Lecture 17 recap

1) Definition of length of a vector, distance, angle, dot product between two vectors.

2) Dot product and matrix product.

3) Some results on dot products.

4) An example on information retrieval.

5) Orthogonality (orthonormality). Orthogonal and orthonormal sets (bases).

6) Usefulness of orthogonal and orthonormal bases.

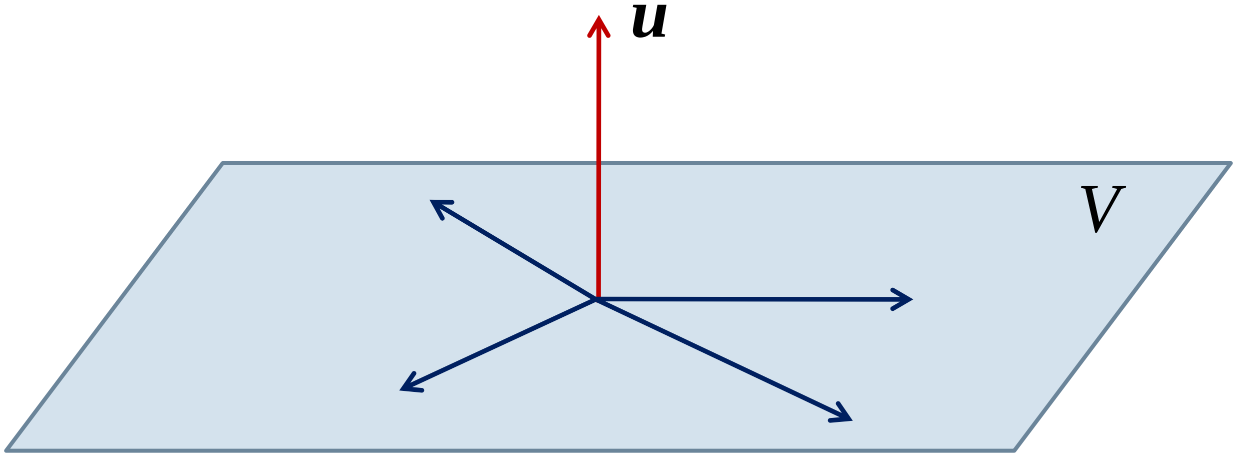
LECTURE 18

Orthogonal and Orthonormal Bases (cont'd)

Best Approximation

Definition (vector orthogonal to a space)

Let V be a subspace of $\mathbb{R}^n$. A vector $\boldsymbol{u}$ is orthogonal (or perpendicular) to V if $\boldsymbol{u}$ is orthogonal to all vectors in V.



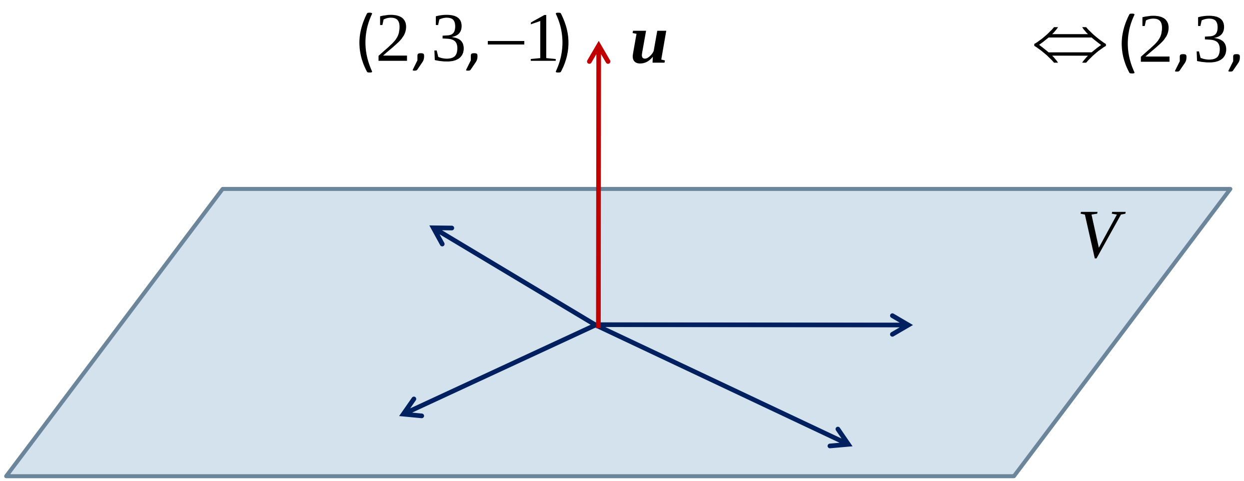
Example (vector orthogonal to a space)

$V = \{(x, y, z) \mid 2x + 3y - z = 0\}$ is a subspace of $\mathbb{R}^3$

(it is a plane in $\mathbb{R}^3$ containing the origin).

$$(x, y, z) \in V \Leftrightarrow 2x + 3y - z = 0$$

$$\Leftrightarrow (2,3,-1) \cdot (x, y, z) = 0$$



$(2,3,-1)$ is orthogonal to all vectors in V.

So $(2,3,-1)$ is orthogonal to V.

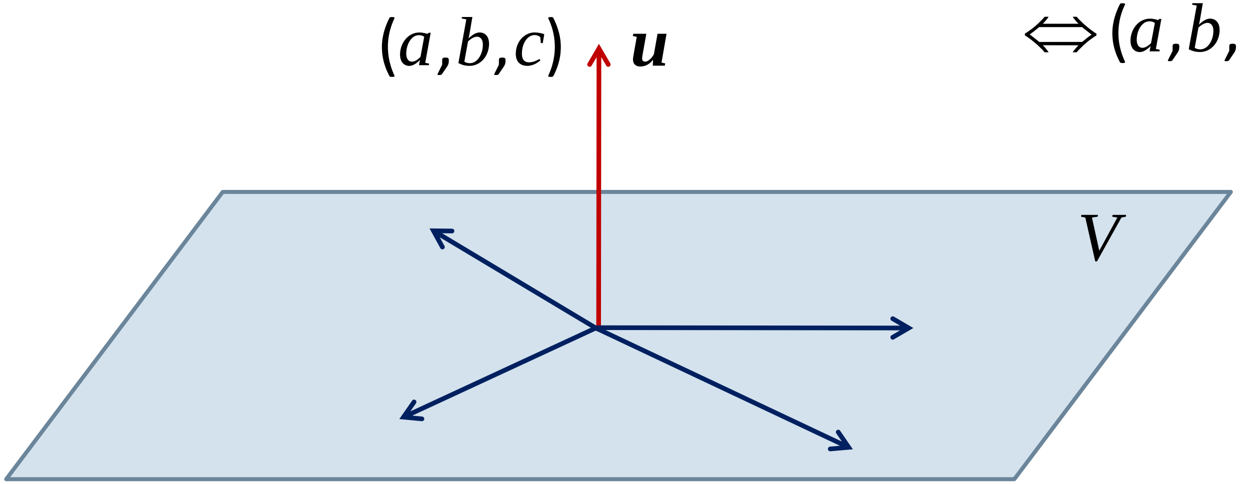
Example (vector orthogonal to a space)

$V = \{(x, y, z) \mid ax + by + cz = 0\}$ is a subspace of $\mathbb{R}^3$

(it is a plane in $\mathbb{R}^3$ containing the origin).

$$(x, y, z) \in V \Leftrightarrow ax + by + cz = 0$$

$$\Leftrightarrow (a, b, c) \cdot (x, y, z) = 0$$



(a,b,c) is orthogonal to all vectors in V.

So (a,b,c) is orthogonal to V.

Example (vector orthogonal to a space)

$V = \{(x, y, z) \mid ax + by + cz = 0\}$ is a subspace of $\mathbb{R}^3$

Let $\boldsymbol{n} = (a, b, c)$.

$$V = \{(x, y, z) \mid ax + by + cz = 0\}$$

$$= \{\boldsymbol{u} \in \mathbb{R}^3 \mid \boldsymbol{n} \cdot \boldsymbol{u} = 0\}$$

$\boldsymbol{n}$ is orthogonal to V. We say that $\boldsymbol{n}$ is a <u>normal vector</u> of V.

Question: If $\boldsymbol{n}$ is orthogonal to V, is $c\boldsymbol{n}$ orthogonal to V for every $c \neq 0$?

YES!

Example (vector orthogonal to a space)

$V = \text{span}\{(1,1,-1,0),(0,1,1,-1)\}$ is a subspace of $\mathbb{R}^4$.

Question: Find all vectors orthogonal to V.

Remember:

$\boldsymbol{v} = (w,x,y,z)$ is orthogonal to V

$\Leftrightarrow$

$\boldsymbol{v} = (w,x,y,z)$ is orthogonal to every vector in V.

Wow! Isn't this very difficult to check??

Example (vector orthogonal to a space)

$V = \text{span}\{(1,1,-1,0),(0,1,1,-1)\}$ is a subspace of $\mathbb{R}^4$.

Question: Find all vectors orthogonal to V.

In general, if $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$, then $\boldsymbol{v}$ is orthogonal to V if and only if $\boldsymbol{v}$ is orthogonal to $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$, that is, $\boldsymbol{v} \cdot \boldsymbol{u}_i = 0$ for all $i = 1, \ldots, k$.

This is much easier!

Example (vector orthogonal to a space)

$V = \text{span}\{(1,1,-1,0),(0,1,1,-1)\}$ is a subspace of $\mathbb{R}^4$.

Question: Find all vectors orthogonal to V.

Definition (orthogonal projection)

Let V be a subspace of $\mathbb{R}^n$.

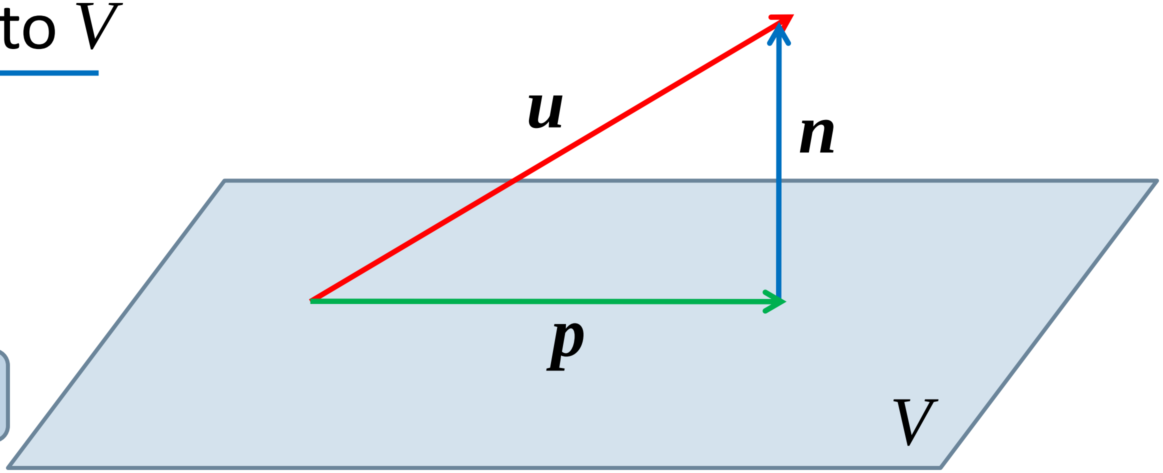
Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$$

where $\boldsymbol{n}$ is orthogonal to V

and $\boldsymbol{p}$ belongs to V.

$\boldsymbol{p}$ is called the (orthogonal) projection of $\boldsymbol{u}$ onto V.



Wait a minute…

Let V be a subspace of $\mathbb{R}^n$.

Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$$

where $\boldsymbol{n}$ is orthogonal to V

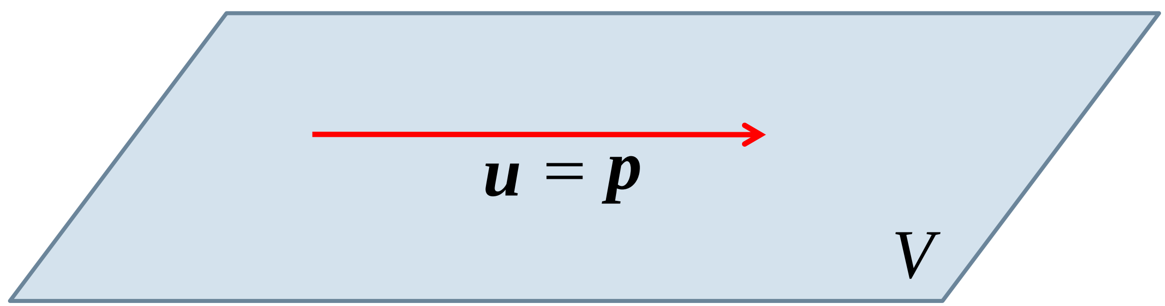
and $\boldsymbol{p}$ belongs to V.

$$\boldsymbol{n} = \boldsymbol{0}$$

$$\boldsymbol{u} = \boldsymbol{0} + \boldsymbol{u}$$

What happens if

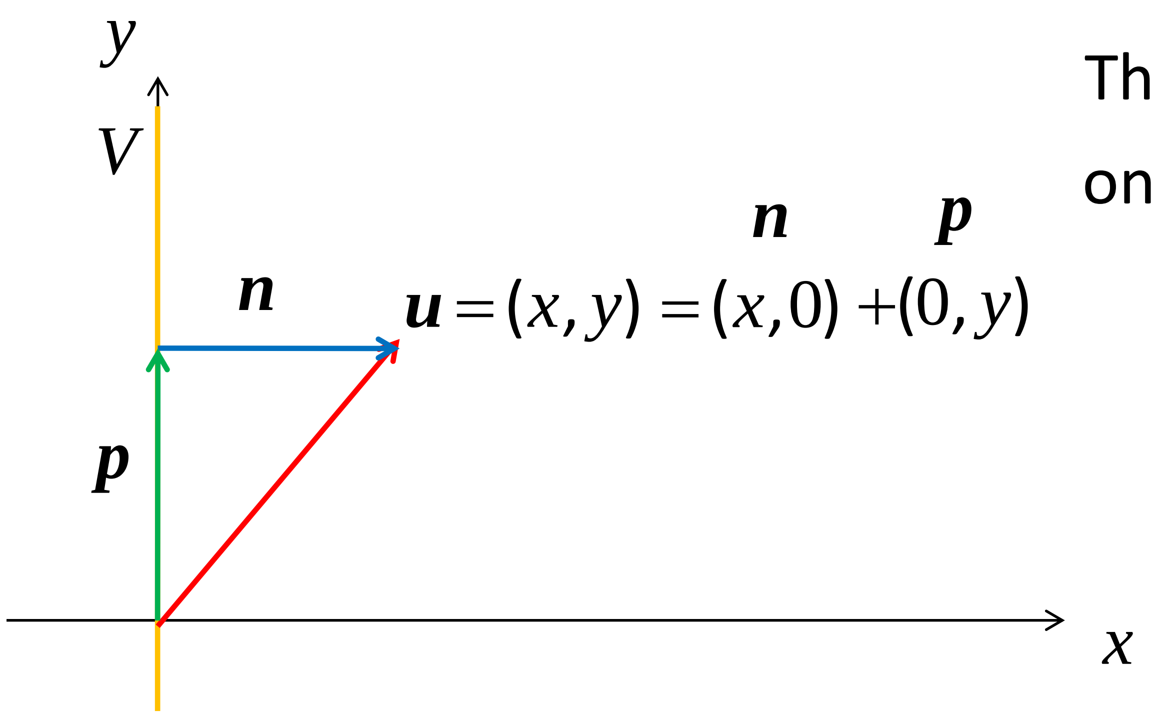
$\boldsymbol{u}$ belongs to V?



Projection in R^2 and R^3

Let V be a subspace of $\mathbb{R}^2$.

$$V = \{(0, y) \mid y \in \mathbb{R}\} \text{ (} V \text{ is the } y\text{-axis)}$$



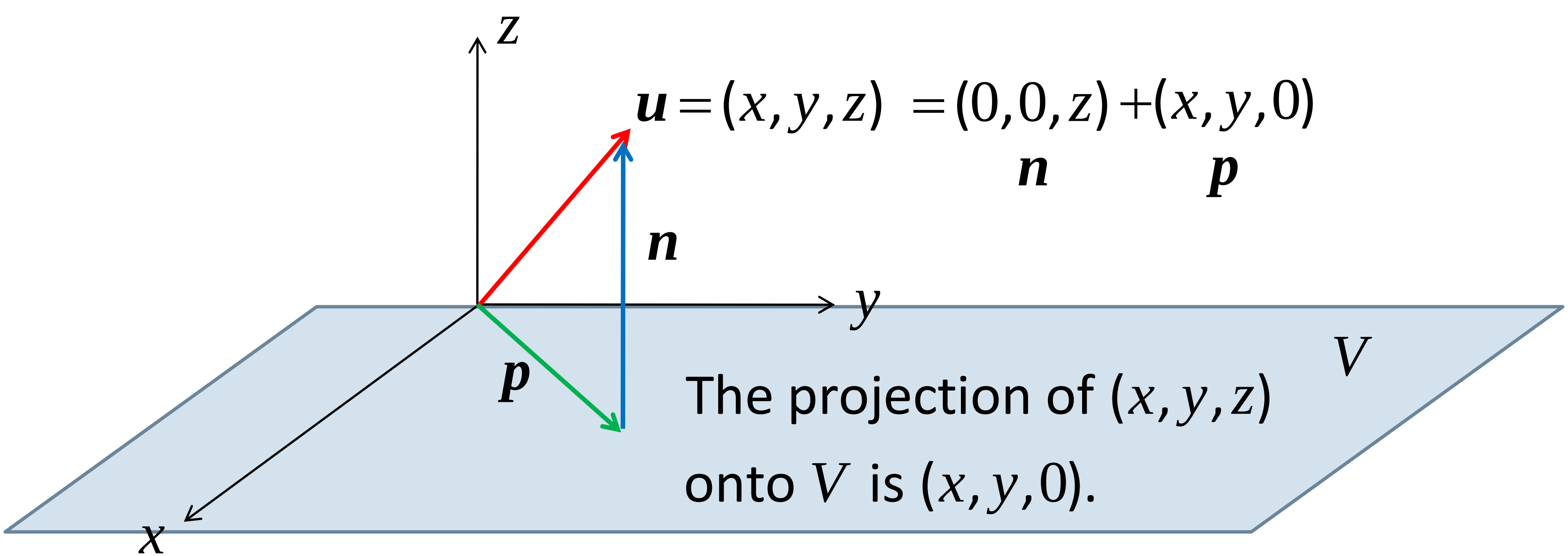
$$\boldsymbol{u} = (x, y) = \overset{\boldsymbol{n}}{(x, 0)} + \overset{\boldsymbol{p}}{(0, y)}$$

The projection of (x, y) onto V is $(0, y)$.

Projection in R^2 and R^3

Let V be a subspace of $\mathbb{R}^3$.

$$V = \{(x, y, 0) \mid x, y \in \mathbb{R}\} \text{ (} V \text{ is the } xy\text{-plane)}$$



Wait a(nother) minute…

Let V be a subspace of $\mathbb{R}^2$.

$$V = \{(0, y) \mid y \in \mathbb{R}\} \text{ (} V \text{ is the } y\text{-axis)}$$

Let V be a subspace of $\mathbb{R}^3$.

$$V = \{(x, y, 0) \mid x, y \in \mathbb{R}\} \text{ (} V \text{ is the } xy\text{-plane)}$$

Questions:

What if the subspaces are not so 'trivial'?

How can we compute orthgonal projections in general?

What is required?

Theorem (orthogonal projection)

Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$.

If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then the projection of $\boldsymbol{w}$ onto V is

$$\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \ldots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

Looks familiar?

If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthonormal basis for V, then the projection of $\boldsymbol{w}$ onto V is

$$(\boldsymbol{w}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w}\cdot\boldsymbol{v}_2)\boldsymbol{v}_2 + \ldots + (\boldsymbol{w}\cdot\boldsymbol{v}_k)\boldsymbol{v}_k$$

An expression we have seen before

If $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for a vector space V, then for any vector $\boldsymbol{w} \in V$,

$$\boldsymbol{w} = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \ldots + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

From Lect. 17

Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$.

If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then the projection of $\boldsymbol{w}$ onto V is

$$\left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \ldots + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

Why does this make sense?

An expression we have seen before

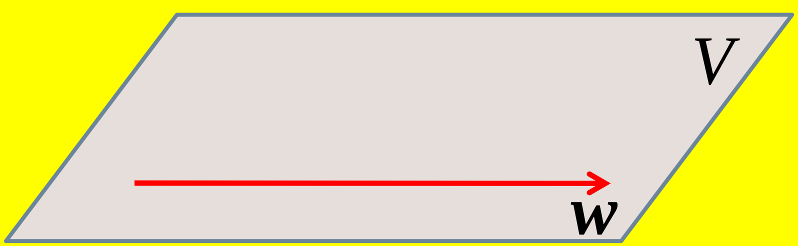
$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V (subspace of $\mathbb{R}^n$)

...then for any vector $\boldsymbol{w} \in V$,

$$\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \ldots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

$$= \boldsymbol{w}$$

$$\boldsymbol{w} = \boldsymbol{w} + \boldsymbol{0}$$

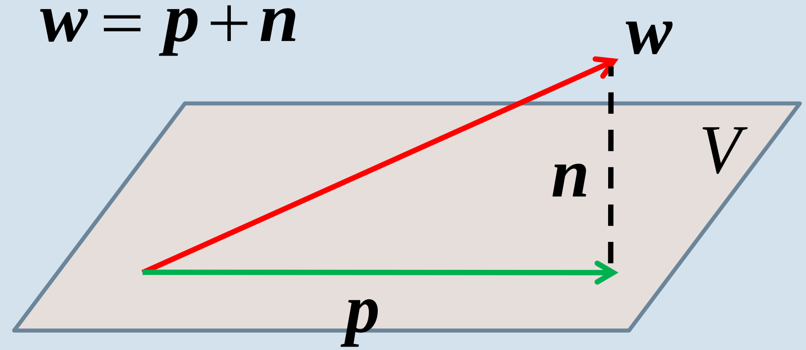


...and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$.

$$\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \ldots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

is the projection of $\boldsymbol{w}$ onto V.

$$\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{n}$$



Theorem (orthogonal projection)

Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$.

If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V, then the projection of $\boldsymbol{w}$ onto V is

$$\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \ldots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthonormal basis for V, then the projection of $\boldsymbol{w}$ onto V is

$$(\boldsymbol{w}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w}\cdot\boldsymbol{v}_2)\boldsymbol{v}_2 + \ldots + (\boldsymbol{w}\cdot\boldsymbol{v}_k)\boldsymbol{v}_k$$

Example (orthogonal projection)

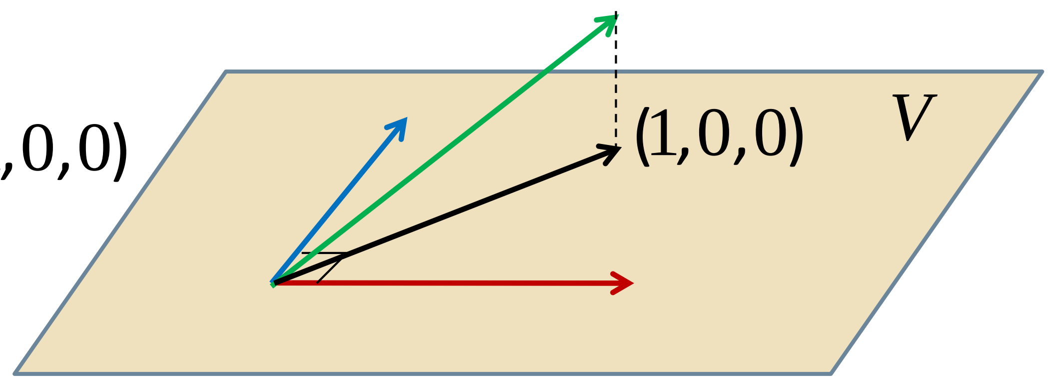
Let $V = \text{span}\{(1,0,1),(1,0,-1)\}$ be a subspace of $\mathbb{R}^3$ (V is a plane).

$$(1,0,1)\cdot(1,0,-1) = 1+0-1 = 0$$

$\Rightarrow \{(1,0,1),(1,0,-1)\}$ is an orthogonal basis for V.

What is the projection of $\boldsymbol{w} = (1,1,0)$ onto V?

$$\frac{(1,1,0)\cdot\boxed{(1,0,1)}}{\|(1,0,1)\|^2}(1,0,1) + \frac{(1,1,0)\cdot\boxed{(1,0,-1)}}{\|(1,0,-1)\|^2}(1,0,-1) = (1,0,0)$$



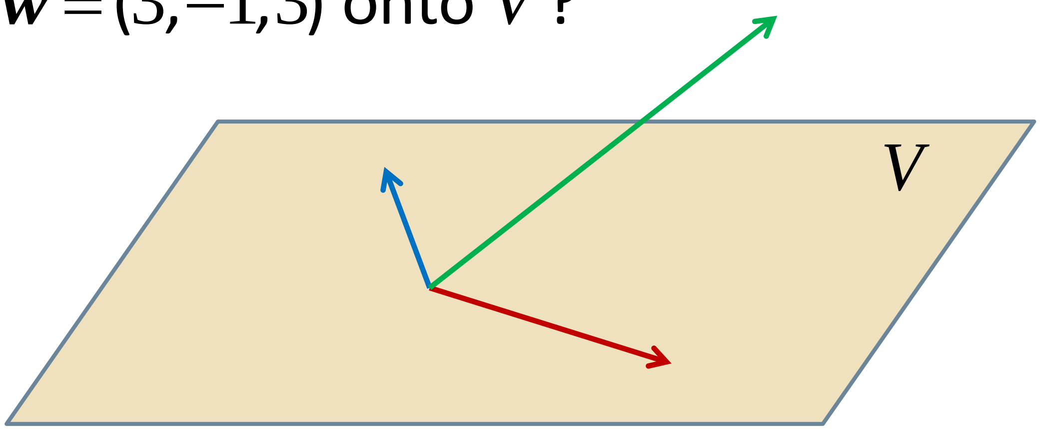
Example (orthogonal projection)

Let $V = \text{span}\{(1,1,1),(1,3,-1)\}$ be a subspace of $\mathbb{R}^3$ (V is a plane).

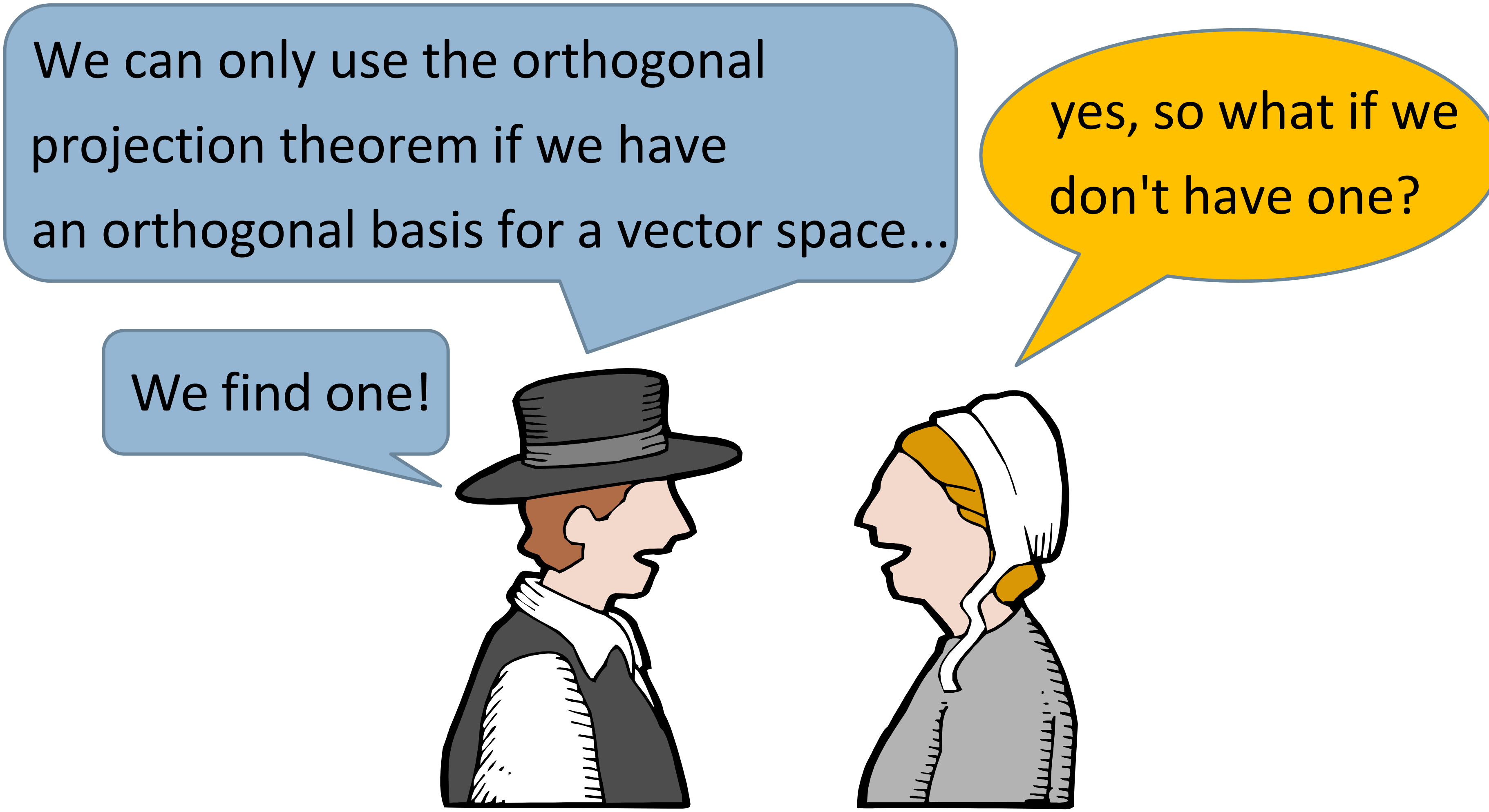
$(1,1,1)\cdot(1,3,-1) = 3 \neq 0$

$\Rightarrow \{(1,1,1),(1,3,-1)\}$ is a basis, but not an orthogonal basis for V.

What is the projection of $\boldsymbol{w} = (3,-1,3)$ onto V?



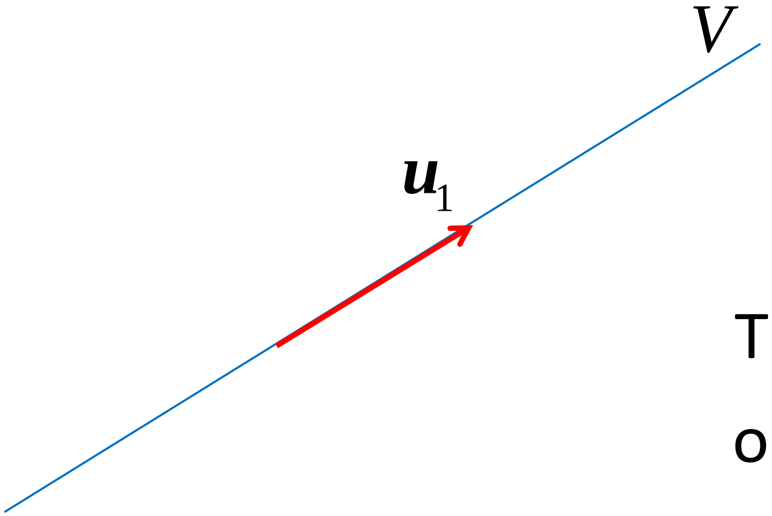
Finding orthogonal bases



Finding orthogonal bases

Let V be a subspace of $\mathbb{R}^n$.

1) V is one-dimensional, that is, $V = \text{span}\{\boldsymbol{u}_1\}$, $\boldsymbol{u}_1 \neq \mathbf{0}$.

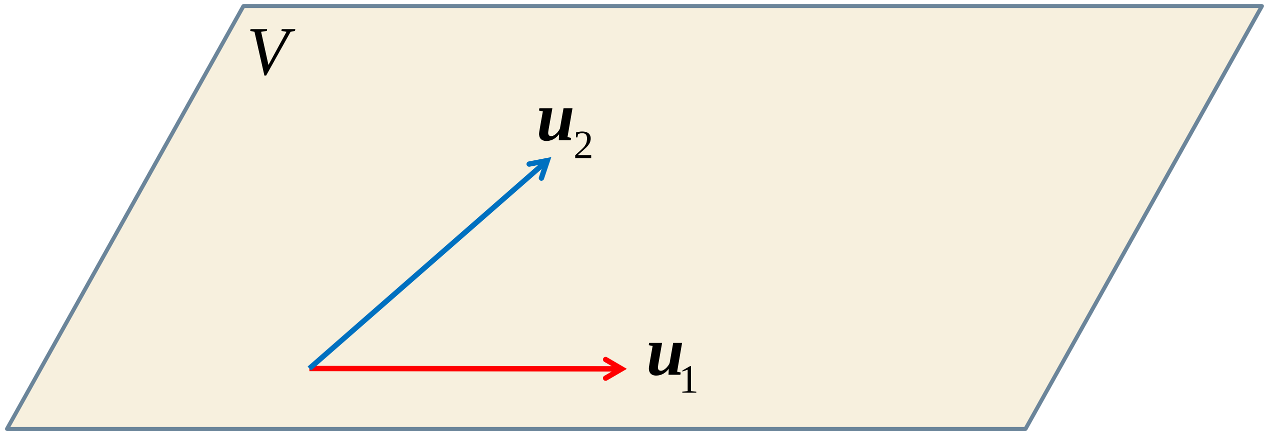


Trivially, $\{\boldsymbol{u}_1\}$ is an orthogonal basis for V.

Finding orthogonal bases

Let V be a subspace of $\mathbb{R}^n$.

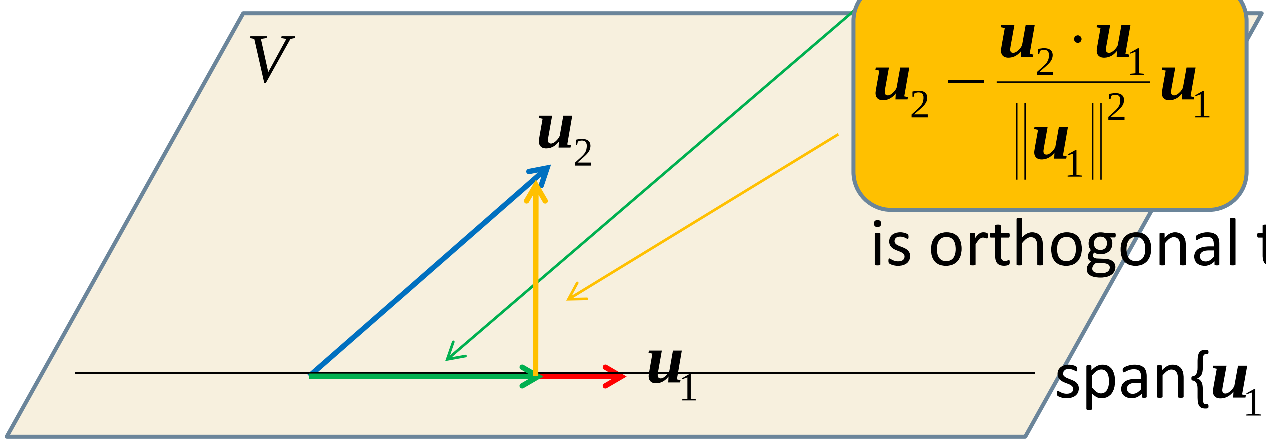
2) V is two-dimensional, that is, $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are linearly independent vectors.



Finding orthogonal bases

span$\{\boldsymbol{u}_1\}$ is a subspace of V and $\{\boldsymbol{u}_1\}$ is an orthogonal basis for span$\{\boldsymbol{u}_1\}$.

By orthogonal projection theorem, $\dfrac{\boldsymbol{u}_2 \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1$ is the projection of $\boldsymbol{u}_2$ onto span$\{\boldsymbol{u}_1\}$

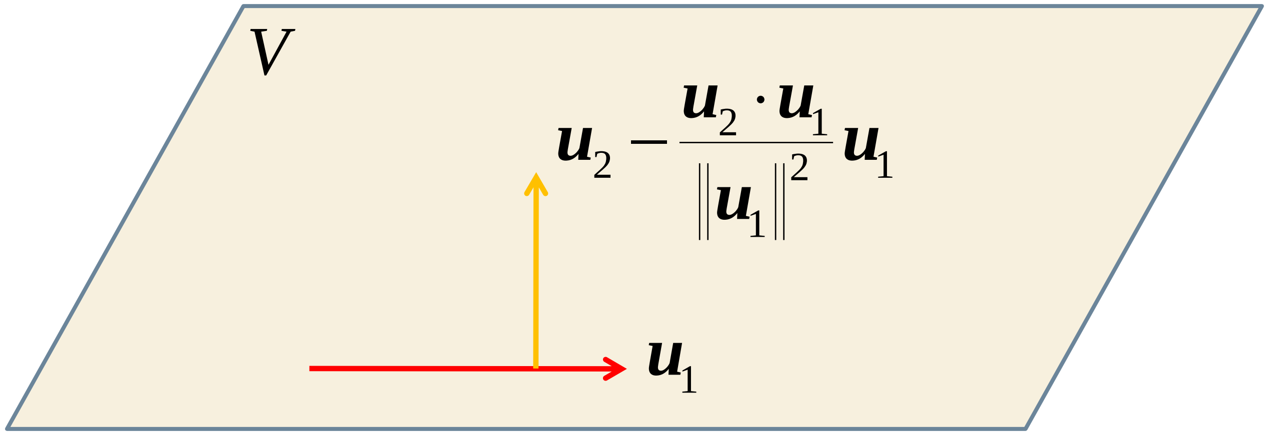


$\boldsymbol{u}_2 - \dfrac{\boldsymbol{u}_2 \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1$ is orthogonal to $\boldsymbol{u}_1$

Finding orthogonal bases

2) V is two-dimensional, that is, $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are linearly independent vectors.

$$\left\{\boxed{\boldsymbol{u}_1}\ ,\ \boxed{\boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1}\right\}$$ is an orthogonal basis for V.

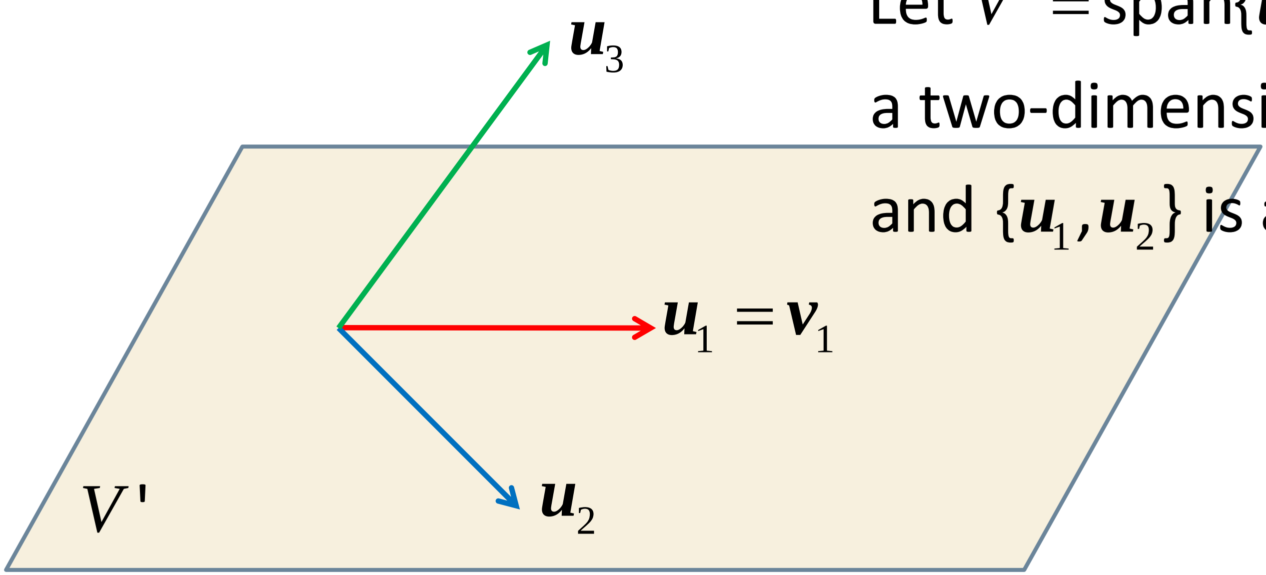


Finding orthogonal bases

Let V be a subspace of $\mathbb{R}^n$.

3) V is three-dimensional, that is, $V = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ are linearly independent vectors.

Let $V' = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, then V' is a two-dimensional subspace of V and $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is a basis for V'.

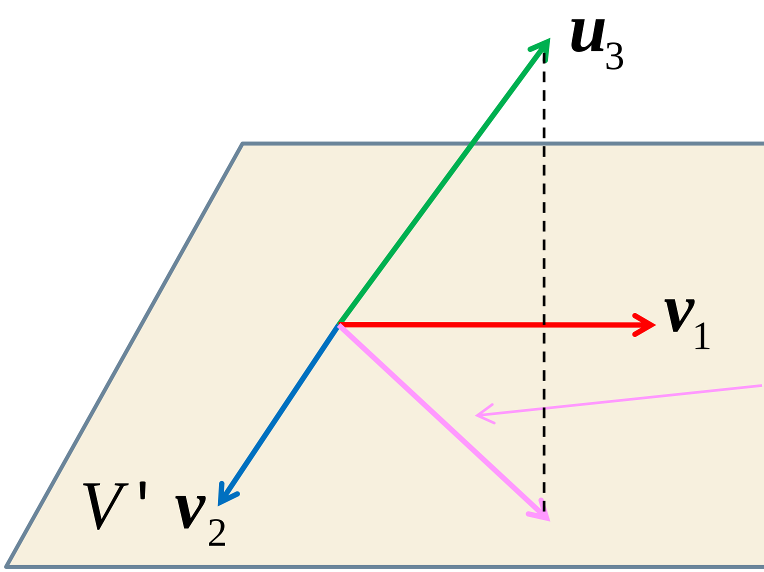


Let $\boldsymbol{v}_1 = \boldsymbol{u}_1$.

Finding orthogonal bases

Let V be a subspace of $\mathbb{R}^n$.

By previous discussion, $\{\boldsymbol{v}_1, \underbrace{\boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1}_{\boldsymbol{v}_2}\}$ is an orthogonal basis for V'.



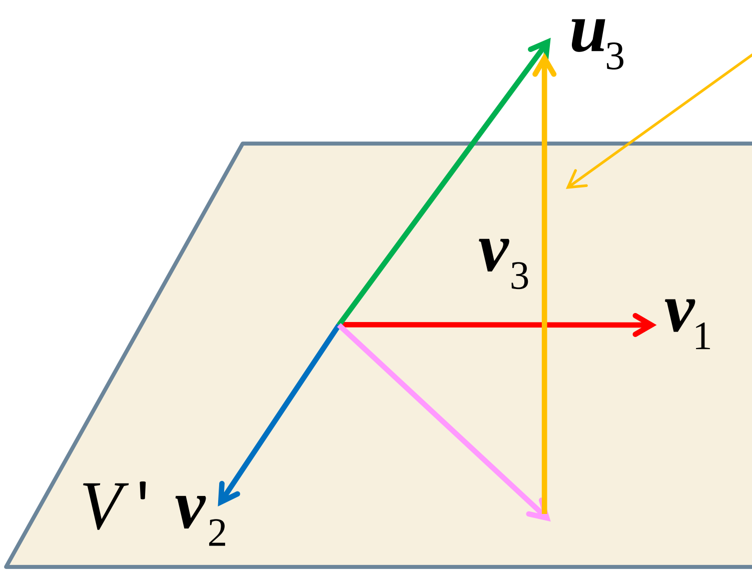
By orthogonal projection theorem, $\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2$ is the projection of $\boldsymbol{u}_3$ onto V'.

Finding orthogonal bases

$\boldsymbol{u}_3 - (\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2)$ is orthogonal to $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$

Let $\boxed{\boldsymbol{v}_3} = \boldsymbol{u}_3 - (\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2)$.

Then $\boldsymbol{v}_3$ is orthogonal to V'.



By orthogonal projection theorem,

$\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2$ is the projection

of $\boldsymbol{u}_3$ onto V'.

Finding orthogonal bases

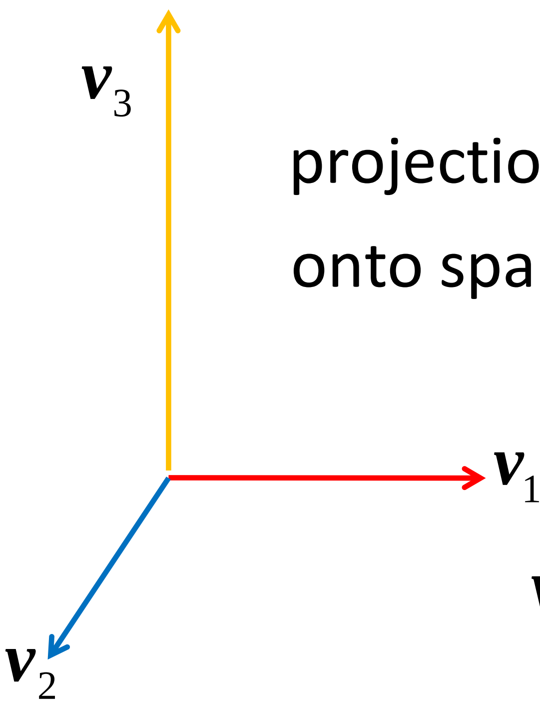
$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthogonal basis for V.

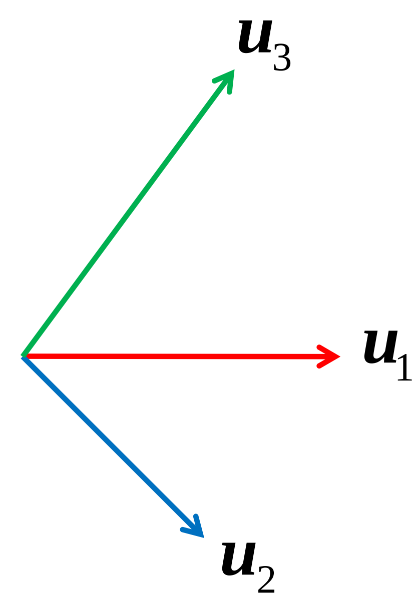
$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis for V.

$$\boldsymbol{v}_1 = \boldsymbol{u}_1 \qquad \boldsymbol{v}_2 = \boldsymbol{u}_2 - \boxed{\frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1}$$

projection of $\boldsymbol{u}_2$ onto span$\{\boldsymbol{v}_1\}$

projection of $\boldsymbol{u}_3$ onto span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$





$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \boxed{\left(\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2\right)}$$

Theorem (Gram-Schmidt Process)

Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a basis for a vector space V.

Let $\boldsymbol{v}_1 = \boldsymbol{u}_1$;

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1; \qquad \boldsymbol{v}_3 = \boldsymbol{u}_3 - \left(\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2\right);$$

$$\vdots$$

$$\boldsymbol{v}_k = \boldsymbol{u}_k - \left(\frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \ldots + \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_{k-1}}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}\right);$$

Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for V.

Theorem (Gram-Schmidt Process)

Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for V.

Let $\boldsymbol{v}_1 = \boldsymbol{u}_1$;

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1; \qquad \boldsymbol{v}_3 = \boldsymbol{u}_3 - \left(\frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2\right);$$

$$\vdots$$

$$\boldsymbol{v}_k = \boldsymbol{u}_k - \left(\frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \ldots + \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_{k-1}}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}\right);$$

$\left\{\frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1, \frac{1}{\|\boldsymbol{v}_2\|}\boldsymbol{v}_2, \ldots, \frac{1}{\|\boldsymbol{v}_k\|}\boldsymbol{v}_k\right\}$ is an orthonormal basis for V.

Example (Gram-Schmidt Process)

Apply Gram-Schmidt Process to transform

$$\{(1,0,1),(0,1,2),(2,1,0)\}$$

into an orthogonal basis for $\mathbb{R}^3$.

Remark: You may choose

$$\boldsymbol{u}_1=(1,0,1),\ \boldsymbol{u}_2=(0,1,2),\ \boldsymbol{u}_3=(2,1,0).$$

The concept of approximations

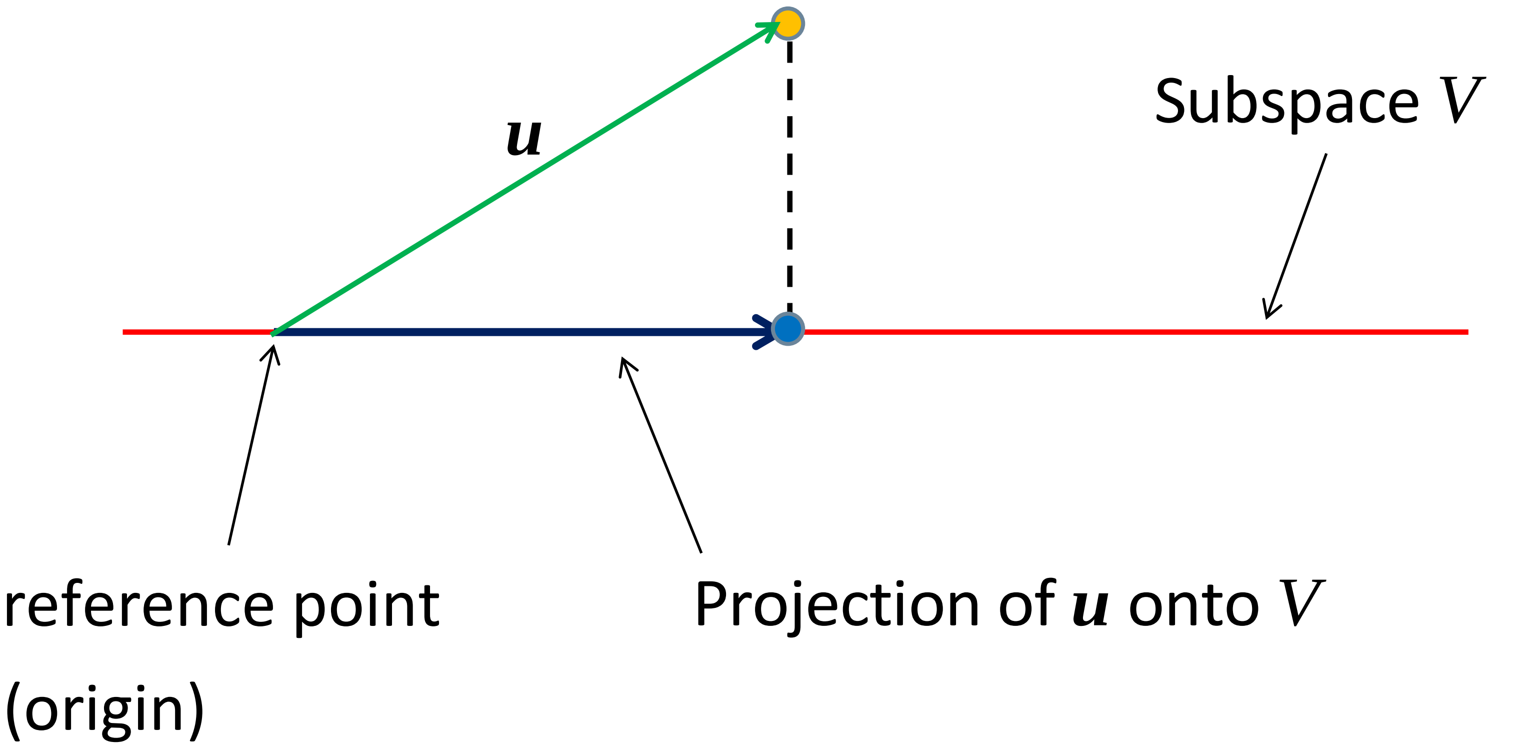
There are many computations where exact answers are not possible (or not necessary). This gives rise to the need for approximation.

The concept of orthogonality is central in the study of approximations.

Although the setting used here is the Euclidean space, the following discussions on approximations can be extended to general (abstract) vector spaces (e.g. functions).

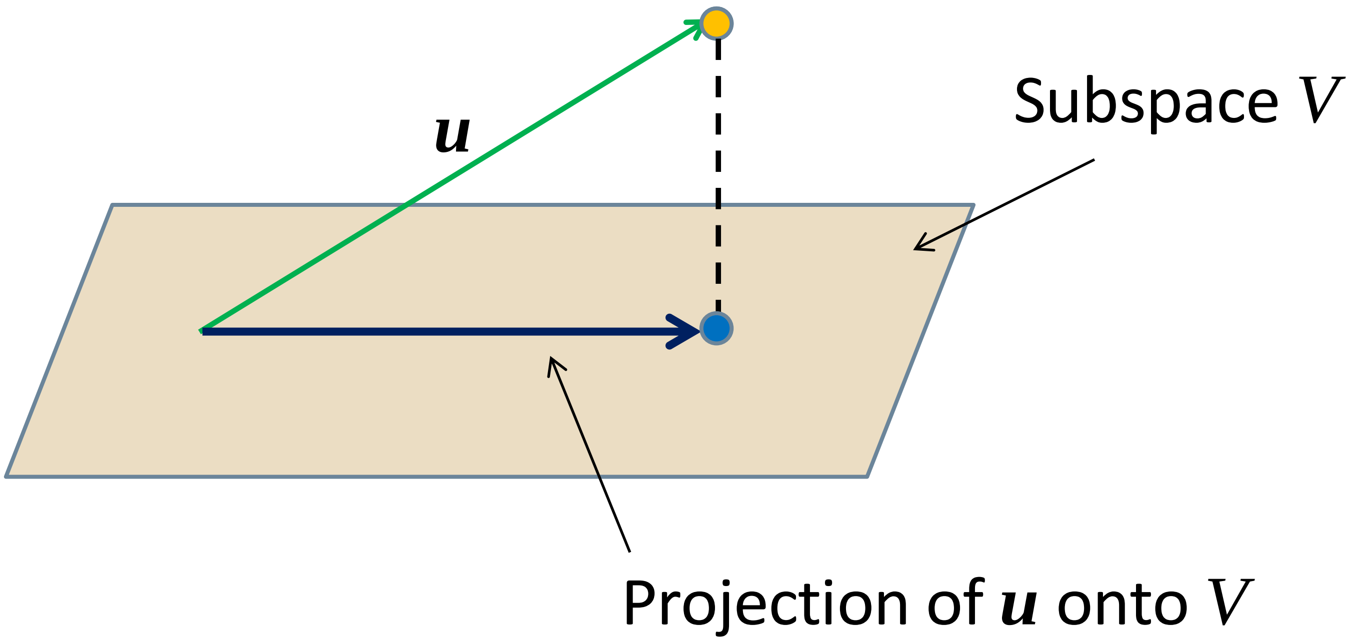
A simple question

Find a point on the line that is 'closest' to the given point.



A simple question

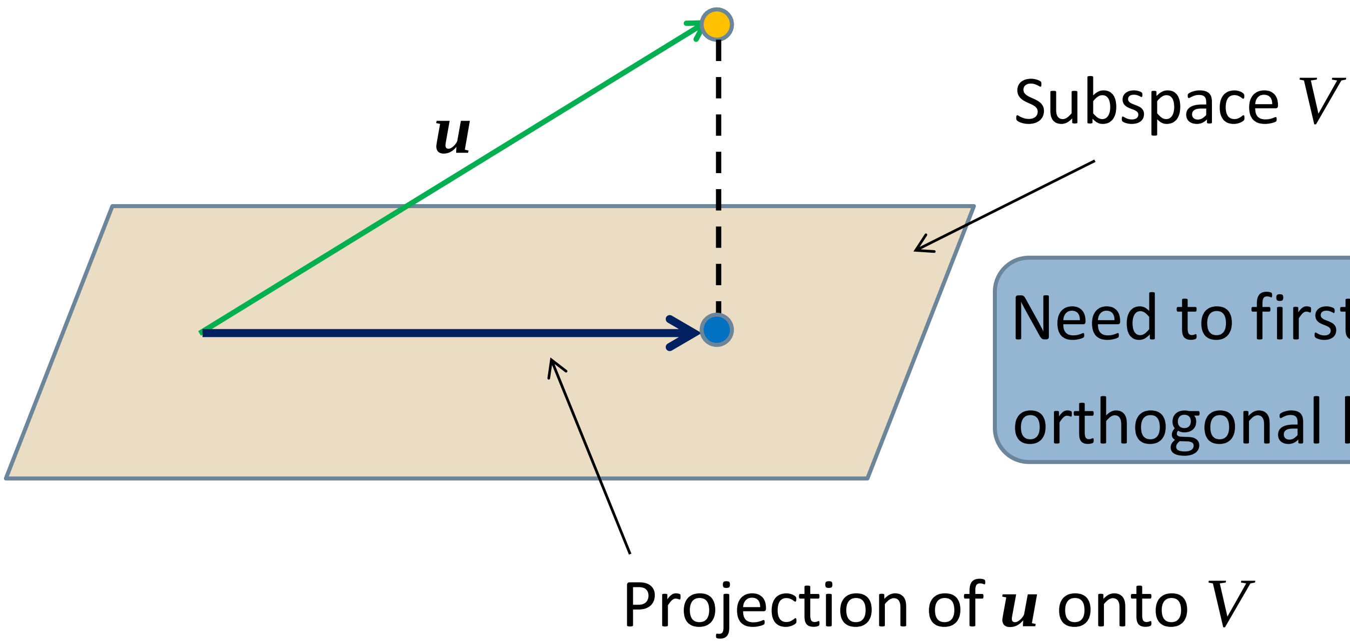
Find a point on the plane that is 'closest' to the given point.



Example

$V = \text{span}\{(1,0,1),(1,1,1)\}$ (a plane in $\mathbb{R}^3$ containing origin).

Find the (shortest) distance from $\boldsymbol{u} = (1,2,3)$ to V.



Need to first find an orthogonal basis for V.

Example

$V = \text{span}\{(1,0,1),(1,1,1)\}$ (a plane in $\mathbb{R}^3$ containing origin).

Find the (shortest) distance from $\boldsymbol{u} = (1,2,3)$ to V.

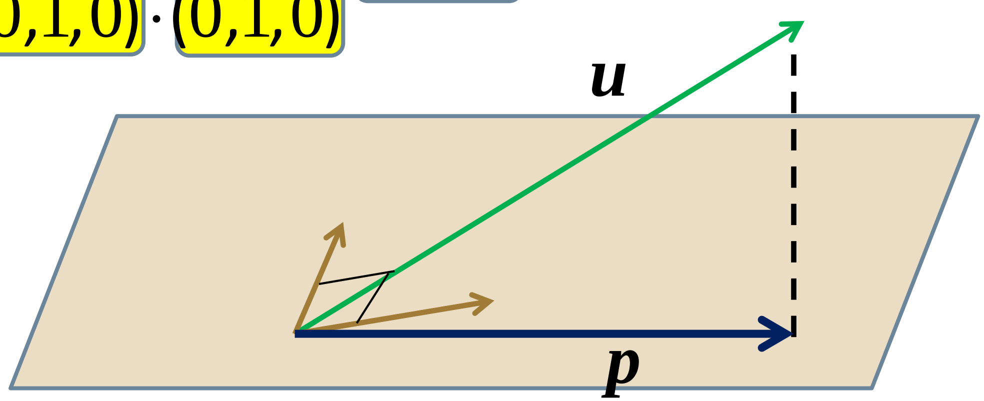
By Gram-Schmidt Process, $\boxed{(1,0,1)}$ and $\boxed{(0,1,0)}$

forms an orthogonal basis for V.

$$\boldsymbol{p} = \frac{(1,2,3)\cdot\boxed{(1,0,1)}}{\boxed{(1,0,1)}\cdot\boxed{(1,0,1)}}(1,0,1) + \frac{(1,2,3)\cdot\boxed{(0,1,0)}}{\boxed{(0,1,0)}\cdot\boxed{(0,1,0)}}\boxed{(0,1,0)} = (2,2,2)$$

$$\text{Distance} = \|\boldsymbol{u} - \boldsymbol{p}\|$$

$$= \|(-1,0,1)\| = \sqrt{2}$$

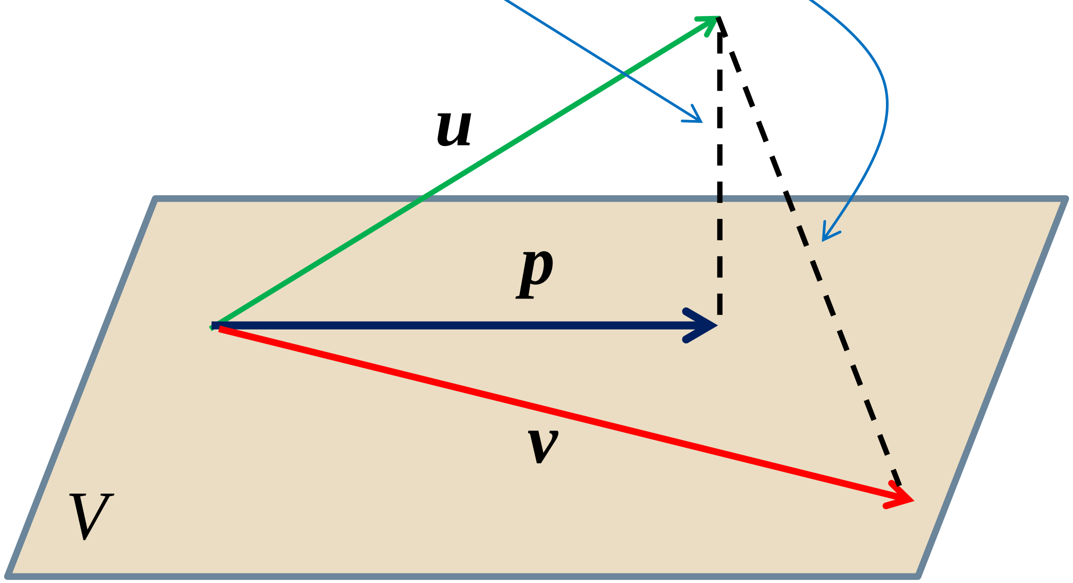


Theorem

Let V be a subspace in $\mathbb{R}^n$.

If $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{p}$ is the projection of $\boldsymbol{u}$ onto V, then

$$d(\boldsymbol{u}, \boldsymbol{p}) \leq d(\boldsymbol{u}, \boldsymbol{v}) \text{ for all } \boldsymbol{v} \in V.$$



that is, $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.

Proof: Omitted.

An example on experiments

You believe that physical quantities r, s and t are related according to the equation

$$t = cr^2 + ds + e$$

where constants c, d, e are to be determined.

A series of experiments were conducted to measure t given different values of r and s.

A total of six 'data points' are collected.

An example on experiments

A total of six 'data points' are collected.

$$t = cr^2 + ds + e$$

i	1	2	3	4	5	6
r_i	0	0	1	1	2	2
s_i	0	1	2	0	1	2
t_i	0.5	1.6	2.8	0.8	5.1	5.9

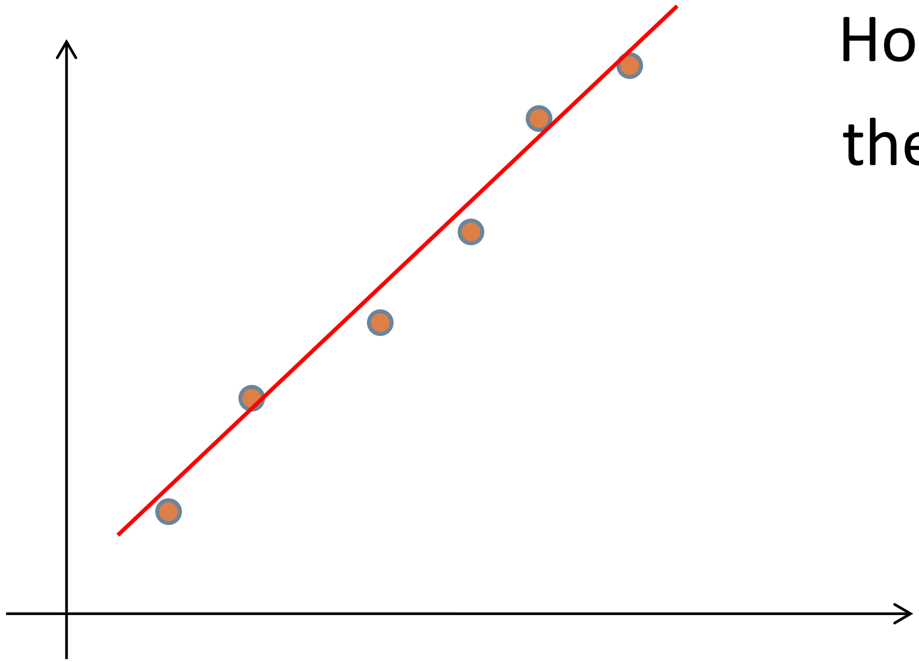
Can we find c, d, e such that

$t_i = cr_i^2 + ds_i + e$ for each $i = 1, \ldots, 6$?

Of course not!

An example on experiments

Can you draw a straight line that passes through all the points?



How do you draw the 'best' line?

An example on experiments

If there were no experimental errors, c, d, e would satisfy (solve) the following 6 equations:

$$\begin{cases} t_1 = cr_1^2 + ds_1 + e \\ t_2 = cr_2^2 + ds_2 + e \\ \qquad\vdots \\ t_6 = cr_6^2 + ds_6 + e \end{cases} \Leftrightarrow \begin{pmatrix} r_1^2 & s_1 & 1 \\ r_2^2 & s_2 & 1 \\ \vdots & \vdots & \vdots \\ r_6^2 & s_6 & 1 \end{pmatrix} \begin{pmatrix} c \\ d \\ e \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_6 \end{pmatrix} \Leftrightarrow \boldsymbol{Ax} = \boldsymbol{b}$$

$$\boldsymbol{A} = \begin{pmatrix} r_1^2 & s_1 & 1 \\ r_2^2 & s_2 & 1 \\ \vdots & \vdots & \vdots \\ r_6^2 & s_6 & 1 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} c \\ d \\ e \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_6 \end{pmatrix}$$

An example on experiments

Since there are no exact solutions for c, d, e, the linear system $\boldsymbol{Ax}=\boldsymbol{b}$ is inconsistent.

$$\begin{pmatrix} r_1^2 & s_1 & 1 \\ r_2^2 & s_2 & 1 \\ \vdots & \vdots & \vdots \\ r_6^2 & s_6 & 1 \end{pmatrix} \begin{pmatrix} c \\ d \\ e \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_6 \end{pmatrix} \Leftrightarrow \boldsymbol{Ax}=\boldsymbol{b}$$

We now seek to find c, d, e such that

observed

predicted

$$[\boxed{t_1} - \boxed{(cr_1^2 + ds_1 + e)}]^2 + [\boxed{t_2} - \boxed{(cr_2^2 + ds_2 + e)}]^2 + \ldots + [\boxed{t_6} - \boxed{(cr_6^2 + ds_6 + e)}]^2$$

$$= \sum_{i=1}^{6} [t_i - (cr_1^2 + ds_i + e)]^2$$

Sum of squares of errors

$$= \|(\boldsymbol{b} - \boldsymbol{Ax})\|^2$$

is minimized

An example on experiments

Summary of problem:

A linear system $\boldsymbol{Ax}=\boldsymbol{b}$ is inconsistent.

We wish to find $\boldsymbol{x}$ such that $\|(\boldsymbol{b}-\boldsymbol{Ax})\|^2$ (or equivalently) $\|(\boldsymbol{b}-\boldsymbol{Ax})\|$ is minimized.

Remark: If $\boldsymbol{Ax}=\boldsymbol{b}$ is consistent, then we simply choose $\boldsymbol{x}$ to be a solution so that $\boldsymbol{b}-\boldsymbol{Ax}=\boldsymbol{0}$ which means $\|(\boldsymbol{b}-\boldsymbol{Ax})\|=0$, the smallest possible value.

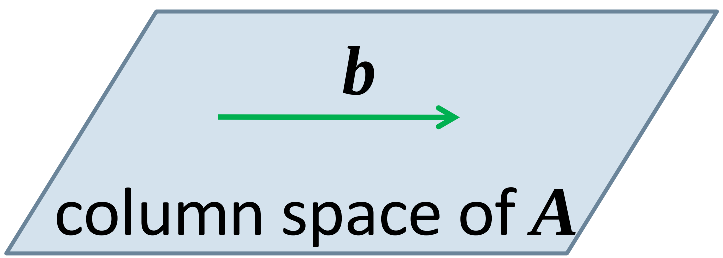
Definition (Least squares solution)

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system where $\boldsymbol{A}$ is a $m \times n$ matrix.

A vector $\boldsymbol{u} \in \mathbb{R}^n$ is called a least squares solution to the linear system if $\|\boldsymbol{b} - \boldsymbol{Au}\| \leq \|\boldsymbol{b} - \boldsymbol{Av}\|$ for all $\boldsymbol{v} \in \mathbb{R}^n$.

When is a linear system consistent?

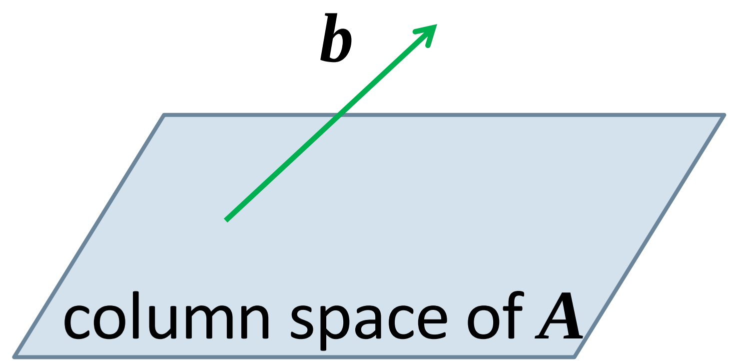
Recall that $\boldsymbol{Ax}=\boldsymbol{b}$ is consisent if and only if $\boldsymbol{b}$ belongs to the column space of $\boldsymbol{A}$.



$\boldsymbol{Ax}=\boldsymbol{b}$ is consistent.

Least squares solution

$=$Exact solution.

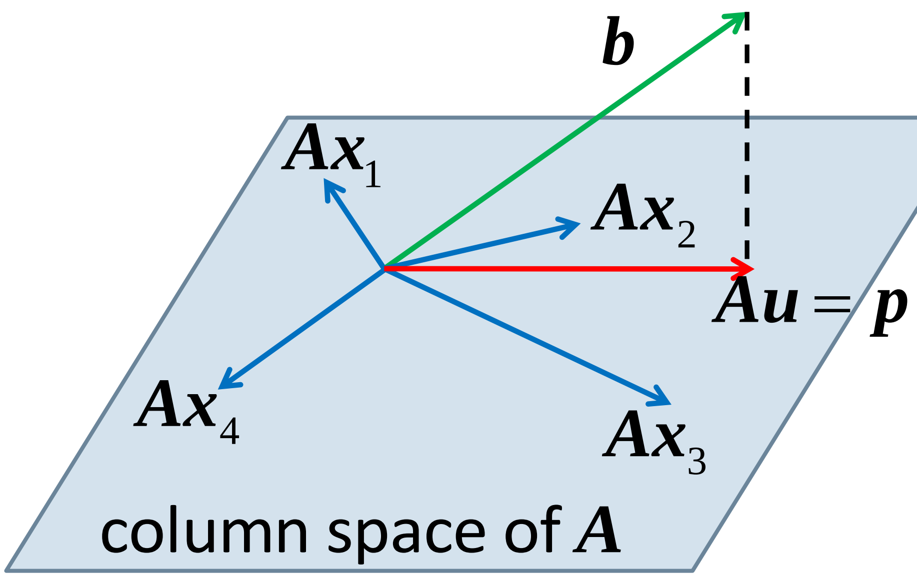


$\boldsymbol{Ax}=\boldsymbol{b}$ is inconsistent.

Least squares solution $=?$

When is a linear system consistent?

Recall that $\boldsymbol{Ax} = \boldsymbol{b}$ is consisent if and only if $\boldsymbol{b}$ belongs to the column space of $\boldsymbol{A}$.



Which $\boldsymbol{u}$ will be such that $\boldsymbol{Au}$ is 'closest' to $\boldsymbol{b}$?

Answer: An $\boxed{\boldsymbol{u} \text{ such that } \boldsymbol{Au} = \boldsymbol{p}}$ where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$.

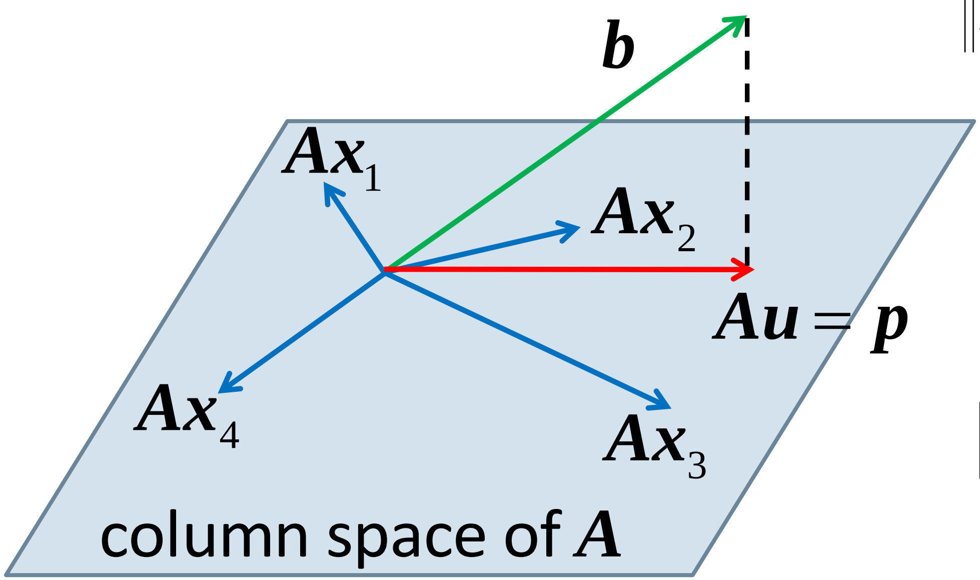
$$\boxed{\|(\boldsymbol{b} - \boldsymbol{Au})\| \le \|(\boldsymbol{b} - \boldsymbol{Av})\| \text{ for all } \boldsymbol{v} \in \mathbb{R}^n}$$

To find least squares solution $\boldsymbol{u}$, we solve $\boldsymbol{Ax} = \boldsymbol{p}$.

Theorem

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system, where $\boldsymbol{A}$ is an $m \times n$ matrix, and let $\boldsymbol{p}$ be the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$. Then

$$\boxed{\left\|(\boldsymbol{b}-\boldsymbol{p})\right\| \leq \left\|(\boldsymbol{b}-\boldsymbol{Av})\right\| \text{ for all } \boldsymbol{v} \in \mathbb{R}^n}$$



$$\left\|\boldsymbol{b}-\boldsymbol{p}\right\| = d(\boldsymbol{b},\boldsymbol{p}) \leq d(\boldsymbol{b},\boldsymbol{w}) = \left\|\boldsymbol{b}-\boldsymbol{w}\right\|$$

for all $\boldsymbol{w} \in V$ (column space of $\boldsymbol{A}$)

Since $V = \{\boldsymbol{Av} \mid \boldsymbol{v} \in \mathbb{R}^n\}$

$$\left\|(\boldsymbol{b}-\boldsymbol{p})\right\| \leq \left\|(\boldsymbol{b}-\boldsymbol{Av})\right\| \text{ for all } \boldsymbol{v} \in \mathbb{R}^n$$

Example

$$\boldsymbol{A}=\begin{pmatrix}1 & 1\\0 & 1\\1 & 1\end{pmatrix},\boldsymbol{b}=\begin{pmatrix}1\\2\\3\end{pmatrix} \text{ and } V=\text{span}\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$$

(column space of $\boldsymbol{A}$)

By earlier example, projection of $\boldsymbol{b}$ onto V is $\boldsymbol{p}=\begin{pmatrix}2\\2\\2\end{pmatrix}$.

$\begin{pmatrix}x\\y\end{pmatrix}$ is a least squares solution if and only if

$$\boldsymbol{A}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2\\2\\2\end{pmatrix}\Leftrightarrow\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\2\end{pmatrix}.$$

END OF LECTURE 18

Lecture 19

Best approximation (cont'd)

Orthogonal matrices (till end of Chapter 5)

Orthogonal diagonalization (Section 6.3)