Linear Algebra:

Vectors in $\mathbb{R}^n$

Matrices (operations)

linear system

- vector spaces
- Basis

$\begin{cases} \text{sum} \\ \text{scalar prod.} \end{cases}$

$A\vec{x} = \lambda \vec{x}$

Eigenvalue & Eigen Space

Diagonalization

$A = PDP^{-1}$

Jordan $A = PJP^{-1}$

$\begin{cases} \text{length.} \\ \text{angle.} \end{cases}$

Geometry of vector spaces

"inner product"

Dynamical System

$\vec{v}_t = A^t \vec{v}_0$

$A^t = PD^tP^{-1}$ or $A^t = PJ^tP^{-1}$

(**Perron-Frobenius** Thm

$A^\infty \rightsquigarrow A^{1000}$

symmetric matrices $A^T = A$

$A = PDP^{-1} = PDP^T$

Northeastern University, Department of Mathematics

SVD $\Rightarrow$ PCA

MATH 5110: Applied Linear Algebra and Matrix Analysis. (Fall 2022)

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

§6 Inner product spaces

Contents

1. Inner Product Spaces

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

$$\vec{u}\cdot\vec{v} := u_1v_1 + \cdots + u_nv_n$$

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

$$(\vec{u}, \vec{v}) \longrightarrow \vec{u}\cdot\vec{v}$$

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

$$\text{dist}(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|$$

angle between $\vec{u}$, $\vec{v}$

$$\cos\angle(\vec{u},\vec{v}) = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}$$

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

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

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

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

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

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

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

(1.) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.
(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**.

(1), (3) $\Rightarrow$ $\langle \vec{u}, c\vec{v} \rangle = \bar{c} \langle \vec{u}, \vec{v} \rangle$

Ex: In $\mathbb{C}^n$, the dot prod. is $\vec{u} \cdot \vec{v} = u_1 \bar{v}_1 + \cdots + u_n \bar{v}_n$

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

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

$$\langle \vec{u}, \vec{v} \rangle_W := \sum_{i=1}^{n} c_i u_i v_i$$

$$\boxed{:= \underset{1\times n}{\vec{u}^T} \underset{n\times n}{W} \underset{n\times 1}{\vec{v}}}$$

$$W = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{bmatrix}$$

More generally

for W · symmetric
· W positive definite (all eigenvalues >0)

$V =$

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

$\mathbb{C}$

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

$$V \times V \to \mathbb{R}$$

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

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

2. Norms (length)

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

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

- $||\vec{v}|| \geq 0.$ $\quad ||\vec{v}|| = 0 \Leftrightarrow \vec{v} = \vec{0}$
- Unit vector $\vec{u}$ if $||\vec{u}|| = 1$

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

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

$$||c \cdot \vec{v}||^2 = \langle c\vec{v}, c\vec{v} \rangle = c^2 \langle \vec{v}, \vec{v} \rangle = c^2 ||\vec{v}||^2$$

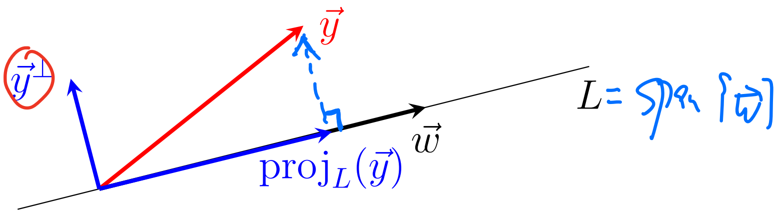
Theorem 8 (Pythagorean Theorem)**.** *If two vectors $\vec{u}, \vec{v} \in V$ are orthogonal, then they satisfy the **Pythagorean Relation***

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

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

$||$

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



Definition 9. Let $L = \text{Span}\{\vec{w}\}$ 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**

$$\begin{cases} \text{proj}_L(\vec{y}) := c\vec{w} \\ \langle \underbrace{y - \text{proj}_L \vec{y}}_{= \vec{y}^\perp}, \vec{w} \rangle = 0 \end{cases}$$

Proposition 10. *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 $\vec{w}$.*

$\langle \vec{y} - c\vec{w}, \vec{w} \rangle = 0$ $\Longrightarrow$ $c = \dfrac{\langle \vec{y}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle}$

$\langle y, \vec{w} \rangle - c\langle \vec{w}, \vec{w} \rangle = 0$

Theorem 11 (Cauchy-Schwarz inequality).

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

equality hold $\Leftrightarrow$ $\vec{y} = c\vec{x}$

$-1 \leq \dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} \leq 1$

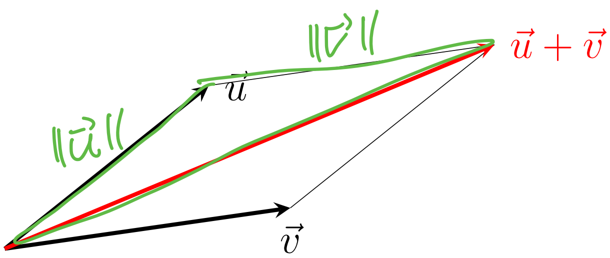
$\vec{y} = \text{proj}_{\vec{x}} \vec{y} + \vec{y}^\perp$

$$\|\vec{y}\|^2 = \langle \vec{y}, \vec{y} \rangle = \|proj_{\vec{x}}\vec{y}\|^2 + \|\vec{y}^{\perp}\|^2 \geq \|proj_{\vec{x}}\vec{y}\|^2$$

$$= \frac{\langle \vec{y}, \vec{x} \rangle^2}{\|\vec{x}\|^2}$$

Proposition 12 (Triangle Inequality). *Two vectors $\vec{u}, \vec{v} \in V$ satisfy*

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



Definition 13. (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

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

$\vec{x} \to \|\vec{x}\|$

$\|\ \| : V \to \mathbb{R}$

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

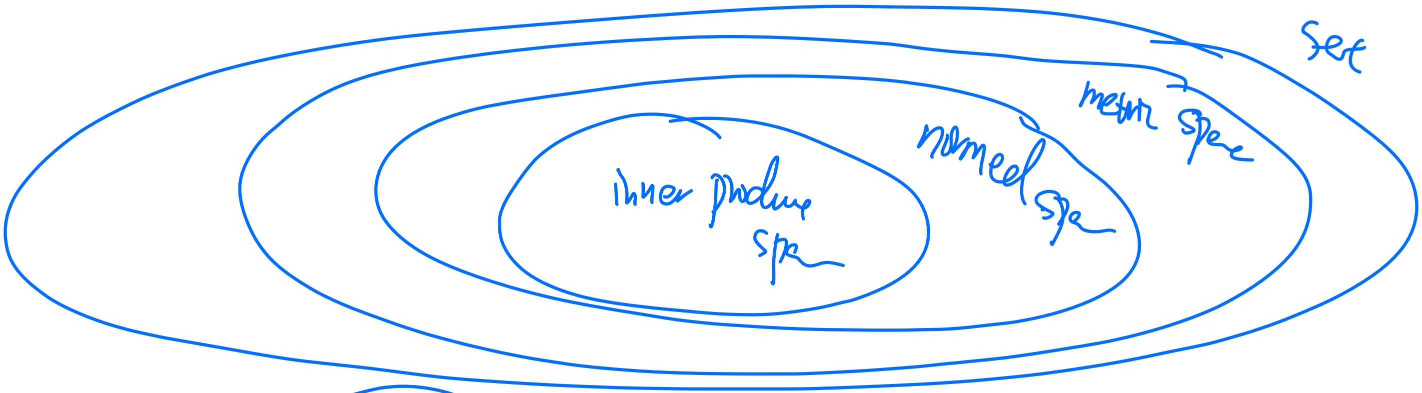
$\|\vec{x}\|$

More general

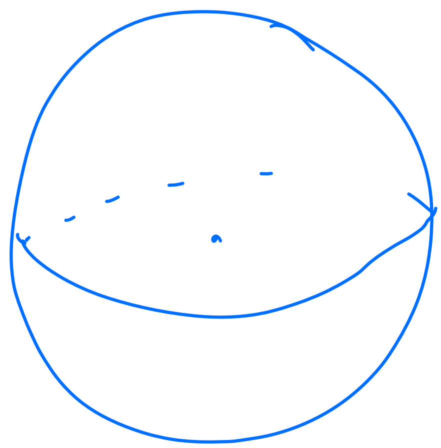
Definition 14. A **norm** on V is a map from V to $\mathbb{F}$ such that
(1) $||\vec{x}|| \geq 0$ for all $\vec{x} \in V$. $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}$.
(2) $||c\vec{x}|| = |c|||\vec{x}||$ for all $\vec{x} \in V$ and $c \in \mathbb{F}$.
(3) The triangle inequality $||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ holds for all vectors in V.

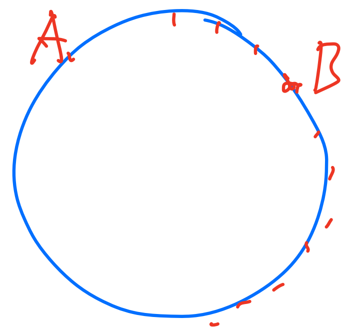
Ex: $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$

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



Ex: discrete metric $\quad \text{dist}(A,B) = \begin{cases} 0 & \text{if } A = B \\ 1 & \text{if } A \neq B \end{cases}$





Distance (metric) on a set S is a binary map

$$\text{dist}(\ ,\ ) : S \times S \to \mathbb{R}$$

such that
① $\text{dist}(A,B) = \text{dist}(B,A)$

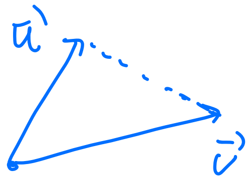
② $\text{dist}(A,B) = 0 \iff A = B$

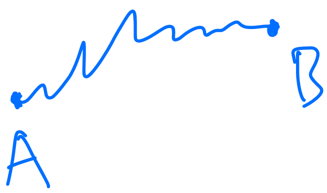
③ $\text{dist}(A,C) \leq \text{dist}(A,B) + \text{dist}(B,C)$

→ inner prod.

Definition 15 (Distance Between Vectors)**.** The **distance** $\mathrm{dist}(\vec{u},\vec{v})$ between vectors $\vec{u},\vec{v}\in\mathbb{R}^n$ is defined as

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





$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \vec{u} \in \mathbb{R}^n$$

Example 16. (l^p spaces) Let $1\le p<\infty$, it is natural to define l^p norms on $\mathbb{F}^n$ $\mathbb{R}^n$.

l^1 norm $\quad \|\vec{u}\|_{L_1} = |u_1|+|u_2|+\cdots+|u_n|$

$\|\vec{u}\|_{L_p} \neq$ any $\sqrt{\langle\vec{u},\vec{u}\rangle}$ for $p=1,3,4,\ldots$

l^2 norm $\quad \|\vec{u}\|_{L_2} := \sqrt{u_1^2+\cdots+u_n^2} = \sqrt{\vec{u}\cdot\vec{u}} = \|\vec{u}\|$

l^p norm: $\quad \|\vec{u}\|_{L_p} := \sqrt[p]{|u_1|^p+\cdots|u_n|^p}$

Example 17. (l^∞ spaces) It is natural to define l^∞ norms on $\mathbb{F}^n$ $\mathbb{R}^n$

$$\|\vec{x}\|_\infty = \max_{1\le i\le n}\{|x_i|\}$$

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

Example 18. (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

$$\|A\| := \sup\{\|A\vec{x}\| \mid \vec{x}\in\mathbb{R}^n \text{ and } \|\vec{x}\|=1\}$$

↑ smallest upperbound.

Example 19. Infinity norm on $\mathbb{F}^{m\times n}$.

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

$$\begin{bmatrix} |a_{11}|+|a_{12}| & \cdots & +|a_{1n}| \\ a_{21} & a_{22} \;\cdots & a_{2n} \\ & & \end{bmatrix}$$

Definition 20 (Orthogonal Set)**.** A set $\{\vec{u}_1, \ldots, \vec{u}_p\}$ of vectors in a inner vector space V is called **orthogonal** if

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

Proposition 21.
- *Orthogonal vectors are linear independent.*
- *Orthogonal vectors $\{\vec{u}_1, \ldots, \vec{u}_n\}$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$.*

$\dim V = n$

Definition 22.
- An **orthogonal basis** for a subspace W of an inner product space V is any basis for W which is also an orthogonal set.
- If each vector is a **unit** vector in an orthogonal basis, then it is an **orthonormal basis**.

Theorem 23 (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*

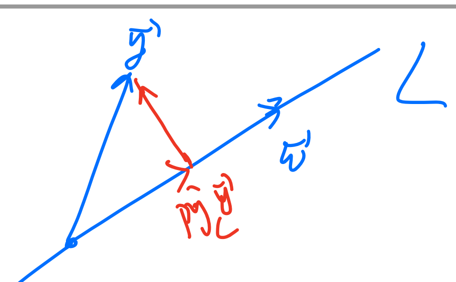
$$\operatorname{proj}_W \vec{y} = \vec{y} = \left(\frac{\langle \vec{y}, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \right) \vec{u}_1 + \cdots + \left(\frac{\langle \vec{y}, \vec{u}_p \rangle}{\langle \vec{u}_p, \vec{u}_p \rangle} \right) \vec{u}_p$$

If $W = \mathbb{R}^n$, then the $\mathscr{B}$-coordinates of $\vec{y}$ are given by:

$$[\vec{y}]_{\mathscr{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{with} \quad c_i = \frac{\langle \vec{y}, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} = \frac{\langle \vec{y}, \vec{u}_i \rangle}{||\vec{u}_i||^2}$$

*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*

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



$\operatorname{proj}_L \vec{y} = c\,\vec{w}$ projection

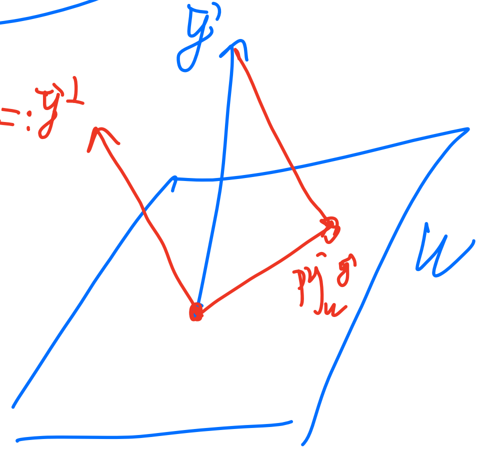
$\langle \vec{y} - \operatorname{proj}_L \vec{y}, \vec{w} \rangle = 0$ orthogonal

- V inner product space.
- $\vec{y} \in V$ and W sub space of V

basis of W is $\{\vec{u}_1 \cdots \vec{u}_p\}$

- $\vec{y} \notin W$

$\vec{y} - \text{proj}_W \vec{y} =: \vec{y}^{\perp}$



Def: $\boxed{\text{Orthogonal projection of } \vec{y} \text{ onto } W.}$

$$\begin{cases} \text{proj}_W \vec{y} = c_1 \vec{u}_1 + \cdots + c_p \vec{u}_p \\ \langle \vec{y}^{\perp}, W \rangle = 0 \end{cases} \iff \langle \vec{y}^{\perp}, \vec{u}_i \rangle = 0 \text{ for } i=1, \ldots, p$$

If $\vec{u}_1 \cdots \vec{u}_p$ orthogonal then

$$\begin{cases} \langle \vec{y} - (c_1\vec{u}_1 + \cdots + c_p\vec{u}_p), \vec{u}_1 \rangle = 0 & \Rightarrow c_1 = \_\_\_ \\ \\ \langle \vec{y} - (c_1\vec{u}_1 + \cdots + c_p\vec{u}_p), \vec{u}_p \rangle = 0 & \Rightarrow c_p = \_\_\_ \end{cases}$$

Theorem 24 (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 25 (Orthogonal Decomposition). *If $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an* orthogonal *basis for W, then*

$$\mathrm{proj}_W(\vec{y}) = \underbrace{\left(\frac{\langle \vec{y}, \vec{u}_1\rangle}{\langle \vec{u}_1, \vec{u}_1\rangle}\right)}_{= c_1} \vec{u}_1 + \cdots + \underbrace{\left(\frac{\langle \vec{y}, \vec{u}_p\rangle}{\langle \vec{u}_p, \vec{u}_p\rangle}\right)}_{= c_p} \vec{u}_p$$

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

inner product space

Definition 26 (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.

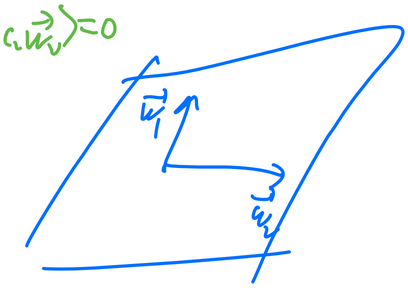
$W^{\perp} :=$ { all vectors in V orthogonal to W }

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

Ex: $W = \left\{ \vec{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \ \vec{w}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right\} \subset \mathbb{R}^3$ $\quad A = [\vec{w}_1 \ \vec{w}_2] = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$

$\mathrm{im} A = S := \mathrm{Span}(W)$

$S^{\perp} = W^{\perp} = \left\{ \vec{v} \in V \;\middle|\; \begin{matrix} \langle \vec{v}, \vec{w}_1 \rangle = 0 \\ \langle \vec{v}, \vec{w}_2 \rangle = 0 \end{matrix} \right\} \to \langle \vec{v}, c_1\vec{w}_1 + c_2\vec{w}_2 \rangle = 0$



$= \left\{ \vec{v} \in V \;\middle|\; \begin{matrix} v_1 + 2v_2 + 3v_3 = 0 \\ 2v_1 + 3v_2 + 4v_3 = 0 \end{matrix} \right\}$

$= \ker\left(\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}\right) = \boxed{\ker(A^T) = (\mathrm{im} A)^{\perp}}$

Theorem 27. *Let S be a subset of V. Let $W = \mathrm{Span}(S)$, then*

(1) $W^\perp = S^\perp$

(2) $S^\perp$ is a subspace of V.

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

H.W. ☆ (4) $\dim(W) + \dim(W^\perp) = \dim V$.

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

Theorem 28. *Let W be a subspace of V, then*

$V = W + W^\perp$

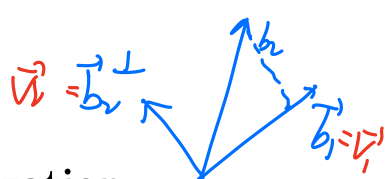
$$V = W \oplus W^\perp$$

Theorem 29. *Let A be an $m \times n$ matrix, then*

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

More over,

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



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 30 (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*

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

$$\vec{v}_2 = \vec{b}_2^{\perp} = \vec{b}_2 - \text{proj}_{\vec{v}_1} \vec{b}_2 = \vec{b}_2 - \frac{\langle \vec{b}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1$$

$$\vec{v}_3 = \vec{b}_3^{\perp} = \vec{b}_3 - \text{proj}_{\vec{v}_1, \vec{v}_2} \vec{b}_3 = \vec{b}_3 - \frac{\langle \vec{b}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 - \frac{\langle \vec{b}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \vec{v}_2$$

$$\vdots$$

Then $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W.

Theorem 31 (Gram-Schmidt (Normalize)). *If $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W, then*

$$\left\{ \vec{w}_i = \frac{\vec{v}_i}{\|\vec{v}_i\|} \quad \text{for } i = 1, \ldots, p \right\} \text{ is orthonormal basis.}$$

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

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

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

Given a vector in W, let's compare their coordinates:

$$\begin{tikzcd}
{[\vec{x}]_{\mathscr{B}}} \arrow[dr, "M"] \arrow[dd, "R"'] & \\
& \vec{x} \\
{[\vec{x}]_{\mathscr{U}}} \arrow[ur, "Q"'] &
\end{tikzcd}$$

Each matrix defines an isomorphism. So, $M = QR$.

Here $M = [\vec{b}_1 \ \dots \ \vec{b}_p]$ and $Q = [\vec{u}_1, \dots, \vec{u}_p]$.

$n \times p$ $\quad$ $\in \mathbb{R}^n$

Theorem 32. *Given a $n \times p$ matrix $M = [\vec{b}_1 \ \dots \ \vec{b}_p]$ with independent columns. There is a unique decomposition*

$$M = QR$$

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

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

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

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

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

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

$\vec{u}_i = \frac{\vec{v}_i}{||\vec{v}_i||}$

So,

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

$Q^T = \begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vec{u}_3^T \end{bmatrix}$ □

$\underset{n\times p}{M} = \underset{n\times p}{Q}\ \underset{p\times p}{R}$

$\text{rank } M = p$ $\qquad Q^T Q = I_p$ $\qquad QQ^T \neq I_n$

5. Orthogonal Transformations and Orthogonal Matrices

$n \times n$

Let V be a inner product space. over $\mathbb{F} = \mathbb{R}$

Definition 33. 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 34. 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 $T_u : \mathbb{R}^n \to \mathbb{R}^n$ is called an **orthogonal matrix**.

$\vec{x} \to U\vec{x}$

$||T(\vec{x})|| = ||\vec{x}||$

Theorem 35. *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) $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$. $\Longrightarrow$ Thu presene the angle.
(3) $\langle U\vec{x}, U\vec{y}\rangle = 0$ *if and only if* $\langle \vec{x}, \vec{y}\rangle = 0$.

Proposition 36. *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 37. *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 38. *The $n \times n$ matrix U is orthogonal if and only if $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal set.*

Thm: U is orthogonal $\iff ||U\vec{x}|| = ||\vec{x}||$ for all $\vec{x}$

$\iff \langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$

$\{\vec{u}_1 \cdots \vec{u}_n\}$ orthonormal $\Longleftrightarrow$ $\vec{u}_i \cdot \vec{u}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

$\vec{u}_i^T \vec{u}_j$

$\Longleftrightarrow U^T U = I_n$

$$\begin{bmatrix} u_1 \cdot u_1 & u_1 \cdot u_2 & \cdots \\ u_2 \cdot u_1 & u_2 \cdot u_2 & \cdots \\ \vdots & & \end{bmatrix}$$

$\Longleftrightarrow U^{-1} = U^T$

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

$\{\vec{u}_1 \cdots \vec{u}_p\}$ orthonormal $\Longleftrightarrow$ $U^T U = I_p$

$\underbrace{U U^T}_{n \times n} \neq I_n$

$\mathbb{R}^n \xrightarrow{UU^T} \mathbb{R}^n$

Theorem 40. *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 $\vec{y} \in \mathbb{R}^n$,*

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

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

$$P = UU^T$$

Ex: $W = \{a_0 + a_1 t + a_2 t^2\} = \text{Span}\{1, t, t^2\}$

- inner product $\langle f(t), g(t)\rangle = \int_0^1 f(t)\, g(t)\, dt$.
- $\{1, t, t^2\}$ is a basis for W; $1 = b_1$, $t = b_2$, $t^2 = b_3$

Q: Gram-Schmidt process find orthogonal basis for W

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

$$\vec{v}_2 = \vec{b}_2^{\perp} = \vec{b}_2 - \operatorname{proj}_{\vec{v}_1} \vec{b}_2 = t - \frac{\int_0^1 t\, dt}{\int_0^1 1\, dt} \cdot 1 = t - \frac{1}{2}$$

$$\vec{v}_3 = \vec{b}_3^{\perp} = \vec{b}_3 - \frac{\langle \vec{b}_3, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{v}_1 - \frac{\langle \vec{b}_3, \vec{v}_2\rangle}{\langle \vec{v}_2, \vec{v}_2\rangle}\vec{v}_2 = t^2 - \frac{\int_0^1 t^2\, dt}{\int_0^1 1\, dt} 1 - \frac{\int_0^1 t^2(t-\frac{1}{2})\, dt}{\int_0^1 (t-\frac{1}{2})^2 dt}(t - \tfrac{1}{2})$$

$$= t^2 - t + \frac{1}{6}$$

$$\vec{u}_1 = 1$$

$$\vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \sqrt{12}\left(t - \frac{1}{2}\right) \qquad \vec{u}_3 = \sqrt{180}\left(t^2 - t + \frac{1}{6}\right)$$

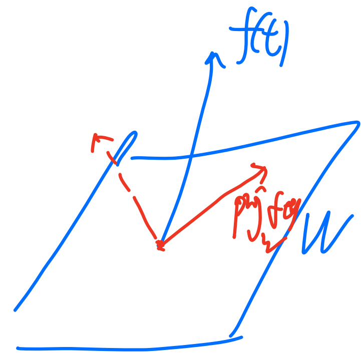
orthonormal basis for W.

$\|f(t)\|$ $\qquad$ $\text{dist}(f(t), g(t)) = \|f(t) - g(t)\|$

Ex: $\cdot\ V = \{\text{all cont. functions } \mathbb{R} \xrightarrow{f} \mathbb{R}\}$ $\quad \langle f(t), g(t)\rangle = \int_0^1 f\cdot g\, dt$

$\cdot\ W = \mathrm{Span}\{1, t, t^2\}$

$\cdot\ f(t) = e^t \in V \qquad f(t) \notin W.$



Q $\cdot\ \mathrm{proj}_W f(t) = ?$

Method 1

Theorem 25 (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 $\vec{y}^\perp = \vec{y} - \mathrm{proj}_W(\vec{y})$.

Method 2. $\cdot\ \{1, t, t^2\}$ is a basis for W.

$\cdot\ \mathrm{proj}_W f(t) \in W \implies \mathrm{proj}_W f(t) = c_0 + c_1 t + c_2 t^2$

$\cdot\ f(t) - \mathrm{proj}_W f(t) \perp W$

$$\Longleftrightarrow \begin{cases} \int_0^1 (c_0 + c_1 t + c_2 t^2 - e^t)(1)\,dt = 0 \\ \int_0^1 (c_0 + c_1 t + c_2 t^2 - e^t)(t)\,dt = 0 \\ \int_0^1 (c_0 + c_1 t + c_2 t^2 - e^t)(t^2)\,dt = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} c_0 + \frac{1}{2}c_1 + \frac{1}{3}c_2 - (e-1) = 0 \\ \frac{1}{2}c_0 + \frac{1}{3}c_1 + \frac{1}{4}c_2 - 1 = 0 \\ \frac{1}{3}c_0 + \frac{1}{4}c_1 + \frac{1}{5}c_2 - (e-2) = 0 \end{cases} \Longleftrightarrow A\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} e-1 \\ 1 \\ e-2 \end{bmatrix}$$

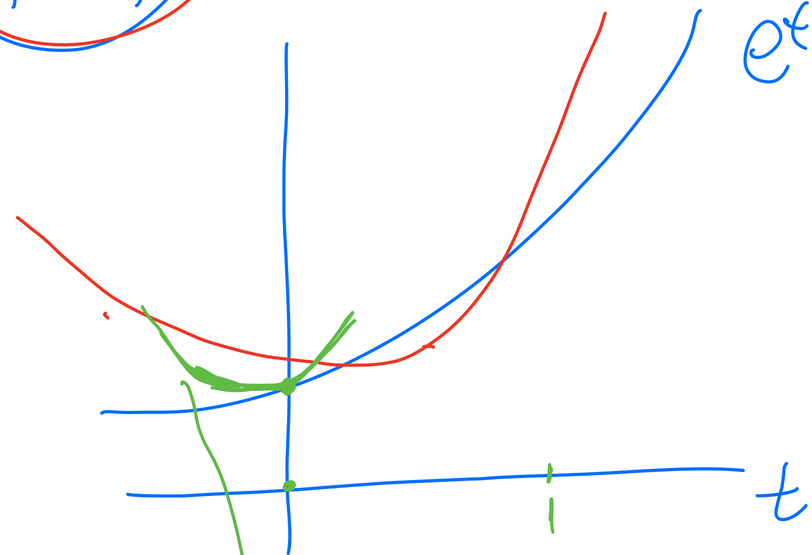
$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1.013 \\ 0.8511 \\ 0.8392 \end{bmatrix}$$

$$\Longrightarrow \operatorname{proj}_{\mathrm{Span}\{1,t,t^2\}} e^t = \boxed{1.013 + 0.8511\, t + 0.8392\, t^2} \in W$$

e^t ~

$e^t = f(t)$



$\operatorname{proj}_W e^t$ is the "closest" vector in W to e^t.

$$\boxed{e^t \approx 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots}$$

Application to complex matrix A. $\mathbb{C}$

The **conjugate transpose** of a complex matrix A is $A^* := \overline{A}^T$

A complex matrix A is called **Hermiltian** if $A^* = A$. sym

---

Def: A matrix $U \in \mathbb{C}^{n\times n}$ is call "unitary" ($U = [\vec{u}_1 \cdots \vec{u}_n]$)

$\iff \|U\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{C}^n$ where $\|\vec{x}\| = \sqrt{\vec{x}^T \overline{\vec{x}}}$

thm, $\iff \langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$ where $\langle \vec{x}, \vec{y}\rangle := \vec{x}^T \overline{\vec{y}}$

$\iff \vec{u}_1 \cdots \vec{u}_n$ are orthonormal.

$\iff U^T \overline{U} = I_n \iff \underbrace{\overline{U}^T}_{U^*} U = I_n$

$\iff U^{-1} = U^*$

---

prop: real matrix A. $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $\vec{v} \to A\vec{v}$

$$(A\vec{x}) \cdot \vec{y} = \vec{x} \cdot (A^T \vec{y})$$

prop complex matrix A $T_A : \mathbb{C}^n \to \mathbb{C}^m$

$$(A\vec{x}) \cdot \vec{y} = \vec{x} \cdot (\overline{A^T} \vec{y})$$

$\langle A\vec{x}, \vec{y}\rangle = \langle \vec{x}, \overline{A^T}\vec{y}\rangle$

- $(A\vec{x}) \cdot \vec{y} = (A\vec{x})^T \overline{\vec{y}} = (\vec{x}^T A^T) \overline{\vec{y}}$

dot over $\mathbb{C}$

$= \vec{x}^T (A^T \overline{\vec{y}})$

$= \vec{x}^T \overline{\overline{A^T} \vec{y}}$

$= \vec{x} \cdot (\overline{A^T} \vec{y})$

$c A$
1×1 $n\times1$

6. The adjoint of a linear operator

An generalization of transpose of a real matrix and conjugate transpose of a complex is the joint operator.

Definition 41. Let $T : V \to W$ be a linear map between two inner product spaces. A a **joint operator** of T is a linear map $T^* : W \to V$ such that

$$\langle T(\vec{v}), \vec{w} \rangle_W = \langle \vec{v}, T^*(\vec{w}) \rangle_V.$$

Theorem 42. *Let $T : V \to W$ be a linear map between two inner product spaces. Then T has a unique a joint operator T^*.*

Theorem 43. *Let $T : V \to W$ be a invertible linear map between two inner product spaces. Then its joint operator T^* is also invertible and $(T^*)^{-1} = (T^{-1})^*$.*

$(A^T)^{-1} = (A^{-1})^T$ $\quad$ $(A^*)^{-1} = (A^{-1})^*$