} \geq 0$

(5.) $\vec{u}\cdot\vec{u} = 0 \iff \vec{u} = \vec{0}$

$\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$

Definition 2 (Inner Product)**.** Let V be a real vector space. An **inner product** on V is a binary function

complex

$$\langle -, - \rangle : V \times V \to \mathbb{R}, \text{ or } \mathbb{C}$$

such that for vectors $\vec{u}, \vec{v}, \vec{w} \in V$ and a scalar $c \in \mathbb{R}$, the following hold:

(1.) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.   (1'). $\langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$ $\mathbb{C}$

(2.) $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.

(3.) $\langle c\vec{u}, \vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$.

(4.) $\langle \vec{u}, \vec{u} \rangle \geq 0$

(5.) $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if $\vec{u} = \vec{0}$.

We call V an **inner product space**.

Prop: over $\mathbb{R}$: $\langle \vec{u}, c\vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$

over $\mathbb{C}$: $\langle \vec{u}, c\vec{v} \rangle = \bar{c}\langle \vec{u}, \vec{v} \rangle$

$$\overline{\langle c\vec{v}, \vec{u} \rangle} = \overline{c\langle \vec{v}, \vec{u} \rangle} = \bar{c}\,\overline{\langle \vec{v}, \vec{u} \rangle}$$

Ex: dot prod. on $\mathbb{C}^n$

$$\vec{u} \cdot_{\mathbb{C}} \vec{v} := u_1\bar{v}_1 + \cdots + u_n\bar{v}_n$$

if $\vec{u} \cdot \vec{v} = \bar{u}_1 v_1 + \cdots + \bar{u}_n v_n$

$\vec{v} \cdot \vec{u} = \bar{v}_1 u_1 + \cdots + \bar{v}_n u_n$

Example 3. (Weighted dot products) Let $c_1, ..., c_n$ be positive numbers. The weighted inner product on $\mathbb{F}^n$ is

> 0 dot $\mathbb{R}$ $\mathbb{C}$

$$\langle \vec{u}, \vec{v} \rangle_w := c_1 u_1 \bar{v}_1 + \cdots + c_n u_n \bar{v}_n$$

$$= \vec{u}^T \begin{bmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{bmatrix} \bar{\vec{v}}$$

A is a positive definite ~~symmetric~~.

$\langle \vec{u}, \vec{v} \rangle := \vec{u}^T A \bar{\vec{v}}$

Example 4. Let $P_n(\mathbb{F})$ be the vector space of polynomials of degree at most n with coefficient in $\mathbb{F} = \mathbb{R}$ ~~or $\mathbb{C}$~~. An inner product on $P_n(\mathbb{F})$ can be defined as

function $\mathbb{C}$

$$\langle f(t), g(t) \rangle := \int_{a}^{b} f(t)\overline{g(t)}\,dt$$

Definition 5. Two vectors $\vec{u}$ and $\vec{v}$ are called **orthogonal** if

$$\langle \vec{u}, \vec{v} \rangle = 0$$

?

inner prod $\Longrightarrow$

2. Norms

$$\|e^x\| = \sqrt{\langle e^x, e^x \rangle}$$
$$= \sqrt{\int_0^1 (e^x)^2 dx}$$

Definition 6 (Norm of a Vector)**.** Let V be a inner product space over $\mathbb{F}$. The **length** or **norm** of a vector $\vec{v} \in V$ induced by inner product , denoted by $||\vec{v}||$, is defined as

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

Prop: $\|\vec{v}\| \geq 0$

$$\|\vec{v}\| = 0 \iff \vec{v} = \vec{0}$$

Def: $\vec{v}$ is unit vector if $\|\vec{v}\| = 1$.

Proposition 7. *For any vector $\vec{v} \in V$ and any scalar $c \in \mathbb{F}$ one obtains*

$$||c \cdot \vec{v}|| = |c| \cdot ||\vec{v}||.$$

Theorem 8 (Pythagorean Theorem)**.** *If two vectors $\vec{u}, \vec{v} \in V$ are orthogonal, then they satisfy the **Pythagorean Relation***

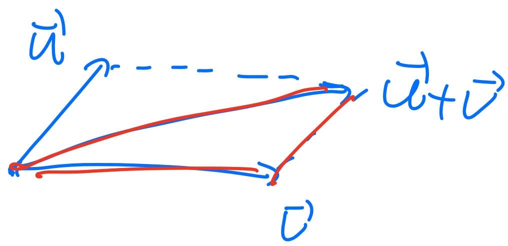
$$||\vec{u} + \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2.$$

$$\langle \vec{u}+\vec{v}, \vec{u}+\vec{v} \rangle$$

Theorem 9 (**Cauchy-Schwarz inequality**).

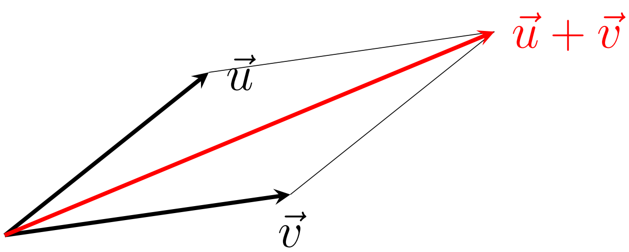
$$\left| \langle \vec{x}, \vec{y} \rangle \right| \leq \|\vec{x}\| \|\vec{y}\|$$

equality holds $\iff$ $\vec{y} = c\vec{x}$.



Proposition 10 (Triangle Inequality)**.** *Two vectors* $\vec{u}, \vec{v} \in V$ *satisfy*

$$||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||.$$



Definition 11. (Angles Between Vectors) The **angle between two nonzero vectors** $\vec{u}, \vec{v} \in V$ is the the angle $0 \leq \theta \leq \pi$ satisfying

$$\cos \vec{\theta} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \|\vec{v}\|}$$

A vector space V with norm is called a **normed vector space**.

Definition 12. A **norm** on V is a map from V to $\mathbb{F}$ such that

(1) $||\vec{x}|| \geq 0$ for all $\vec{x} \in V$. $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}$.
(2) $||c\vec{x}|| = |c|||\vec{x}||$ for all $\vec{x} \in V$ and $c \in \mathbb{F}$.
(3) The triangle inequality $||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ holds for all vectors in V.

Example 13. (l^p spaces) Let $1 \leq p < \infty$, it is natural to define l^p norms on $\mathbb{F}^n$

$\vec{V} \in \mathbb{R}^n$

$$\vec{V} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$\|\vec{V}\|_p = \sqrt[p]{|v_1|^p + \cdots + |v_n|^p}.$$

e.g. $p=2 \Rightarrow \|\vec{V}\|_2 = \|\vec{V}\| = \sqrt{v_1^2 + \cdots + v_n^2} = \sqrt{\langle \vec{V}, \vec{V} \rangle} \Leftarrow$ dot prod.

$p=1 \Rightarrow \|\vec{V}\|_1 = |v_1| + \cdots + |v_n|$

• If $p \neq 2$, $\|\vec{V}\|_p$ is not induced from any inner prod.

Example 14. (l^∞ spaces) It is natural to define l^∞ norms on $\mathbb{F}^n$

$$\|\vec{V}\|_\infty = \max_{1 \leq i \leq n} \{ |v_i| \}$$

Example 15. (Norms on $\mathbb{F}^{m\times n}$ induced by norms on $\mathbb{F}^n$) Normed matrix vector spaces $\mathbb{F}^{m\times n}$. Using norms on $\mathbb{F}^n$, one can define norms on matrix vector spaces

$$\|A\| = \sup \{ \|A\vec{x}\| \mid \vec{x} \in \mathbb{F}^n \text{ with } \|\vec{x}\| = 1 \}$$

$$= \sup \left\{ \frac{\|A\vec{x}\|}{\|\vec{x}\|} \,\middle|\, \vec{x} \in \mathbb{F}^n, \vec{x} \neq \vec{0} \right\}$$

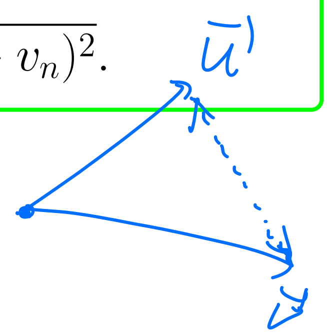
smallest upper bound.

Example 16. Infinity norm on $\mathbb{F}^{m\times n}$.

$$\|A\|_\infty = \max_{1\le i\le n}\left\{\sum_{j=1}^{n} |a_{ij}|\right\}$$

Definition 17 (Distance Between Vectors)**.** The **distance** $\operatorname{dist}(\vec{u},\vec{v})$ between vectors $\vec{u},\vec{v}\in\mathbb{R}^n$ is defined as

$$\operatorname{dist}(\vec{u},\vec{v}) = ||\vec{u}-\vec{v}|| = \sqrt{(u_1-v_1)^2+\cdots+(u_n-v_n)^2}.$$



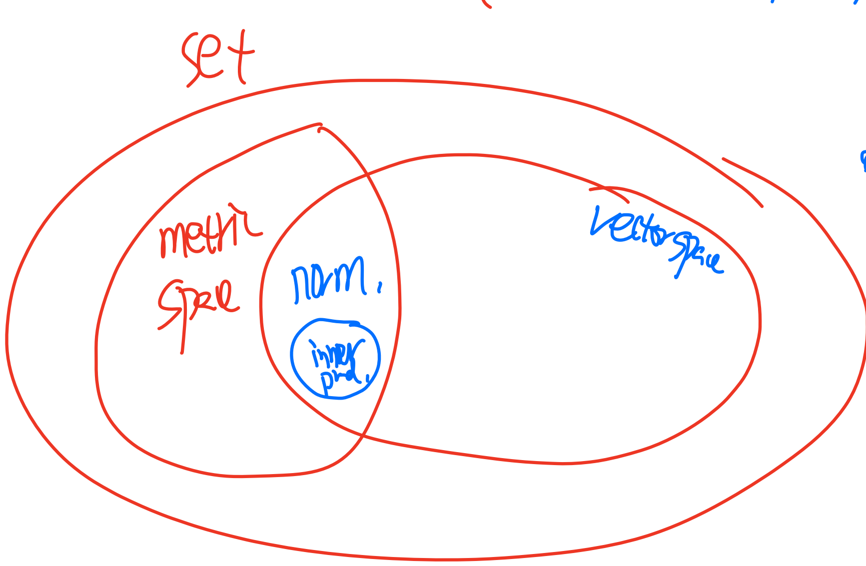
Def: Let S be a set.

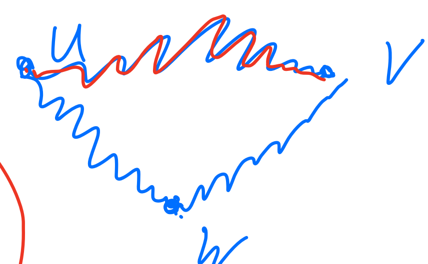
The (distance) metric of $u, v \in S$ is

$$\operatorname{dist}(-,-): S\times S \longrightarrow \mathbb{R}$$

such that:

1. $\operatorname{dist}(u,v) = \operatorname{dist}(v,u)$
2. $\operatorname{dist}(u,v) = 0 \iff u = v$
3. $\operatorname{dist}(u,v) \le \operatorname{dist}(u,w) + \operatorname{dist}(v,w)$





Ex: $d(u,v) = \begin{cases} 0 & \text{if } u=v \\ 1 & \text{if } u\neq v \end{cases}$

- Inner product Space V with $+$, scalar prod, $\langle\ ,\ \rangle$

3. Orthogonal Projections and Orthonormal Bases

non-zero

Definition 18 (Orthogonal Set)**.** A set $\{\vec{u}_1, \dots, \vec{u}_p\}$ of vectors in a inner vector space V is called **orthogonal** if

$$\langle \vec{u}_i, \vec{u}_j \rangle = 0 \quad \text{for all } i \neq j.$$

Proposition 19. ☆ (1) • *"Orthogonal vectors $\{\vec{u}_1 \cdots \vec{u}_p\}$ are linear independent."*

• *Orthogonal vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$.* $\dim \mathbb{R}^n = n$

(1) • $x_1 \vec{u}_1 + \cdots + x_p \vec{u}_p = \vec{0}$

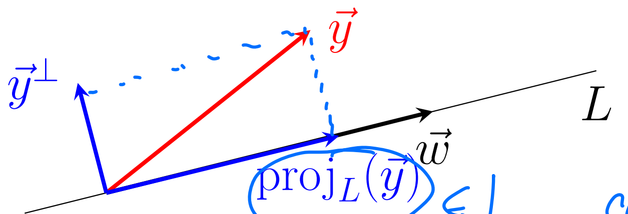
For any i: $\langle \vec{u}_i, x_1\vec{u}_1 + \cdots + x_p \vec{u}_p \rangle = 0$

$x_1 \langle \vec{u}_i, \vec{u}_1 \rangle + \cdots + x_p \langle \vec{u}_i, \vec{u}_p \rangle = 0 \Rightarrow x_i \langle \vec{u}_i, \vec{u}_i \rangle = 0 \Longrightarrow x_i = 0.$

Definition 20. • An **orthogonal basis** for a subspace W of an inner product space V is any basis for W which is also an orthogonal set. $\{\vec{u}_1 \cdots \vec{u}_p\}$

☆ • If each vector is a **unit** vector in an orthogonal basis, then it is an **orthonormal basis**. $\Longleftrightarrow$ $\langle u_i, u_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$ and $W = \mathrm{Span}\{\vec{u}_1 \cdots \vec{u}_p\}$

Let $L = \mathrm{Span}\{\vec{w}\}$ be the subspace in V spanned by $\vec{w} \in V$. For a given vector $\vec{y} \in V$, the **orthogonal projection of $\vec{y}$ onto** L



$\in L$ and $\boxed{\langle \vec{y} - \mathrm{proj}(\vec{y}), \vec{w} \rangle = 0}$ $\Leftarrow$ $\vec{y}^{\perp} \perp L.$

Let $\vec{w}$ be a nonzero vector in V. Any vector $\vec{y} \in V$ can be uniquely written as the sum of a scalar product of $\vec{w} \in V$ and a vector orthogonal to $\vec{w}$.

• $\mathrm{proj}_L \vec{y} = \dfrac{\langle \vec{y}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \cdot \vec{w} \quad \in L$

• Check $y^{\perp} := \vec{y} - \mathrm{proj}\, \vec{y}$

proj/corr.

Theorem 21 (Coordinates with respect to an orthogonal basis). *Let $\mathscr{B} = \{\vec{u}_1, \ldots, \vec{u}_p\}$ be an **orthogonal** basis for a subspace W of an inner product space V, and let $\vec{y}$ be any vector in W. Then*

$$\mathrm{proj}_W \vec{y} = \vec{y} = \left(\frac{\langle \vec{y}, \vec{u}_1\rangle}{\langle \vec{u}_1, \vec{u}_1\rangle}\right)\vec{u}_1 + \cdots + \left(\frac{\langle \vec{y}, \vec{u}_p\rangle}{\langle \vec{u}_p, \vec{u}_p\rangle}\right)\vec{u}_p$$

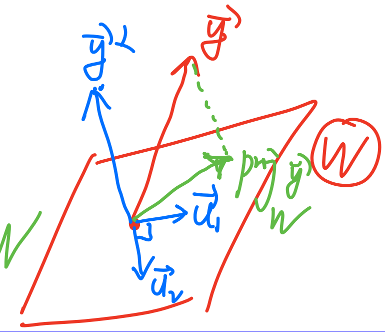
In particular, let $\mathscr{B} = \{\vec{u}_1, \ldots, \vec{u}_p\}$ be an **orthonormal** basis for a subspace W of $\mathbb{R}^n$, and let $\vec{y}$ be any vector in W. Then

$$\vec{y} = \langle \vec{y}, \vec{u}_1\rangle \vec{u}_1 + \cdots + \langle \vec{y}, \vec{u}_p\rangle \vec{u}_p$$

$\mathrm{proj}_W \vec{y} = c_1 \vec{u}_1 + \cdots + c_p \vec{u}_p$

- $\mathrm{proj}_W \vec{y} \in W$
- $\vec{y}^\perp := \vec{y} - \mathrm{proj}_W \vec{y} \perp W$

Check $\langle \vec{y} - \mathrm{proj}_W \vec{y}, \vec{u}_i \rangle = 0$



Theorem 22 (Orthogonal Decomposition). *Let W be any subspace of V and let $\vec{y} \in V$ be any vector. Then there exists a unique decomposition*

$$\vec{y} = \mathrm{proj}_W(\vec{y}) + \vec{y}^\perp$$

with $\mathrm{proj}_W(\vec{y}) \in W$ and $\vec{y}^\perp$ is perpendicular to W.

Theorem 23 (Orthogonal Decomposition). *If $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an* orthogonal basis *for W, then*

$$\mathrm{proj}_W(\vec{y}) = \left(\frac{\langle \vec{y}, \vec{u}_1\rangle}{\langle \vec{u}_1, \vec{u}_1\rangle}\right)\vec{u}_1 + \cdots + \left(\frac{\langle \vec{y}, \vec{u}_p\rangle}{\langle \vec{u}_p, \vec{u}_p\rangle}\right)\vec{u}_p$$

and $\vec{y}^\perp = \vec{y} - \mathrm{proj}_W(\vec{y})$.

$W \subset V$ subspace.

Thm: $V = W + U$

$W \subsetneq V$

Definition 24 (Orthogonal Complements). Given a nonempty **subset** (finite or infinite) W of V, its **orthogonal complement** $W^\perp$ is the set of all vectors $\vec{v} \in V$ orthogonal to W.

$W^\perp = \{$ all vectors in V "perpendicular" to $W \}$

$= \{ \vec{v} \in V \mid \langle \vec{v}, \vec{w} \rangle = 0$ for all $\vec{w} \in W \}$

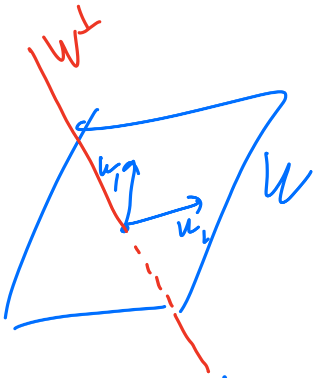
· Ex: $S = \left\{ \vec{w}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \vec{w}_2 = \begin{bmatrix}2\\3\\4\end{bmatrix} \right\} \subset \mathbb{R}^3 = V$

$\langle , \rangle$ = dot prod.

$A = \begin{bmatrix}1&2\\2&3\\3&4\end{bmatrix}$

$\mathrm{im}\begin{bmatrix}1&2\\2&3\\3&4\end{bmatrix} = W = \underline{\mathrm{Span}\{\vec{w}_1, \vec{w}_2\}} \subset V$

$W^\perp \stackrel{\cdot}{=} S^\perp = \left\{ \vec{v} \in \mathbb{R}^3 \;\middle|\; \begin{matrix} \langle \vec{v}, \vec{w}_1 \rangle = 0 \\ \langle \vec{v}, \vec{w}_2 \rangle = 0 \end{matrix} \right\}$ $\qquad \vec{v} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \in \mathbb{R}^3$



$= \underline{\left\{ \vec{v} \in \mathbb{R}^3 \;\middle|\; \begin{matrix} x_1 + 2x_2 + 3x_3 = 0 \\ 2x_1 + 3x_2 + 4x_3 = 0 \end{matrix} \right\}} = \ker \underbrace{\begin{bmatrix}1&2&3\\2&3&4\end{bmatrix}}_{A^T}$

$= \mathrm{Span}\left\{ \begin{bmatrix}1\\-1\\1\end{bmatrix} \right\}$

☆ Thm: $(\mathrm{im}\,A)^\perp = \ker(A^T)$

Ex: · If $\langle \vec{v}, \vec{w}_1 \rangle = 0$, $\langle \vec{v}, \vec{w}_2 \rangle = 0$, then

$\langle \underline{\vec{v}}, \underline{x_1 \vec{w}_1 + x_2 \vec{w}_2} \rangle = 0.$

$\{x_1\vec{b}_1 + \dots + x_p\vec{b}_p \mid \text{all } x_i\}$

$\{\vec{b}_1 \dots \vec{b}_p\}$ $\quad \| \quad \subset V$

Theorem 25. *Let S be a subset of V. Let $W = \mathrm{Span}(S)$, then*

(1) $S^\perp = W^\perp$ is a subspace of V.

(2) $(W^\perp)^\perp = W$

(3) $\dim W + \dim W^\perp = \dim V$

(4) $W \cap W^\perp = \{\vec{0}\}$

$\vec{u} \in W \quad \vec{u} \in W^\perp$

$\Rightarrow \langle \vec{u}, \vec{u} \rangle = 0 \Rightarrow \vec{u} = \vec{0}$

Theorem 26. *Let W be a subspace of V, then*

$$V = W \oplus W^\perp$$

$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$

$\mathrm{Row}\, A := \mathrm{im}\, A^T$

Theorem 27. *Let A be an $m \times n$ matrix, then*

$$(\mathrm{Row}\ A)^\perp = \ker(A) \quad \text{and} \quad (\mathrm{im}\ A)^\perp = \ker A^T. \quad = W^\perp$$

More over,

$$\mathbb{R}^m = \ker A^T \oplus \mathrm{im}\ A$$

$W = \mathrm{im}\, A$

$\mathbb{R}^m \xrightarrow{A^T} \mathbb{R}^n$

4. **Gram-Schmidt process and QR-factorization**

The **Gram-Schmidt process** is an algorithm that produces an orthogonal (or orthonormal) basis for any subspace W of V by starting with any basis for W.

Theorem 28 (Gram-Schmidt (Orthogonalize)). *Let W be a subspace of V and let $\vec{b}_1, \cdots, \vec{b}_p$ be a basis for W. Define vectors $\vec{v}_1, \ldots, \vec{v}_p$ as*

(1) $\vec{v}_1 = \vec{b}_1$

(2) $\vec{v}_2 = \vec{b}_2^{\perp} = \vec{b}_2 - \dfrac{\langle \vec{b}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1$

$\vdots$

(p) $\vec{v}_p = \vec{b}_p^{\perp} = \vec{b}_p - \dfrac{\langle \vec{b}_p, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 - \dfrac{\langle \vec{b}_p, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \vec{v}_2 - \cdots - \dfrac{\langle \vec{b}_p, \vec{v}_{p-1} \rangle}{\langle \vec{v}_{p-1}, \vec{v}_{p-1} \rangle} \vec{v}_{p-1}$

· Then $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$ is an orthogonal basis for W

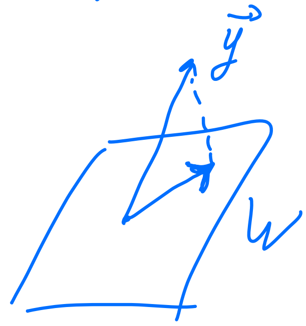
Theorem 29 (Gram-Schmidt (Normalize)). *If $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W, then*

$\left\{ \dfrac{\vec{v}_1}{\|\vec{v}_1\|}, \cdots, \dfrac{\vec{v}_p}{\|\vec{v}_p\|} \right\}$ is an orthonormal basis for W.

$$\text{Basis} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis} \xrightarrow{\text{normalize}} \text{Orthonormal basis.}$$

- <u>Inner product space</u> $V, \underline{+, \text{ scalar prod}, \langle \ , \ \rangle}$

- <u>orthogonal basis</u> / <u>orthonormal basis</u> → • $\{\vec{v}_1 \cdots \vec{v}_p\}$ for $W \subset V$
  - $\vec{y} \in V$



- ① how to use it ?

  ② how to find it ?

① $$\boxed{\text{proj}_W \vec{y} = \frac{\langle \vec{y}, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 + \cdots + \frac{\langle \vec{y}, \vec{v}_p \rangle}{\langle \vec{v}_p, \vec{v}_p \rangle} \vec{v}_p .}$$

② <u>Gram-Schmidt</u>

- $W^\perp$

$V = W \oplus W^\perp$

①

<u>Ex</u>: · <u>Inner product</u> on $C[0,1] = \{ f : [0,1] \to \mathbb{R} \text{ cont.} \}$

$$\langle f, g \rangle := \int_0^1 f(t) g(t)\, dt.$$

- $B = \{ \underset{\vec{u}_0(t)}{1}, \underset{\vec{u}_1(t)}{\cos 2\pi t}, \underset{\vec{v}_1(t)}{\sin 2\pi t}, \underset{\vec{u}_2(t)}{\cos 4\pi t}, \underset{\vec{v}_2(t)}{\sin 4\pi t}, \cdots \cos 2n\pi t, \ \sin 2n\pi t \cdots$

- <u>Thm</u>: B is <u>orthogonal</u>.

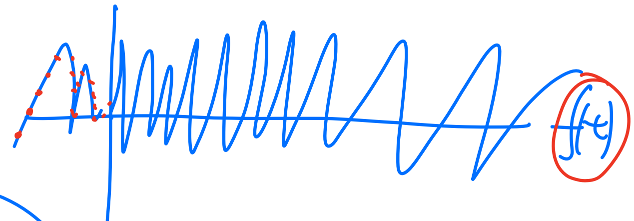
- $\boxed{\vec{u}_0(t)=1}$

$$\|\vec{u}_0(t)\| := \sqrt{\langle \vec{u}_0(t), \vec{u}_0(t)\rangle} = \sqrt{\int_0^1 1\,dt} = \sqrt{x\big|_0^1} = 1$$

$$\|\sin 2\pi nt\|^2 = \int_0^1 \sin 2\pi nt \, \sin 2\pi nt \, dt = \frac{1}{2}$$

$$\|\cos 2\pi nt\|^2 = \int_0^1 \cos 2\pi nt \, \cos 2\pi nt \, dt = \frac{1}{2}$$

- Def : $W = \operatorname{span} B$.
- B is an orthogonal basis for W.
- Fourier Transform



for any $f(t)$: $[0,1] \to \mathbb{R}$

$$f(t) = a_0 + a_1 \sin 2\pi t + b_1 \cos 2\pi t + \cdots$$

$\mathrm{proj}_W f(t)$

$$+ a_n \sin 2\pi n t + b_n \cos 2\pi n t + \cdots$$

$a_0 =$

$$a_n = \frac{\langle f(t), \sin 2\pi n t \rangle}{1/2} = 2 \int_0^1 f(t) \sin 2\pi n t \, dt.$$

$b_n =$

---

② $\boxed{\text{Gram-Schmidt}}$.

basis $\longrightarrow$ orthogonal basis $\longrightarrow$ orthonormal.

- $V = P_2 = \{ a_0 + a_1 t + a_2 t^2 \}$

- $\langle f, g \rangle := \int_0^1 f(t) g(t) \, dt.$

- $\mathcal{B} = \{ \underset{\vec{b}_1}{1}, \underset{\vec{b}_2}{t}, \underset{\vec{b}_3}{t^2} \}$ is a basis for P_2

- $\mathcal{B}$ is not orthogonal. $\langle 1, t \rangle = \int_0^1 t \, dt = \frac{t^2}{2}\Big|_0^1 = \frac{1}{2}$

Gram-Schmidt

$$\begin{cases} \vec{v}_1 = \vec{b}_1 = 1 \\ \vec{v}_2 = \vec{b}_2 - \dfrac{\langle \vec{b}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 = t - \dfrac{\int_0^1 t\, dt}{\int_0^1 1\, dt} 1 = t - \dfrac{1}{2} \\ \vec{v}_3 = \vec{b}_3 - \dfrac{\langle \vec{b}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 - \dfrac{\langle \vec{b}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \vec{v}_2 \end{cases}$$

$$= t^2 - \frac{1}{3} - \frac{1/12}{1/12}\left(t - \frac{1}{2}\right) = t^2 - t + \frac{1}{6}$$

- $\left\{ 1, \; t - \frac{1}{2}, \; t^2 - t + \frac{1}{6} \right\}$ is an orthogonal basis for P_2

- $$\begin{cases} \vec{u}_1 = 1 \\ \vec{u}_2 = \dfrac{t - \frac{1}{2}}{\left\| t - \frac{1}{2} \right\|} = \sqrt{12}\left(t - \frac{1}{2}\right) \\ \vec{u}_3 = \dfrac{\vec{v}_3}{\|\vec{v}_3\|} = \dfrac{t^2 - t - \frac{1}{6}}{\sqrt{\int_0^1 \left(t^2 - t - \frac{1}{6}\right)^2 dt}} = \sqrt{180}\left(t^2 - t - \frac{1}{6}\right) \end{cases}$$

- subspace of $\mathbb{R}^n$
- $\langle , \rangle$ is dot prod.

QR-Factorization.

$\downarrow$ Review of 2331.

QR-Factorization is the matrix version of Gram-Schmidt process for a subspace W of $\mathbb{F}^n$:

$$\text{Basis } \mathscr{B} = \{\vec{b}_1, \ldots, \vec{b}_p\} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis } \mathscr{V} = \{\vec{v}_1, \ldots, \vec{v}_p\}$$
$$\xrightarrow{\text{normalize}} \text{Orthonormal basis } \mathscr{U} = \{\vec{u}_1, \ldots, \vec{u}_p\}.$$

Theorem 30. *Given a $n \times p$ matrix $M = [\vec{b}_1 \ \ldots \ \vec{b}_p]$ with independent columns. There is a unique decomposition*

$$M = QR$$

when $n=p$ $\rightarrow$ $R = Q^{-1}M = Q^TM$

where, $Q = [\vec{u}_1, \ldots, \vec{u}_p]$ has orthonormal columns and R is an $p \times p$ upper triangular matrix with

$$r_{ii} = ||\vec{v}_i|| \text{ for } i = 1, \ldots, p \text{ and } r_{ij} = \langle \vec{u}_i, \vec{b}_j \rangle \text{ for } i < j.$$

Proof. Proof(for $p = 3$): From Gram-Schmidt process, write $\vec{b}_i$ as linear combinations of $\vec{u}_i$.

$$\vec{b}_1 = \vec{v}_1 = ||\vec{v}_1||\vec{u}_1$$

$$\vec{b}_2 = \vec{v}_2 + \frac{\langle \vec{b}_2, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 = ||\vec{v}_2||\vec{u}_2 + \langle \vec{b}_2, \vec{u}_1 \rangle \vec{u}_1$$

$$\vec{b}_3 = \vec{v}_3 + \frac{\langle \vec{b}_3, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\langle \vec{b}_3, \vec{v}_2 \rangle}{||\vec{v}_2||^2}\vec{v}_2 = ||\vec{v}_3||\vec{u}_3 + \langle \vec{b}_3, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{b}_3, \vec{u}_2 \rangle \vec{u}_2$$

So,

$$[\vec{b}_1 \ \vec{b}_2 \ \vec{b}_3] = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3] \begin{bmatrix} ||\vec{v}_1|| & \langle \vec{u}_1, \vec{b}_2 \rangle & \langle \vec{u}_1, \vec{b}_3 \rangle \\ 0 & ||\vec{v}_2|| & \langle \vec{u}_2, \vec{b}_3 \rangle \\ 0 & 0 & ||\vec{v}_3|| \end{bmatrix}$$

□

real

- An $n\times n$ "matrix" A is orthogonal

$\iff \|T_A(\vec{x})\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$ with dot prod.

$\iff \|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$

$\iff \langle A\vec{x}, A\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$ for any $\vec{x}, \vec{y} \in \mathbb{R}^n$

$\iff$ columns of A, $\{\vec{a}_1, \cdots, \vec{a}_n\}$ is orthonormal basis for $\mathbb{R}^n$

$\iff \vec{a}_i \cdot \vec{a}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

$[\ ][\,]\ \vec{a}_i^T \vec{a}_j =$

$\iff A^T A = I_n$

$\iff A^T = A^{-1}$

- $\{\vec{b}_1, \dots \vec{b}_p\}$ is an "orthonormal" set in $\mathbb{R}^n$ $(p \leq n)$

$\iff B = [\vec{b}_1 \cdots \vec{b}_p]$ $n\times p$ matrix satisfy

$$\boxed{B^T B = I_p}$$

$n \times p \quad p \times n$

Q: What is $\boxed{BB^T} = ?$

$n \times n$ matrix.

5. **Orthogonal Transformations and Orthogonal Matrices**

$\vec{x} \to A\vec{x}$

$\mathbb{R}^n \quad \mathbb{R}^n$

Let V be a inner product space.

Definition 31. A linear transformation $T : V \to V$ is called **orthogonal** if

$$||T(\vec{x})|| = ||\vec{x}|| \text{ for all } \vec{x} \in V$$

that is, T preserves the length of vectors.

Example 32. Whether or not the following transformations are orthogonal.

(1.) Rotations $S : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of rotation $S = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

(2.) Reflections $R : \mathbb{R}^2 \to \mathbb{R}^2$ are orthogonal transformations.

The matrix of reflection matrix $R = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ with $a^2 + b^2 = 1$ is orthogonal.

(3.) Orthogonal projections $P : \mathbb{R}^2 \to \mathbb{R}^2$ are NOT orthogonal transformations.

The matrix of an orthogonal transformation $T : \mathbb{F}^n \to \mathbb{F}^n$ is called an **orthogonal matrix**.

Theorem 33. *Let U be an $n \times n$ orthogonal matrix and let $\vec{x}$ and $\vec{y}$ be any vectors in $\mathbb{F}^n$. Then*
(1) $||U\vec{x}|| = ||\vec{x}||$.
(2) $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$.
(3) $\langle U\vec{x}, U\vec{y}\rangle = 0$ *if and only if* $\langle \vec{x}, \vec{y}\rangle = 0$.

Proposition 34. *U is an orthogonal matrix if and only if $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$ for any $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.*

Using the geometric meaning of the orthogonal transformation, we have

Theorem 35. *1. If A is orthogonal, then A is invertible and A^{-1} is orthogonal.*
2. If A and B are orthogonal, then AB is orthogonal.

Theorem 36. *The $n \times n$ matrix U is orthogonal if and only if $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal set.*

Application to real matrix A.

Recall the transpose of a matrix: Given an $m \times n$ matrix A, we define the ***transpose matrix*** A^T as the $n \times m$ matrix whose (i, j)-th entry is the (j, i)-th entry of A. The dot product can be written as matrix product

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$$

Theorem 37. *The $n \times n$ matrix A is orthogonal if and only if $A^T A = I_n$; if and only if $A^{-1} = A^T$.*

$U^T U = I_p$

$\Updownarrow$

Theorem 38. *Let W be any subspace of $\mathbb{R}^n$ with an orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$. Let $U = [\vec{u}_1 \ \vec{u}_2 \cdots \vec{u}_p]$. For any $\vec{y} \in \mathbb{R}^n$,*

$$\boxed{\mathrm{proj}_W(\vec{y})} = UU^T \vec{y}.$$

That is, the ***matrix of the projection*** *onto W is*

$$P = \boxed{UU^T}$$

$$\mathrm{proj}_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$\vec{y} \longrightarrow \mathrm{proj}_W \vec{y} = \underset{n \times n}{P} \ \underset{n \times 1}{\vec{y}}$$