Lecture 17

Orthonormal Basis and Gram-Schmidt Process

Instructor: Ruoyu Sun

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
数据科学学院
School of Data Science

Today's Lecture: Outline

Main topic: Orthonormal basis

1. Orthonormal basis
2. Gram-Schmidt Process
3. Orthogonal matrix

Strang's book: mainly Sec 4.4; a bit of Sec 4.2

Today's Lecture: Learning Goals

After the lecture, you should be able to

1. compute an orthonormal basis from a basis of a subspace
2. compute the projection onto a subspace
3. tell the definition and properties of orthogonal matrix

Review

Recall: Orthogonality and LS Solution

Definition in Lec 16 (Orthogonality)

Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are orthogonal if $\mathbf{u}^\top \mathbf{v} = 0$. Denote $\mathbf{u} \perp \mathbf{v}$

Thm 16.2 and Lemma 16.1 together:

Consider a least squares problem. The following statements are equivalent:

1. $\mathbf{y}$ minimizes $||A\mathbf{x} - \mathbf{b}||$

2. $\mathbf{b} - A\mathbf{y} \perp S$.

3. $A^\top A\mathbf{y} = A^\top \mathbf{b}$

Part I Orthonormal Basis and Orthogonal Matrix

$u \perp v$: relation of two vectors

Orthonormal Set of Vectors

Definition 17.1 (Orthogonal Set)
Let $\{v_1, v_2, \ldots, v_k\}$ be a set of nonzero vectors in $\mathbb{R}^n$. If $\langle v_i, v_j \rangle = 0$ for any $i \neq j$, then this set is called an **orthogonal set.** pairwise relation

Examples: $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$, $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix} \right\}$.

$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix} \right\}$

Definition 17.2 (Orthonormal Set)
An orthonormal set $\{v_1, v_2, \ldots, v_k\}$ is an orthogonal set of unit-norm vectors, i.e., $\|v_i\| = 1$ for all $i = 1, \ldots, n$

Examples: $\left\{ \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \right\}$. $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, $\left\{ \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}, \begin{bmatrix} -3/\sqrt{13} \\ 2/\sqrt{13} \end{bmatrix} \right\}$.

From an Orthogonal Set to an Orthonormal Set

Question

Can you create an orthonormal set based on an orthogonal set?

$\{u_1, u_2, \ldots, u_n\}$, orthogonal set

$\downarrow$

$\left\{ \frac{u_1}{\|u_1\|}, \ldots, \frac{u_n}{\|u_n\|} \right\}$ orthonormal set

Examples

Example

Consider $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ in $\mathbb{R}^3$

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & 2/\sqrt{14} & 4/\sqrt{42} \\ \frac{1}{\sqrt{3}} & 1/\sqrt{14} & -5/\sqrt{42} \\ \frac{1}{\sqrt{3}} & -3/\sqrt{14} & 1/\sqrt{42} \end{bmatrix}$$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix}$$

then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, the orthonormal set is

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

Is an orthonormal basis unique for a given inner product space?

What is Special about an Orthogonal Set?

Knowledge new

Next, we check the relation of orthogonal set and other notions:

Orthogonal v.s. independent.

Orthogonal v.s. basis.

Question If $\{v_1, v_2, \ldots, v_k\}$ is an orthogonal set, any "redundancy" ?

span$(v_1, \ldots, v_k)$ by fewer vectors

Let students vote.

Yes fewer vec. can span

No, fewer vec cannot span

Proof of Proposition 17.1

Proposition17.1

Let $S = \{v_1, \ldots, v_k\}$ be an orthogonal set. Then $v_1, \ldots, v_k$ are linearly independent.

Proof for $k=2$:

skip proof in class

Exercise

Remark: Orthogonal: pairwise relation, (v_i, v_j)

Indep: all-vec (mutual) relation, $v_1, \ldots, v_n$.

Proof for $k=2$: $v_1 \perp v_2$, then $\alpha_1 v_1 + \alpha_2 v_2 = 0$, ① iff $\alpha_1 = \alpha_2 = 0$ ②

Proof: ① $\Rightarrow 0 = v_1^T(\alpha_1 v_1 + \alpha_2 v_2)$

$= \alpha_1 v_1^T v_1 + \underbrace{\cancel{\alpha_2 v_1^T v_2}}_{=0}$.

$= \alpha_1 \|v_1\|^2$

$\Rightarrow \alpha_1 = 0$ or $\|v_1\| = 0$.

$\xRightarrow{\text{def: } v_1 \neq 0} \alpha_1 = 0$.

Similarly, $\alpha_2 = 0$. Then ① $\Rightarrow \alpha_1 = \alpha_2 = 0$. Thus v_1, v_2 are indep. $\square$

Orthonormal Basis (标准正交基)

Definition 17.3

A set of vectors $S = \{v_1, \ldots, v_n\}$ is called an **orthonormal basis** of a linear space V if

(1) S is an orthonormal set

(2) Vectors in S form a basis of V.

Example

(1). Standard basis $e_1, \cdots, e_n$, $e_i \perp e_j$, $\|e_i\|$, form basis

(2). $\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^3$

Benefit of Orthogonal: Help Verify Basis.

Q: Prove $\{e_1, e_2, \cdots e_n\}$ is basis of $\mathbb{R}^n$.

Method 1. (i) Indep. or span $\mathbb{R}^n$. (ii) n vectors.

need linear systems

Method 2. (i) Orthogonal. $e_i \perp e_j$, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$\Rightarrow$ indep.

(ii) n vec.

$\rightarrow$ basis

no need to solve linear system to check span & indep.

Just need to compute inner product.

Orthogonal Matrix

Definition 17.4 (Orthogonal Matrix) Not orthogonal matrix

An orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ is a real square matrix whose columns form an orthonormal basis in $\mathbb{R}^n$ $Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n \times n}$

Examples:

$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $Q = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}$, q_i's form orthonormal set.

n vector $\rightarrow$ basis.

orthonormal $\Rightarrow$ indep. $\rightarrow$ basis

Claim: If dim(V) = n, then an orthonormal set $S = \{v_1, \ldots, v_n\}$ is an orthonormal basis. [based on what result of past lectures?]

Corollary: If the columns of $Q \in \mathbb{R}^{n \times n}$ form an orthonormal set of $\mathbb{R}^n$, then it is an orthogonal matrix. no need to check span.

Orthogonal Matrix

Proposition 17.2 (Orthogonal Matrix)

A matrix $Q \in \mathbb{R}^{n\times n}$ is an orthogonal matrix if and only if

$$Q^\top Q = I_n .$$

$Q^{-1} = Q^\top$.

Example For any θ

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad Q^{-1} = Q^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Proof of Proposition 17.4

Express $A^{\top}A = I_n$ row.

Proposition 17.2 (Orthogonal Matrix)
A matrix $Q \in \mathbb{R}^{n\times n}$ is an orthogonal matrix if and only if

$$Q^{\top}Q = I_n.$$

Proof: (exercise)

$$Q = [\vec{q}_1, \dots, \vec{q}_n]$$

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I_n = Q^{\top}Q = \begin{bmatrix} \vec{q}_1^{\top}\vec{q}_1 & \cdots & \vec{q}_1^{\top}\vec{q}_n \\ \vdots & & \\ \vec{q}_n^{\top}\vec{q}_1 & \cdots & \vec{q}_n^{\top}\vec{q}_n \end{bmatrix}$$

$$\Rightarrow \begin{cases} 1 = q_i^{\top}q_i = \|q_i\|^2, \ \forall i & \text{[unit norm]} \\ 0 = q_i^{\top}q_j, \ \forall i \neq j & \text{[orthogonal pair]} \end{cases}$$

Midterm exam:
Express $Q^{\top}Q = I_n$ using row-form.

$$Q = [\,|\ |\ \cdots\ |\,] = \begin{bmatrix} \equiv \end{bmatrix}$$

$$Q^{\top}Q = \sum \square\,\square^{\top}$$

sum of outer products

$$= \begin{bmatrix} \square\,\square & \cdots & \square\,\square \\ \vdots & & \\ & \cdots & \end{bmatrix}$$

collection of inner products

$\boxed{\text{Lemma}}$ Left inverse $\Rightarrow$ inverse.

If $\underbrace{A B}_{\text{square}} = I_n$, then $BA = I_n$

and $A^{-1} = B$.

Proof: (Long).

Properties of Orthogonal Matrix

Property 17.1 (Orthogonal Matrix)

If $Q \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, then

(a) The column vectors of Q form an orthonormal basis.

(b) $Q^\top Q = I_n$.

(c) $Q^{-1} = Q^\top$. $QQ^\top = I_n$

(d) $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. [inner product preserving] 保内积、

(e) $\|Q\mathbf{x}\| = \|\mathbf{x}\|$. let $y=x$ [norm preserving] 保范数、

$$\langle Qx, Qy\rangle = (Qy)^T Qx = y^T(Q^TQ)x = y^Tx = \langle y, x\rangle.$$

$$\|Qx\|^2 = x^TQ^TQx = x^Tx = \|x\|^2.$$

These will be very useful for eigenvalues after Lec 21.

Part II Representation and Projection

Recall: Representing Element using Basis

Proposition (Unique Representation via Basis)
Let $\mathcal{U} = \{u_1, \ldots, u_n\}$ be a basis of a linear space V.
Any element $v \in V$ can be **uniquely** represented as a linear combination of $u_1, \ldots, u_n$ [discussed in Lec 15]

Representation $v = \alpha_1 u_1 + \ldots + \alpha_n u_n$

We know such $\alpha_1, \ldots, \alpha_n$ exist. But....how to compute?

Well, if $u_i \in \mathbb{R}^n$, just solve linear system $U \cdot \vec{\alpha} = \vec{v}$

variable

Next, we show, if $\{u_i\text{'s}\}$ is orthonormal basis,

then express MUCH EASIER

Representing Element using Basis

Theorem 17.1 (Representation via Orthonormal Basis)
Let $\mathscr{U} = \{u_1, \ldots, u_n\}$ be an **orthonormal basis** of $\mathbb{R}^n$.
Any vector v can be represented as a linear combination

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

Proof: In the expression $\vec{v} = \sum_i \alpha_i \vec{u}_i$ (*)

coeff $\alpha_i = v^T u_i$, $i = 1, \ldots, n$.

Hint: Two ways.
Via vector.
Via matrix.

Method 1 (vec) We know that $\{\alpha_i\}$ exist since $\{u_i\}$'s form a basis.

Next compute α_i's. $u_k^T \vec{v} = \vec{u}_k^T (\sum_i \alpha_i \vec{u}_i) = \sum_i \alpha_i \vec{u}_k^T \vec{u}_i = \alpha_k \|u_k\|^2 = \alpha_k$.

Method 2. (matrix) (*) $\Leftrightarrow \vec{v} = U \cdot \vec{\alpha} \Rightarrow \vec{\alpha} = U^{-1} \vec{v} \overset{\text{Prop 17.1}}{=\!=} U^T \vec{v} = \begin{bmatrix} \vec{u}_1^T \vec{v} \\ \vdots \\ \vec{u}_n^T \vec{v} \end{bmatrix}$

Method 3. (matrix). $\Big[\underbrace{\sum_{i=1}^{n} (u_i u_i^T)}_{= I_n}\Big] v = U \cdot U^T \cdot v = v$.

Example of Representing using Orthonormal Basis

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

Example

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{\sqrt{42}} \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\}$$

is an orthonormal basis of $\mathbb{R}^3$

For any $\mathbf{x} = [x, y, z]^T \in \mathbb{R}^3$, one has

$$\begin{aligned} \mathbf{x} &= < \mathbf{x}, \mathbf{v}_1 > \mathbf{v}_1 + < \mathbf{x}, \mathbf{v}_2 > \mathbf{v}_2 + < \mathbf{x}, \mathbf{v}_3 > \mathbf{v}_3 \\ &= \frac{x + y + z}{\sqrt{3}} \mathbf{v}_1 + \frac{2x + y - 3z}{\sqrt{14}} \mathbf{v}_2 + \frac{4x - 5y + z}{\sqrt{42}} \mathbf{v}_3 \end{aligned}$$

Orthonormal Basis is Useful

Because

1. Representing any vector by directly computing the "coordinates"

—Useful in, e.g., computing LS-solution (forthcoming)

Projection

2. Useful in eigen-theory (lecture 21)

Central part

Projection

Definition 16.1 [Last lecture]

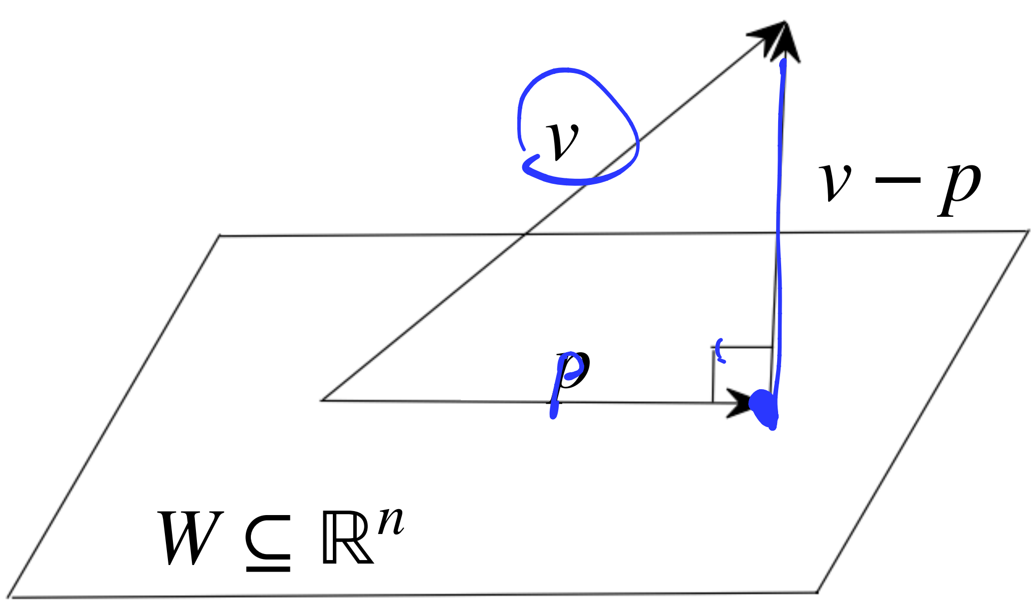
Suppose S is a subspace of $\mathbb{R}^m$.

Suppose $\mathbf{p} \in S$ and $\mathbf{b} - \mathbf{p} \perp S$, then we say $\mathbf{p}$ is the projection of $\mathbf{b}$ onto S.

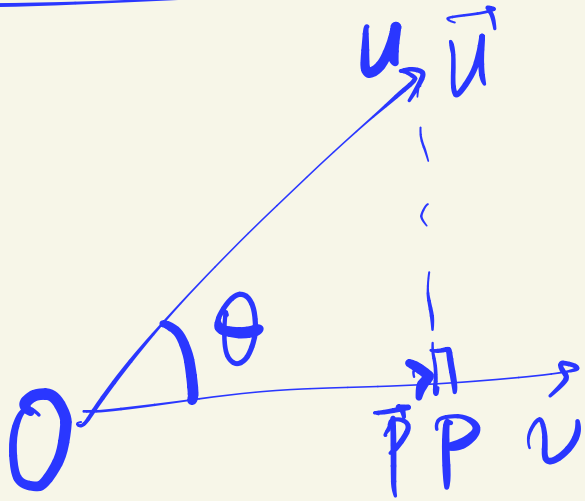
Proposition 17.2

For any subspace $W \subset \mathbb{R}^n$, any vector $\mathbf{v} \in \mathbb{R}^n$, there is a unique vector $\mathbf{p} \in W$ such that $(v - p) \perp W$.

[skip proof]



One-dim subspace



$\vec{p} \triangleq$ Proj of $\vec{u}$ onto $\vec{v}$

span$(\{\vec{v}\}) \triangleq V$

$= \text{Proj}_V(\vec{u})$

Direction of $\vec{p}$? same as $\vec{v}$

$\vec{p} = \|\text{length}\| \cdot \frac{\vec{v}}{\|\vec{v}\|}$ ①

Length of $\vec{p}$? $|OP| = |OU| \cdot \cos\theta$,

$\|\vec{p}\| = \cancel{\|\vec{u}\|} \cdot \frac{\langle \vec{u}, \vec{v} \rangle}{\cancel{\|\vec{u}\|} \cdot \|\vec{v}\|} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|}$ ②

$\vec{p} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v}$

Claim If $\|\vec{v}\| = 1$, then Proj of $\vec{u}$ onto span$(\{\vec{v}\})$ is

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

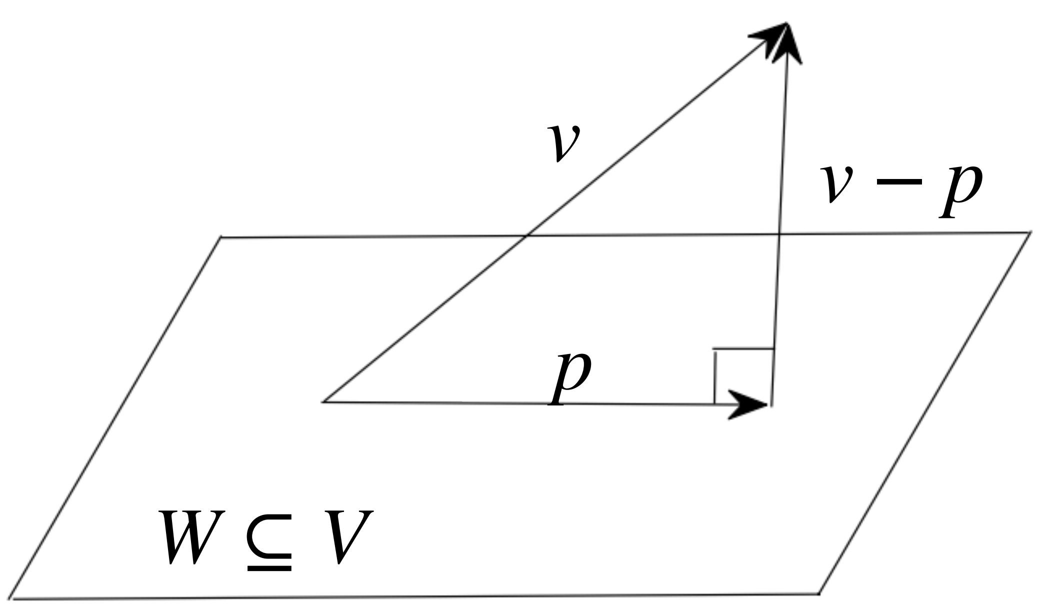
Representing Projection via Orthonormal Basis

Theorem 17.2 (Projection Representation) if you have orthonormal basis.

Let W be a subspace of $\mathbb{R}^n$ and suppose $S = \{w_1, \ldots, w_k\}$ is an **orthonormal basis** of W. The projection of $\mathbf{v}$ onto W can be represented as

$$p = \sum_{i=1}^{k} \langle w_i, v \rangle w_i \quad = \sum_i w_i w_i^T v$$

$$= Z Z^T v, \text{ where } Z = [w_1, \ldots, w_n]$$



Skip proof for general case.

Proof of Representing Projection

Le $S = \{w_1, \ldots, w_k\}$ is an **orthonormal basis** of W.

Prove $p = \sum_{i=1}^{k} \langle w_i, v \rangle w_i$

~~Proof:~~

Hint:
Try k=1.
Then k=2; n=3.

① If you know orthonormal basis (of subspace)

then you know how to compute projection

Part III Gram-Schmidt Process

Key Question in this Part

Suppose $S = \{v_1, \ldots, v_k\}$ is set of linearly independent vectors in $\mathbb{R}^n$.

Key Question:

How to find an orthonormal basis of span(S)?

Equivalent Question: how to find an orthonormal set such that

$$\mathrm{span}(U) = \mathrm{span}(S) \;\; ?$$

Is Finding Orthonormal Basis Simple?

Sometimes, yes. e.g. $\mathbb{R}^n$; e.g. column space of $\begin{bmatrix} 1 & 0 & a & 0 & c \\ 0 & 1 & b & 0 & d \\ 0 & 0 & 0 & 1 & e \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

ortho-basis { orthogonal, unit norm
span others.

Sometimes, no.

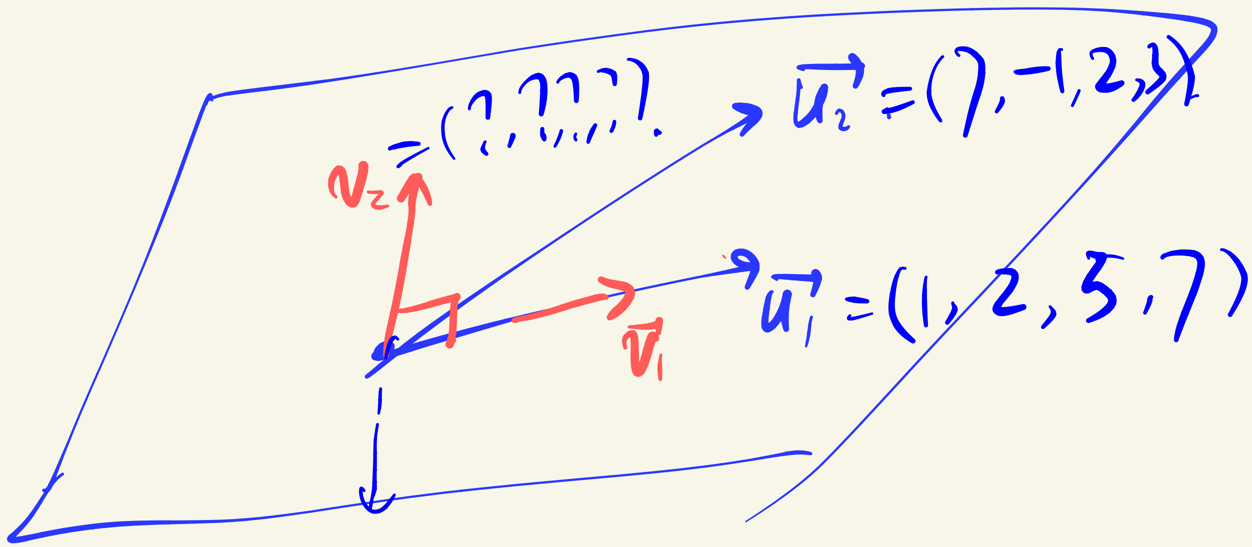
not orthogonal

e.g. Column space of $Z = \begin{bmatrix} -a & -c \\ -b & -d \\ \mathbf{1} & 0 \\ 0 & -e \\ 0 & \mathbf{1} \end{bmatrix}$

Think: Can you use linear combination of columns to find orthonormal basis of a column space?

2D

High



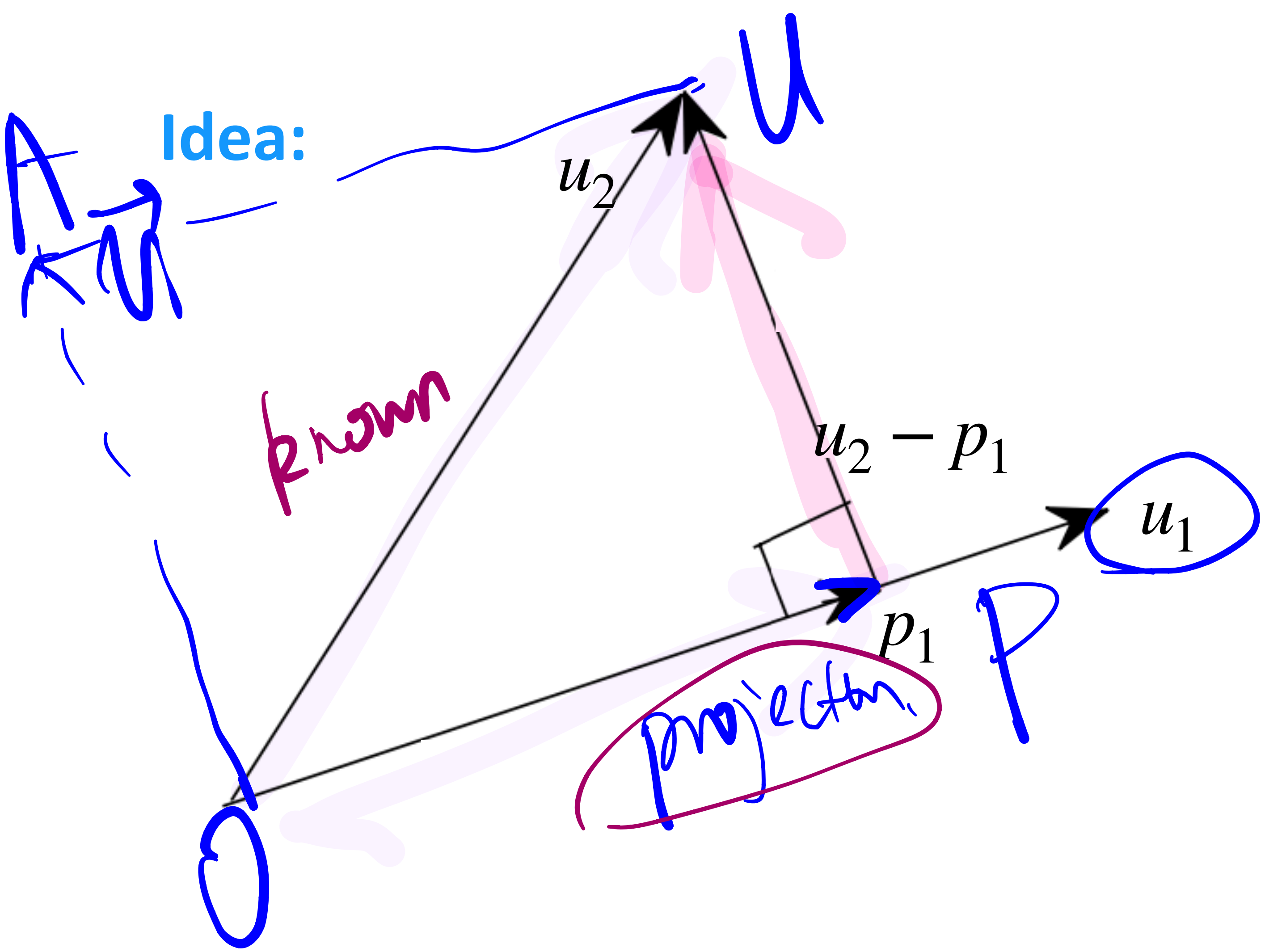
find $\vec{v_2} \perp \vec{v_1}$. on the plane, done.

$$\vec{p}_1 = \left\langle \vec{u_2}, \frac{\vec{u_1}}{\|\vec{u_1}\|} \right\rangle \frac{\vec{u_1}}{\|\vec{u_1}\|} = \left(\begin{bmatrix} 7 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \frac{\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}}{\|(1,2,5,7)\|} \right) \frac{(1,2,5,7)}{\| \sim \|}$$

Exercise: Simple Case

$$Z = \begin{bmatrix} 2 & 4 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} = [u_1, u_2].$$

Exercise: Find an orthonormal basis of span($\{u_1, u_2\}$)



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

$$v_2 = ?$$

$(\because \vec{OP} + \vec{PU} = \vec{OU})$

$\vec{OA} = \vec{PU}$

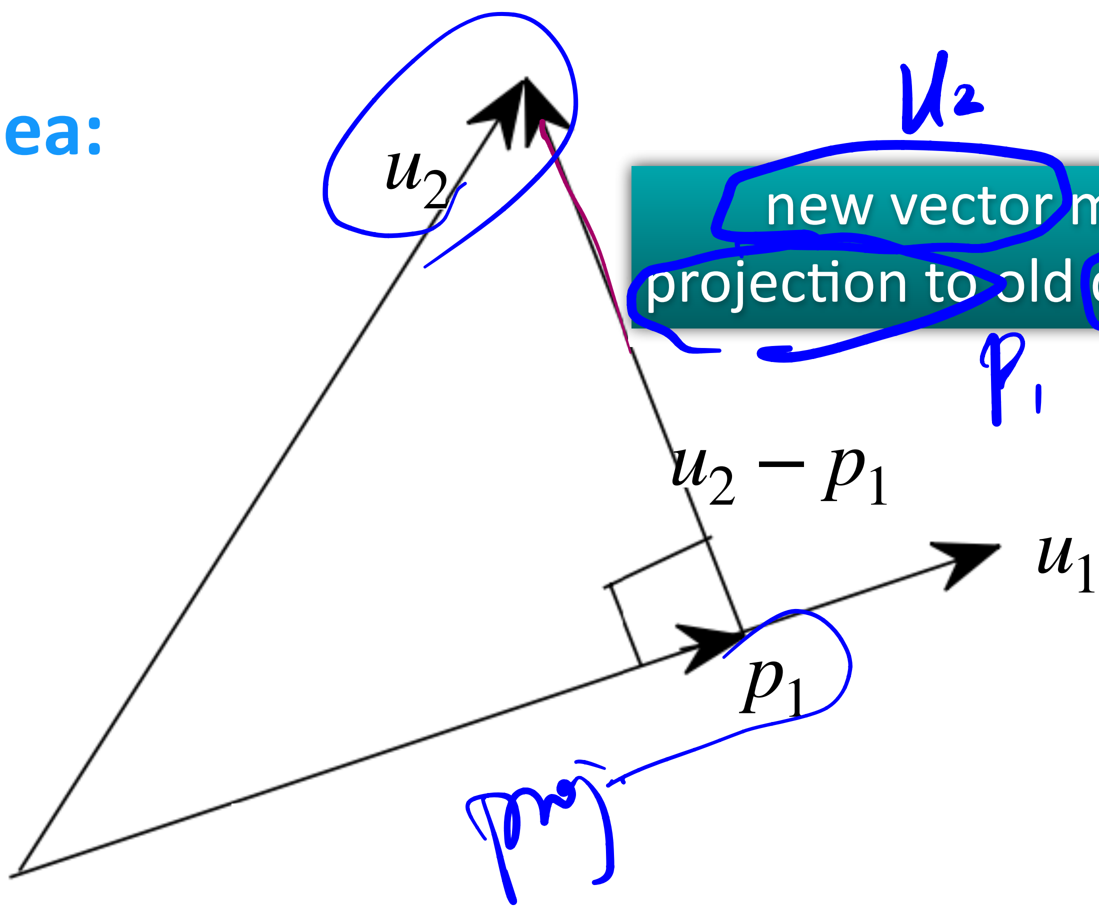
$= \vec{OU} - \vec{OP}$

Exercise: Simple Case

$$Z = \begin{bmatrix} 2 & 4 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} = [u_1, u_2].$$

Exercise: Find an orthonormal basis of span($\{u_1, u_2\}$)

Idea:



new vector minus projection to old directions

u_2 p_1 u_1

old vector

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

$$v_2 = \frac{u_2 - p_1}{\|u_2 - p_1\|}$$

where $p_1 = \langle u_2, v_1 \rangle v_1$

coeff orthonor basis of span{u_1}.

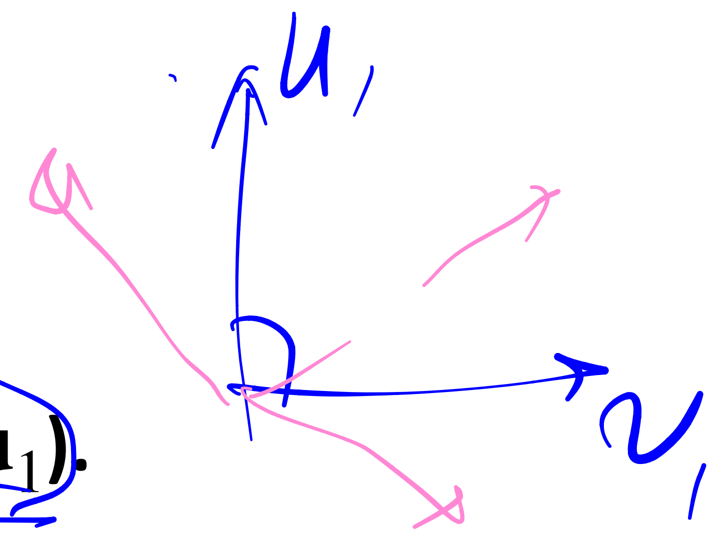
Design Principles of k=2 Case

Find orthonormal set U s.t. $\text{span}(U) = \text{span}(\mathbf{u}_1, \mathbf{u}_2)$.

Idea: **Induction ("数学归纳").**

First, find ortho-basis of $\text{span}(\mathbf{u}_1)$;

Then find ortho-basis of $\text{span}(\mathbf{u}_1, \mathbf{u}_2)$



Step 1: Find the orthonormal basis $\{\mathbf{v}_1\}$ of span($\mathbf{u}_1$).

Step 2: Find the second vector $\mathbf{v}_2$ in the ortho-basis of $\text{span}(\mathbf{u}_1, \mathbf{u}_2)$.

Design principle: This vector should be

i) Orthogonal to $\mathbf{u}_1$;

ii) in $\text{span}(\mathbf{u}_1, \mathbf{u}_2)$.

iii) Norm = 1.

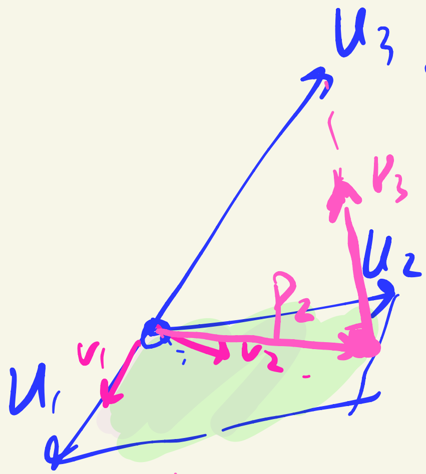
Answer: Direction same as $u_2 - p_1$

Where $p_1 = \text{Proj}_{span(u_1)}(u_2) = \langle u_2, v_1 \rangle v_1$

u_2

new vector minus projection to old directions

p_1 u_1



$v_3 = u_3 - p_2$

$= u_3 - (\text{proj of } u_3)$

Q: How to compute proj of u_3 to $W_2 = \text{span}(\{u_1, u_2\})$?

Step 1 orthonormal basis $\{v_1, v_2\}$ of W_2

Step 2. $\text{proj}_{W_2}(u_3) = (u_3^T v_1)v_1 + (u_3^T v_2)v_2$.

① know orthonormal basis.
of span$\{u_1, u_2\}, u_3\}$ → Compute projection
onto span$\{u_1, u_2\}, u_3\}$

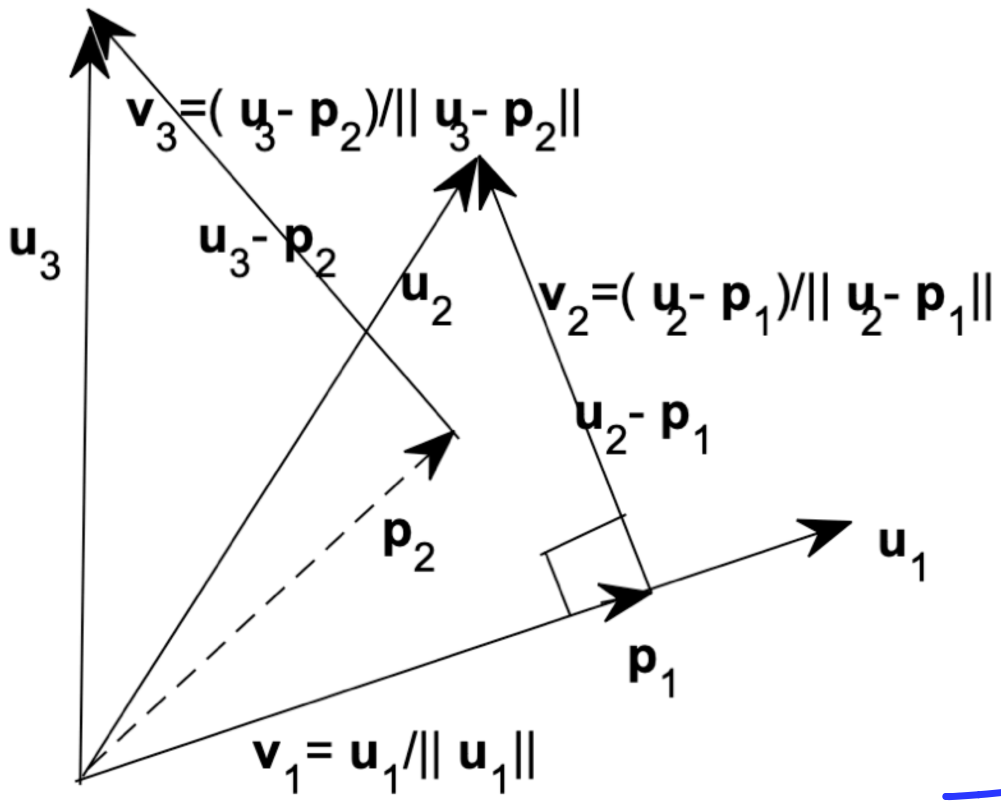
② Compute projection
onto span$\{u_1, u_2\}, u_3$. → Compute orthonormal basis
span$\{u_1, u_2, u_3\}, u_4$.

Two sides of same coin.

Chicken & egg problem

k=3 Case

Find orthonormal set U s.t. $\mathrm{span}(U) = \mathrm{span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\})$.



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

$$v_2 = \frac{u_2 - p_1}{\|u_2 - p_1\|}$$

where $p_1 = \mathrm{Proj}_{span(u_1)}(u_2) = \langle u_2, v_1 \rangle v_1$

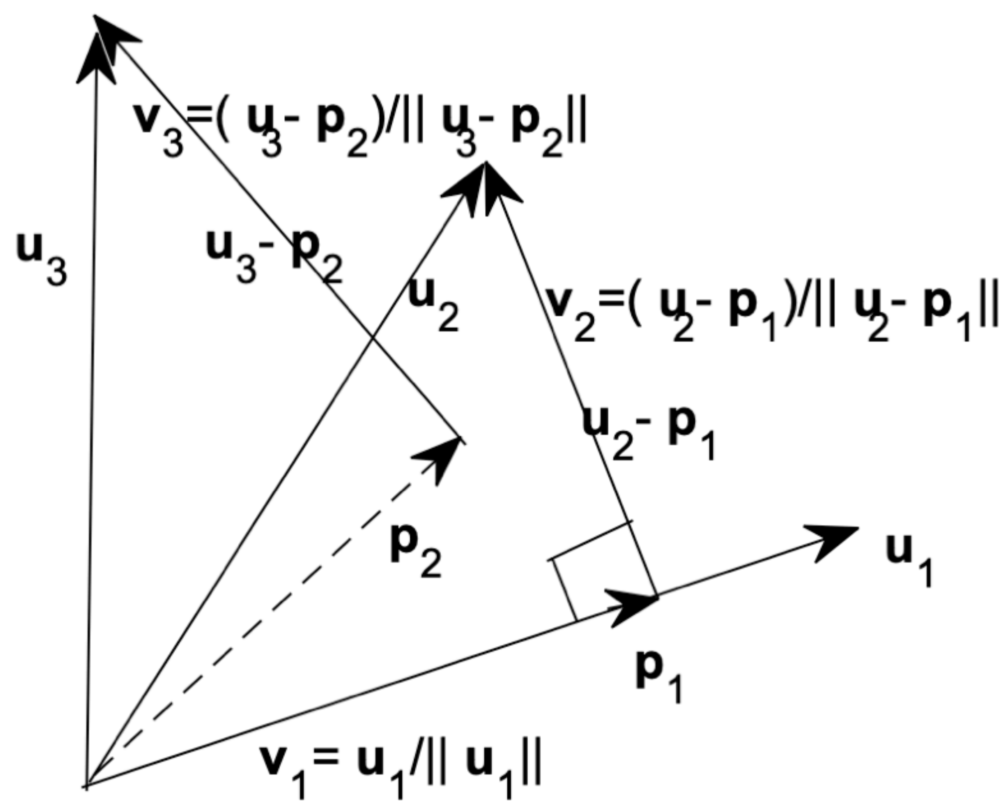
new vector v_3.

$v_3 = ?$

new vector minus projection to old directions

k=3 Case

Find orthonormal set U s.t. $\mathrm{span}(U) = \mathrm{span}(\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\})$.



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

$$v_2 = \frac{u_2 - p_1}{\|u_2 - p_1\|}$$

where $p_1 = \text{Proj}_{span(u_1)}(u_2) = \langle u_2, v_1 \rangle v_1$

new vec

proj

$$v_3 = \frac{u_3 - p_2}{\|u_3 - p_2\|}$$

new vector minus projection to old directions

where $p_2 = \text{Proj}_{span(u_1,u_2)}(u_3) = \langle u_3, v_1 \rangle v_1 + \langle u_3, v_2 \rangle v_2$

ortho basis of span(u1,u2).

Back to the Question: k=2 Case

Find orthonormal set U s.t. $\text{span}(U) = \text{span}(\mathbf{u}_1, \mathbf{u}_2)$.

Steps:

Step 1: Find the orthonormal basis of span($\mathbf{u}_1$).

Step 2: Find the second vector in the ortho-basis of span($\mathbf{u}_1, \mathbf{u}_2$).

Step 3: Find the 3rd vector in the ortho-basis of span($\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$)

This vector should be:

i) Orthogonal to $\mathbf{u}_1, \mathbf{u}_2$;

ii) in $\text{span}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

iii) Norm = 1.

Answer: Direction same as $u_3 - p_2$

where $p_2 = \text{Proj}_{span(u_1,u_2)}(u_3) = \langle u_3, v_1 \rangle v_1 + \langle u_3, v_2 \rangle v_2$

new vector minus projection to old directions

Gram-Schmidt Process

Suppose $S = \{u_1, \ldots, u_k\}$ is a set of linearly independent elements.

Question: Based on S, can we find an orthonormal set U such that

$$\mathrm{span}(U) = \mathrm{span}(S)\ \textbf{?}$$

Gram-Schmidt Process

pseudo code (not runnable in C)

伪代码.

Input $S = \{u_1, \ldots, u_k\}$

For $i = 1, \ldots, k$

If $i = 1$:

$$p_1 = 0$$

Else:

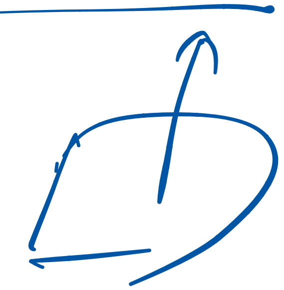
$$p_{i-1} = \sum_{j=1}^{i-1} \langle u_i, v_j \rangle v_j = \textbf{Proj}_{\textbf{span}(u_1,\ldots,u_{i-1})}(u_i).$$

projection to old directions

Set $$v_i = \frac{u_i - p_{i-1}}{\|u_i - p_{i-1}\|} \in \textbf{span}(u_1, \ldots, u_{i-1})^{\perp}$$

new vector minus projection to old directions

Return $U = \{v_1, \ldots, v_k\}$



Gram-Schmidt Process

Suppose $S = \{u_1, \ldots, u_k\}$ is a set of linearly independent vectors.

Question: Based on S, can we find an orthonormal set U such that

$$\mathrm{span}(U) = \mathrm{span}(S) \text{ ?}$$

Proposition 17.3 (Gram-Schmidt Process)

The set $U = \{v_1, \ldots, v_k\}$ returned by the Gram-Schmidt process is an orthonormal basis of $\mathrm{span}(S)$.

Gram-Schmidt Process

Suppose $S = \{u_1, \ldots, u_k\}$ is a set of linearly independent elements.

Question: Based on S, can we find an orthonormal set U such that

$$\mathrm{span}(U) = \mathrm{span}(S) \text{ ?}$$

> Proposition 17.3 (Gram-Schmidt Process)
>
> The set $U = \{v_1, \ldots, v_k\}$ returned by the Gram-Schmidt process is an orthonormal basis of $\mathrm{span}(S)$.

Proof: Orthogonality: $v_{k+1} \perp \mathrm{span}(\{u_1, \ldots, u_k\})$

Norm: $\|v_k\| = 1$.

Linearly Independence: Proposition 17.1

k

$\dim(\mathrm{span}(S)) = k$ implies the result (Coro 17.1)

Gram-Schmidt Process Example

Example

$$S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} \right\} \subseteq \mathbb{R}^4$$

In $\mathbb{R}^n$, the standard inner product is the scalar product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T \mathbf{x}$.

Now find the orthonormal basis for $\mathrm{Span}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

Gram-Schmidt Process Example

Example

$$S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} \right\}$$

Step 1: $\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$

Step 2: calculate

$$\mathbf{u}_2' = \mathbf{u}_2 - < \mathbf{u}_2, \mathbf{v}_1 > \mathbf{v}_1$$

$$= \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix} - \left(\begin{bmatrix} -1 & 4 & 4 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right) \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ -\frac{5}{2} \end{bmatrix}$$

$$\mathbf{v}_2 = \frac{\mathbf{u'}_2}{\| \mathbf{u'}_2 \|} = \frac{1}{5}\begin{bmatrix} -\frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ -\frac{5}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

Gram-Schmidt Process Example

Example

Step3: calculate

$$\mathbf{u}_3' = \mathbf{u}_3 - < \mathbf{u}_3, \mathbf{v}_1 > \mathbf{v}_1 - < \mathbf{u}_3, \mathbf{v}_2 > \mathbf{v}_2$$

$$= \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 4 & -2 & 2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right) \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$- \left(\begin{bmatrix} 4 & -2 & 2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right) \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix}$$

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\mathbf{v}_2 = \frac{\mathbf{u'}_2}{\| \mathbf{u'}_2 \|} = \frac{1}{5} \begin{bmatrix} -\frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ -\frac{5}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{v}_3 = \frac{\mathbf{u'}_3}{\| \mathbf{u'}_3 \|} = \frac{1}{4} \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

Application of Orthonormal Basis: Compute Projection

Application question:

How to compute the projection of u onto a subspace span$(u_1, \ldots, u_k)$?

Answer:

First, find an orthonormal basis $v_1, \ldots, v_k$ by Gram-Schimidt process.

Second, use formula $p = \sum_{i=1}^{k} \langle u, v_i \rangle v_i$

E.g. compute the projection onto span(S), where

$$S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix} \right\} \subseteq \mathbb{R}^4$$

Application of Orthonormal Basis: Solve LS-Problem

Lemma 16.1 (special case S = C(A))

Consider a least squares problem. The following statements are equivalent:

1. $\mathbf{y}$ minimizes $||A\mathbf{x} - \mathbf{b}||$
2. $\mathbf{b} - A\mathbf{y} \perp S$, i.e. $A\mathbf{y} = \text{Proj}_{C(A)}(\mathbf{b})$;

Application question: How to compute LS-solution?

Step 1: Compute $\text{Proj}_{C(A)}(\mathbf{b})$?

How? See the previous slide.

Step 2: Solve $A\mathbf{y} = \mathbf{p}$.

no need to compute A^TA.

Alternative method [Lec 16]

Solve $A^T A\mathbf{y} = A^T\mathbf{b}$.

Summary Today (Write Your Own)

One sentence summary:

Detailed summary:

Review Questions

How many propositions do we have today? [4]
How many theorems do we have today?
What are they?
Which is the most nontrivial one?

How many definitions do we have today?
What are they?

Summary Today (of Instructor)

One sentence summary:

We have studied Orthonormal basis and Gram-Schmidt Process.

Detailed summary:

1) **Orthonormal basis.**

—Basis consisting of unit vectors that are orthogonal to each other.

—**Theorem 17.1**: Representation of vector by ortho-basis of whole space.

—**Theorem 17.2**: Representation of projection by ortho-space of a subspace.

2) **Orthogonal matrix.**

—Square matrix whose columns form a orthonormal basis of $\mathbb{R}^n$.

—**Properties**: $Q^\top Q = I,$ keep the norm $\|Qx\| = \|x\|$;

3) **Gram-Schmidt process.**

—**Goal**: Construct orthonormal basis of a subspace, from a basis of the subspace.

—**Main trick**: every new vector minus projection to old directions

—**Theoretical result**: Gram-Schmidt process indeed returns an orthonormal basis.