Best approximation

The concept of approximations

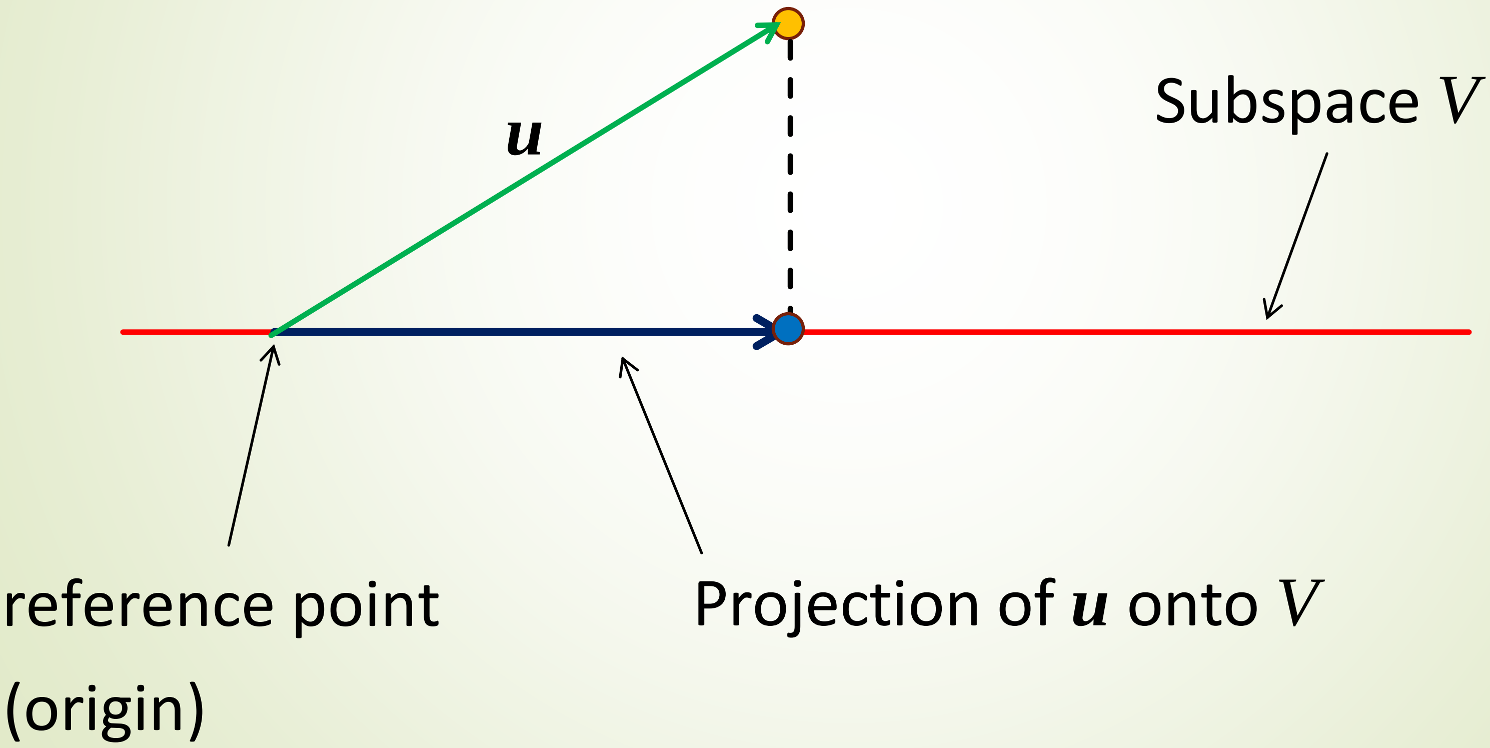
There are many computations where exact answers are not possible (or not necessary). This gives rise to the need for approximation.

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).

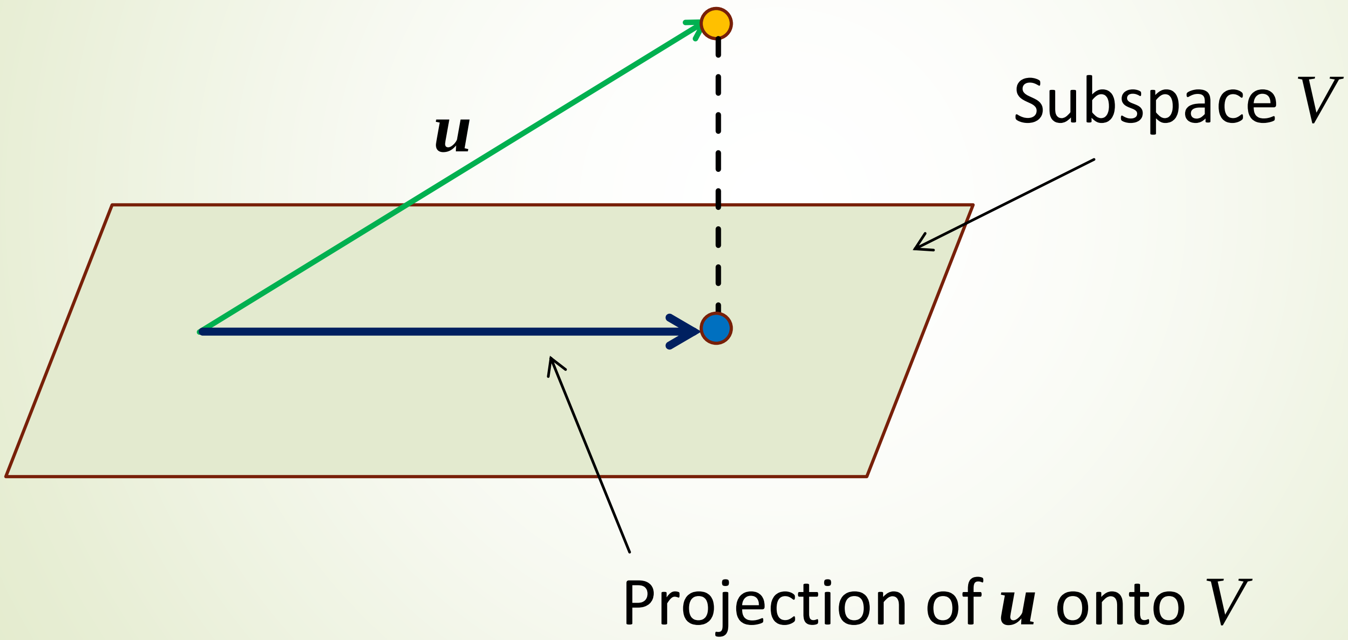
A simple question

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



A simple question

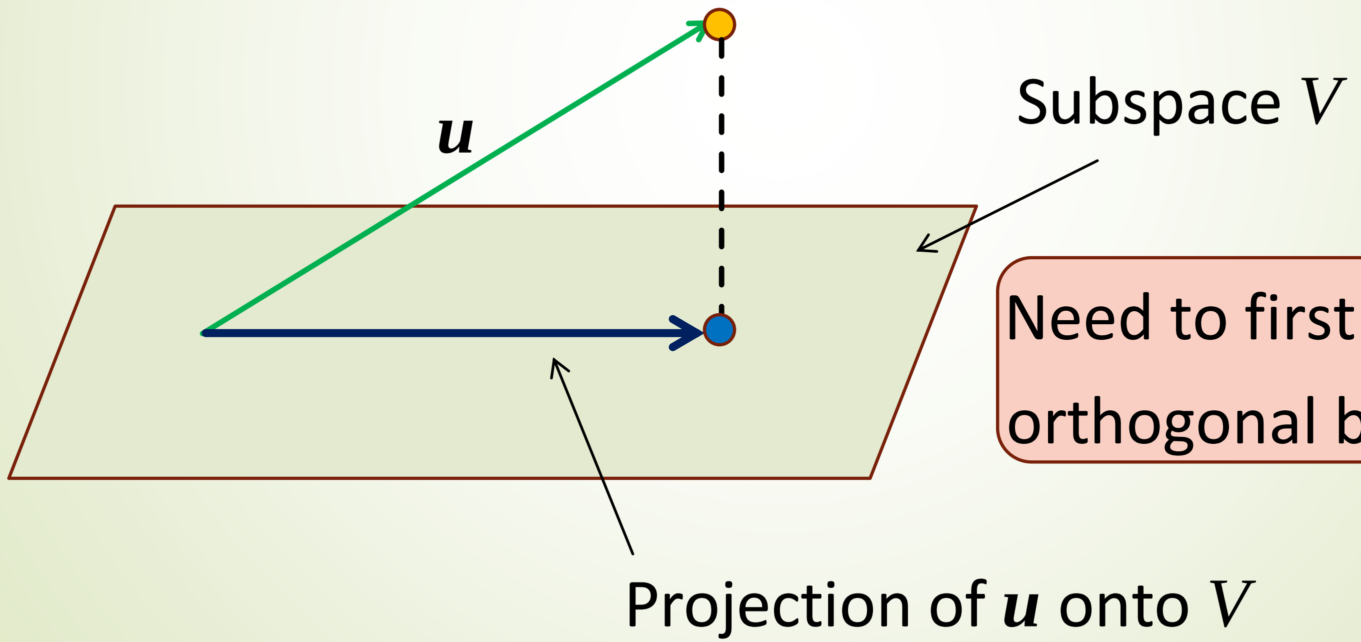
Find a point on the plane that is 'closest' to the given point.



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.

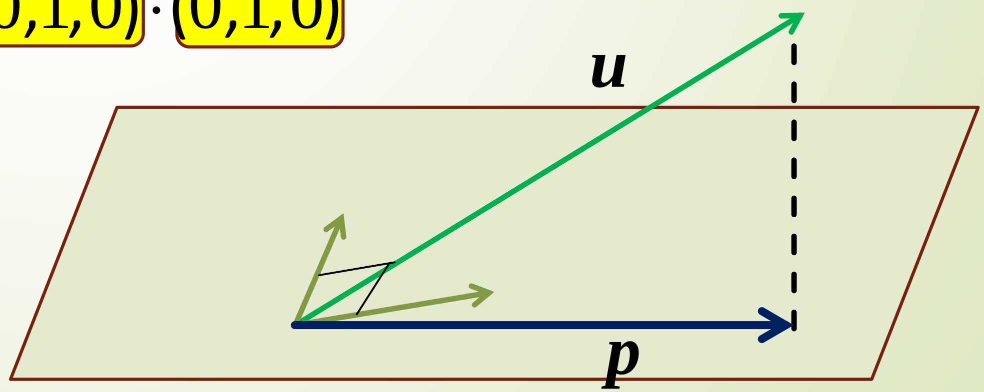
By Gram-Schmidt Process, $\boxed{(1,0,1)}$ and $\boxed{(0,1,0)}$

forms an orthogonal basis for V.

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

Distance $= \left\|\boldsymbol{u} - \boldsymbol{p}\right\|$

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

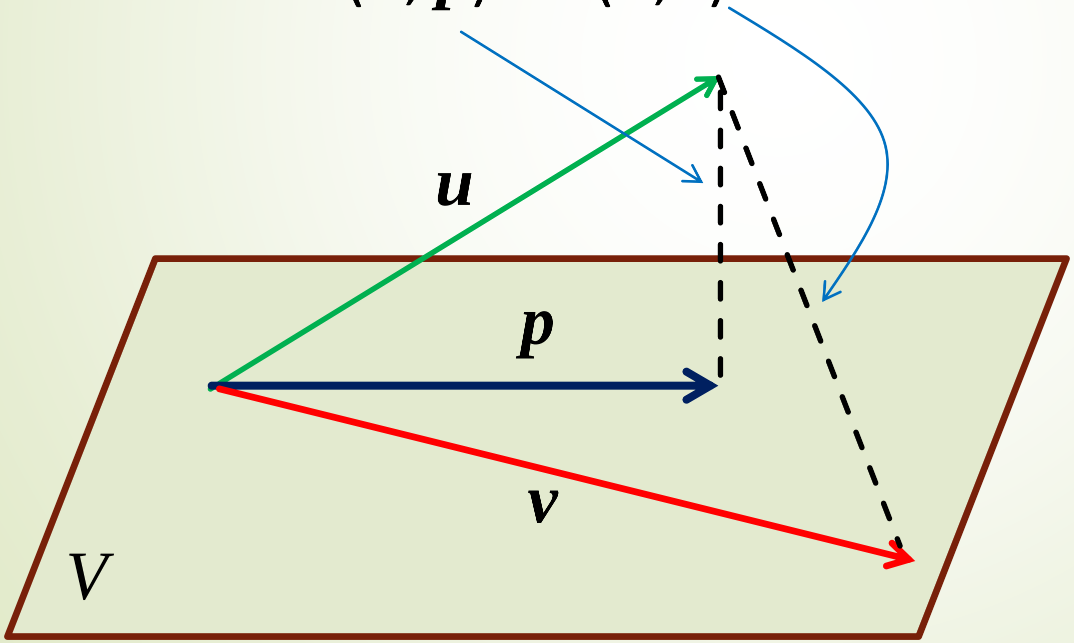


Theorem

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

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



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

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

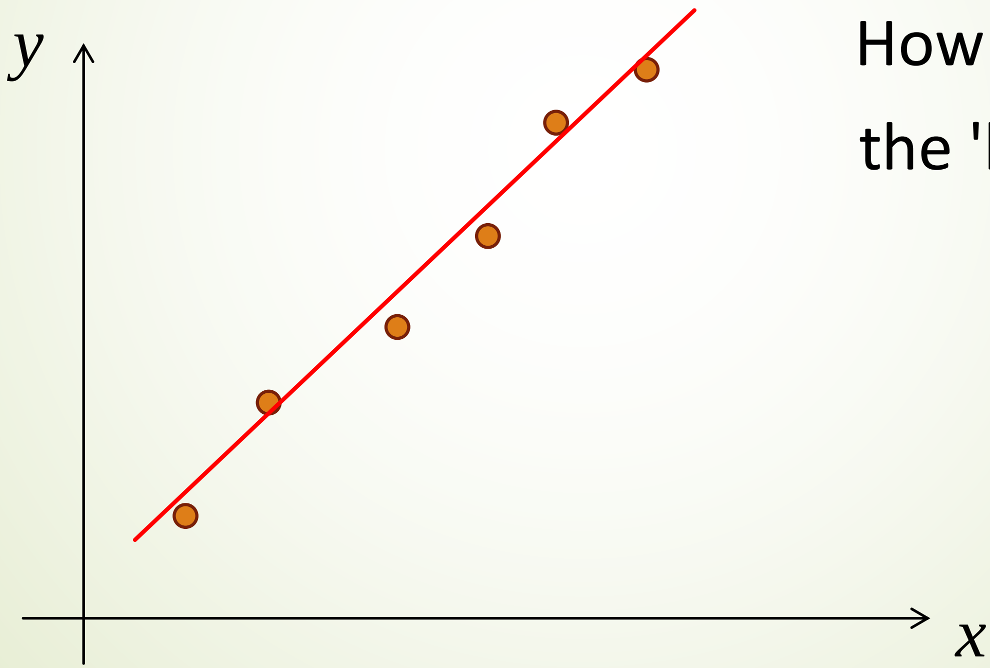
Can we find c, d, e such that

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

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:

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

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

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

$$[t_1 - (cr_1^2 + ds_1 + e)]^2 + [t_2 - (cr_2^2 + ds_2 + e)]^2 + \dots + [t_6 - (cr_6^2 + ds_6 + e)]^2 = \sum_{i=1}^{6} [t_i - (cr_1^2 + ds_i + e)]^2 = \|(\boldsymbol{b} - \boldsymbol{Ax})\|^2$$

is minimized

observed

predicted

Sum of squares of errors

An example on experiments

Summary of problem:

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

We wish to find $\boldsymbol{x}$ such that $\left\|(\boldsymbol{b} - \boldsymbol{Ax})\right\|^2$ (or equivalently) $\left\|(\boldsymbol{b} - \boldsymbol{Ax})\right\|$ 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 $\left\|(\boldsymbol{b} - \boldsymbol{Ax})\right\| = 0$, 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 $\|\boldsymbol{b} - \boldsymbol{Au}\| \le \|\boldsymbol{b} - \boldsymbol{Av}\|$ for all $\boldsymbol{v} \in \mathbb{R}^n$.

Summary

1) If V is a subspace of $\mathbb{R}^n$ and $\boldsymbol{u}$ is a vector in $\mathbb{R}^n$ then the projection of $\boldsymbol{u}$ onto V is the best approximation of $\boldsymbol{u}$ in V.

2) Definition of a least squares solution to a linear system $\boldsymbol{Ax} = \boldsymbol{b}$.