$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}_{m\times n}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m\times 1}}_{b}$$

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.5**

Lecture : **Orthogonality**

Topic : **Orthogonality**

Concept : **Orthogonal Decomposition**

Instructor: **A/P Chng Eng Siong**

TAs: **Zhang Su, Vishal Choudhari**

Rev: 26th June 2020

Orthogonal Decomposition

THEOREM 8

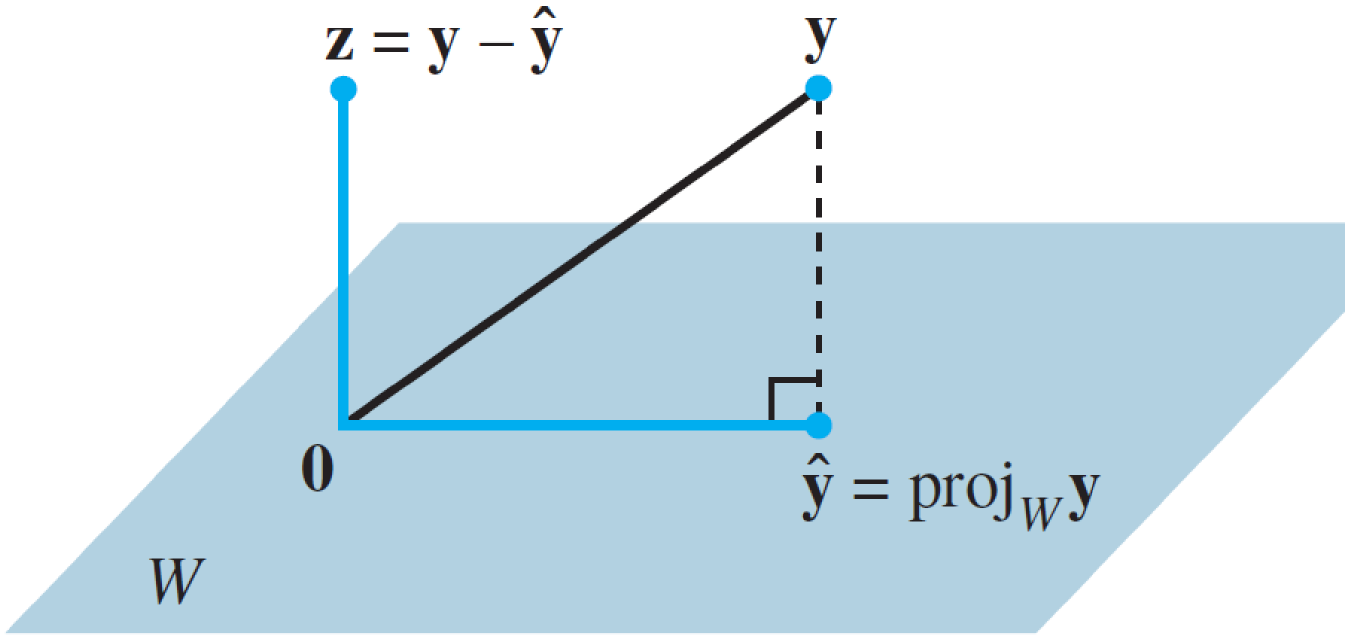


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of y onto** W and often is written as $\text{proj}_W \mathbf{y}$. See Fig. 2.

The Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p \tag{2}$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

Note: $W^\perp$ is the set of all vectors orthogonal to the subspace W.

Lay's Linear Algebra and Applications

348 CHAPTER 6 Orthogonality and Least Squares

[1]We may assume that W is not the zero subspace, for otherwise $W^\perp = \mathbb{R}^n$ and (1) is simply $\mathbf{y} = \mathbf{0} + \mathbf{y}$.

PROOF Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be any orthogonal basis for W, and define $\hat{\mathbf{y}}$ by (2).[1] Then $\hat{\mathbf{y}}$ is in W because $\hat{\mathbf{y}}$ is a linear combination of the basis $\mathbf{u}_1, \dots, \mathbf{u}_p$. Let $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Since $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$, it follows from (2) that

$$\begin{aligned} \mathbf{z} \cdot \mathbf{u}_1 = (\mathbf{y} - \hat{\mathbf{y}}) \cdot \mathbf{u}_1 &= \mathbf{y} \cdot \mathbf{u}_1 - \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 \cdot \mathbf{u}_1 - 0 - \cdots - 0 \\ &= \mathbf{y} \cdot \mathbf{u}_1 - \mathbf{y} \cdot \mathbf{u}_1 = 0 \end{aligned}$$

Thus $\mathbf{z}$ is orthogonal to $\mathbf{u}_1$. Similarly, $\mathbf{z}$ is orthogonal to each $\mathbf{u}_j$ in the basis for W. Hence $\mathbf{z}$ is orthogonal to every vector in W. That is, $\mathbf{z}$ is in $W^\perp$.

To show that the decomposition in (1) is unique, suppose $\mathbf{y}$ can also be written as $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, with $\hat{\mathbf{y}}_1$ in W and $\mathbf{z}_1$ in $W^\perp$. Then $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ (since both sides equal $\mathbf{y}$), and so

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

This equality shows that the vector $\mathbf{v} = \hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ is in W and in $W^\perp$ (because $\mathbf{z}_1$ and $\mathbf{z}$ are both in $W^\perp$, and $W^\perp$ is a subspace). Hence $\mathbf{v} \cdot \mathbf{v} = 0$, which shows that $\mathbf{v} = \mathbf{0}$. This proves that $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1$ and also $\mathbf{z}_1 = \mathbf{z}$. ■

Example

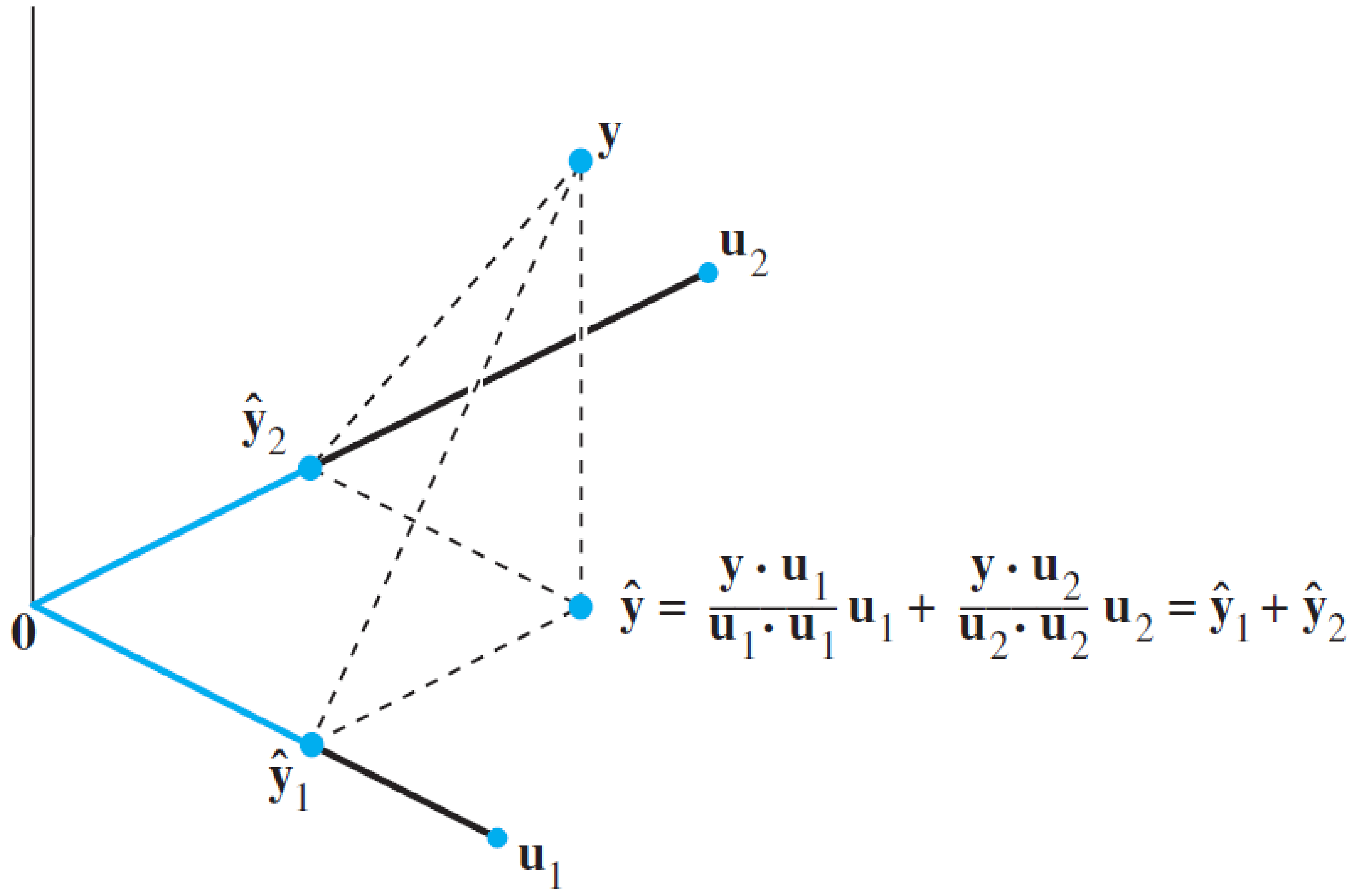


FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

Figure 3 W is a subspace of $\mathbb{R}^3$ spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Here $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$ denote the projections of $\mathbf{y}$ onto the lines spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively. The orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto W is the sum of the projections of $\mathbf{y}$ onto one-dimensional subspaces that are orthogonal to each other.

Lay's Linear Algebra and Applications

348 CHAPTER 6 Orthogonality and Least Squares

EXAMPLE 2 Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

SOLUTION The orthogonal projection of $\mathbf{y}$ onto W is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$$

$$= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

Also

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Theorem 8 ensures that $\mathbf{y} - \hat{\mathbf{y}}$ is in $W^{\perp}$. To check the calculations, however, it is a good idea to verify that $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$ and hence to all of W. The desired decomposition of $\mathbf{y}$ is

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Best Approximation Theorem

THEOREM 9

The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{3}$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

Properties of Orthogonal Projections

If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal basis for W and if $\mathbf{y}$ happens to be in W, then the formula for $\text{proj}_W\, \mathbf{y}$ is exactly the same as the representation of $\mathbf{y}$ given in Theorem 5 in Section 6.2. In this case, $\text{proj}_W\, \mathbf{y} = \mathbf{y}$.

If $\mathbf{y}$ is in $W = \text{Span}\,\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$, then $\text{proj}_W\, \mathbf{y} = \mathbf{y}$.

Lay's Linear Algebra and Applications

350 CHAPTER 6 Orthogonality and Least Squares

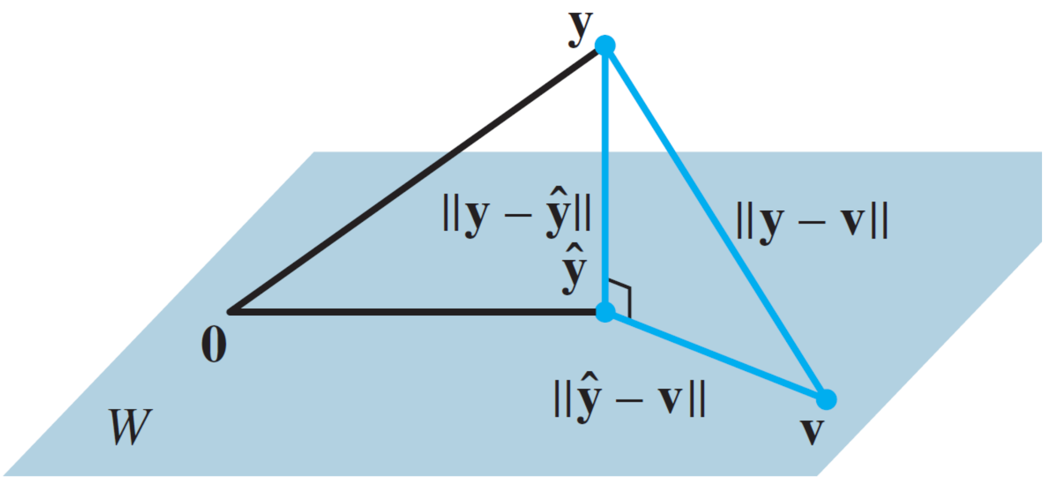


FIGURE 4 The orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$.

The vector $\hat{\mathbf{y}}$ in Theorem 9 is called **the best approximation to y by elements of** W.

PROOF Take $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$. See Fig. 4. Then $\hat{\mathbf{y}} - \mathbf{v}$ is in W. By the Orthogonal Decomposition Theorem, $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W. In particular, $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\hat{\mathbf{y}} - \mathbf{v}$ (which is in W). Since

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v})$$

the Pythagorean Theorem gives

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$$

(See the colored right triangle in Fig. 4. The length of each side is labeled.) Now $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ because $\hat{\mathbf{y}} - \mathbf{v} \neq \mathbf{0}$, and so inequality (3) follows immediately. ∎

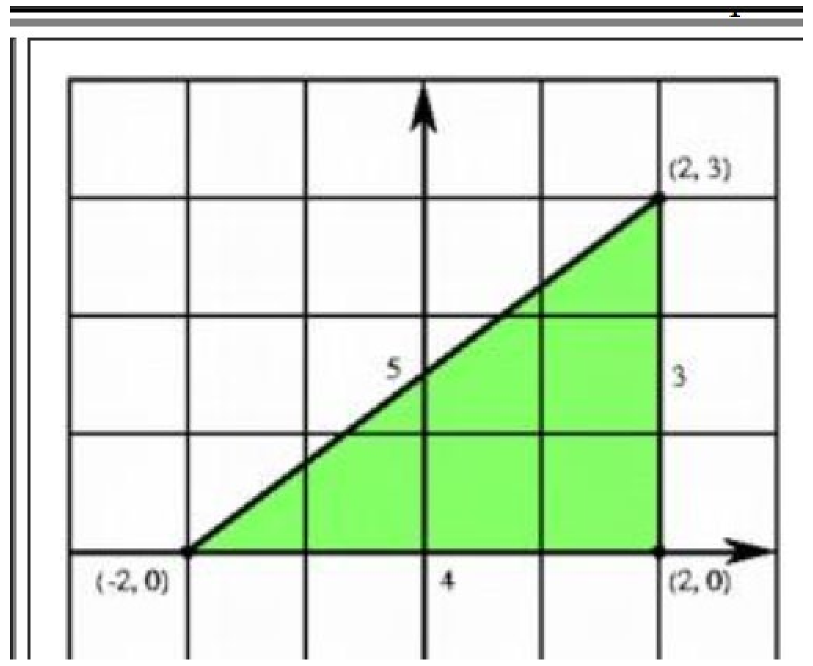
Conclusion: if y is in the space of W, then we can show $\hat{y}$ to be equal to y using theorem 8. Else, it's the best approximation according to theorem 9.

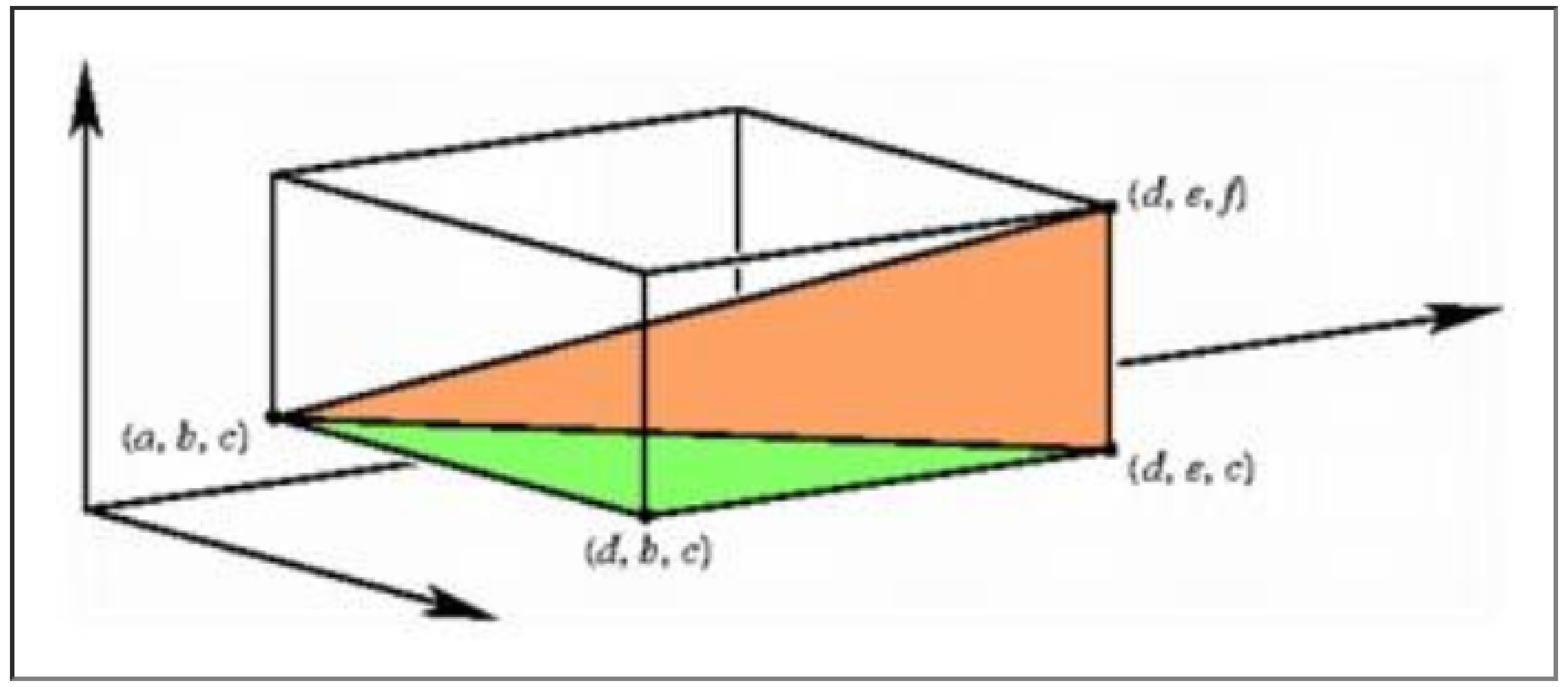
Using Pythagoras Theorem in Higher Dimensions

Ref:

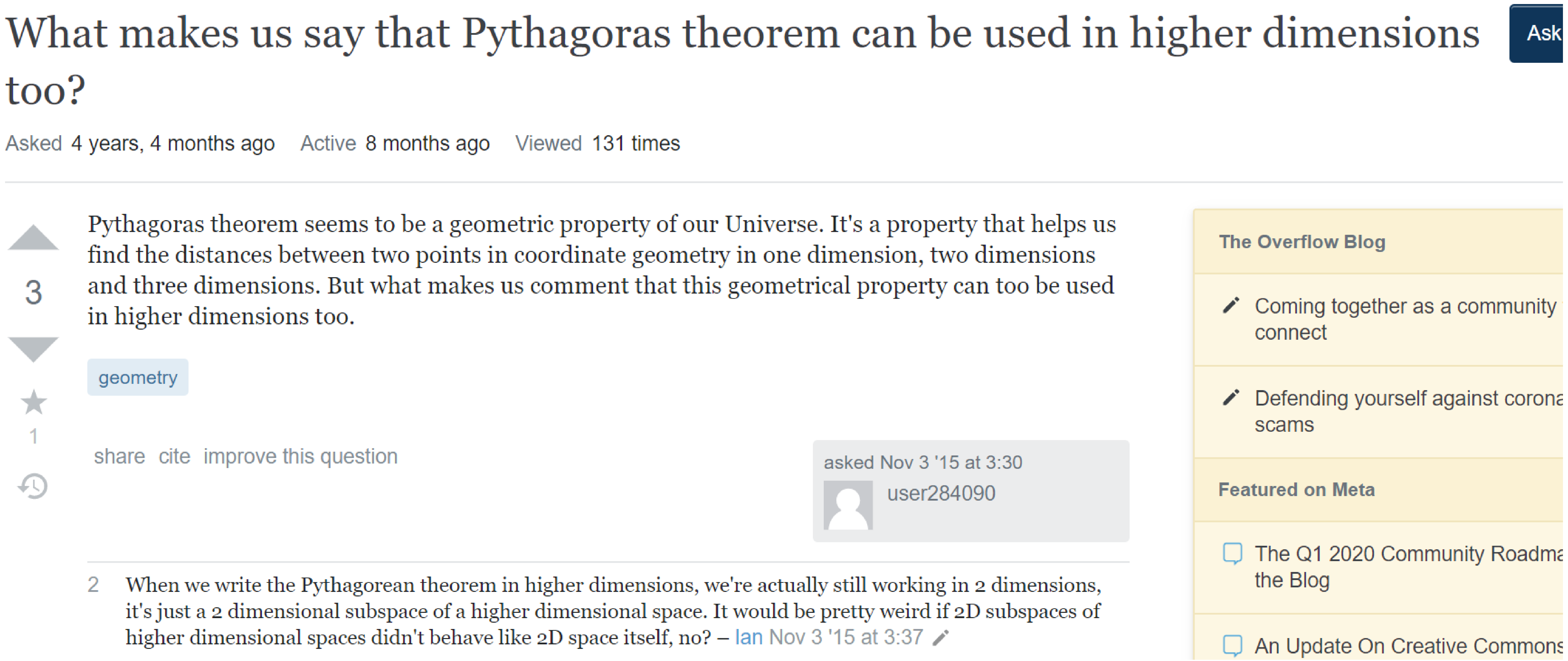
i) https://math.stackexchange.com/questions/1510549/what-makes-us-say-that-pythagoras-theorem-can-be-used-in-higher-dimensions-too

ii) http://www.math.brown.edu/~banchoff/Beyond3d/chapter8/section02.html





To find the distance formula in three-space, we apply the planar distance formula twice, once to find the length of the diagonal of one of the faces, and then to find the hypotenuse of a right triangle having this diagonal and an edge of the cube as its sides.



What makes us say that Pythagoras theorem can be used in higher dimensions too?

Asked 4 years, 4 months ago Active 8 months ago Viewed 131 times

3

Pythagoras theorem seems to be a geometric property of our Universe. It's a property that helps us find the distances between two points in coordinate geometry in one dimension, two dimensions and three dimensions. But what makes us comment that this geometrical property can too be used in higher dimensions too.

geometry

share cite improve this question

asked Nov 3 '15 at 3:30
user284090

2 When we write the Pythagorean theorem in higher dimensions, we're actually still working in 2 dimensions, it's just a 2 dimensional subspace of a higher dimensional space. It would be pretty weird if 2D subspaces of higher dimensional spaces didn't behave like 2D space itself, no? – Ian Nov 3 '15 at 3:37

The Overflow Blog

Coming together as a community connect

Defending yourself against corona scams

Featured on Meta

The Q1 2020 Community Roadma the Blog

An Update On Creative Commons

Example

EXAMPLE 3 If $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, as in Example 2, then the closest point in W to $\mathbf{y}$ is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

■

EXAMPLE 4 The distance from a point $\mathbf{y}$ in $\mathbb{R}^n$ to a subspace W is defined as the distance from $\mathbf{y}$ to the nearest point in W. Find the distance from $\mathbf{y}$ to $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

SOLUTION By the Best Approximation Theorem, the distance from $\mathbf{y}$ to W is $\|\mathbf{y} - \hat{\mathbf{y}}\|$, where $\hat{\mathbf{y}} = \text{proj}_W\,\mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$

The distance from $\mathbf{y}$ to W is $\sqrt{45} = 3\sqrt{5}$. ■

Best Approximation Theorem

THEOREM 10

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W\, \mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p \tag{4}$$

If $U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p\,]$, then

$$\text{proj}_W\, \mathbf{y} = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \tag{5}$$

$Note: p < n$

PROOF Formula (4) follows immediately from (2) in Theorem 8. Also, (4) shows that $\text{proj}_W\,\mathbf{y}$ is a linear combination of the columns of U using the weights $\mathbf{y}\cdot\mathbf{u}_1$, $\mathbf{y}\cdot\mathbf{u}_2, \ldots, \mathbf{y}\cdot\mathbf{u}_p$. The weights can be written as $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, showing that they are the entries in $U^T\mathbf{y}$ and justifying (5). ∎

Lay's Linear Algebra and Applications

Suppose U is an $n \times p$ matrix with orthonormal columns, and let W be the column space of U. Then

$$U^TU\mathbf{x} = I_p\mathbf{x} = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^p \qquad \text{Theorem 6}$$

$$UU^T\mathbf{y} = \text{proj}_W\,\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \qquad \text{Theorem 10}$$

If U is an $n \times n$ (square) matrix with orthonormal columns, then U is an *orthogonal* matrix, the column space W is all of $\mathbb{R}^n$, and $UU^T\mathbf{y} = I\mathbf{y} = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^n$.

Although formula (4) is important for theoretical purposes, in practice it usually involves calculations with square roots of numbers (in the entries of the $\mathbf{u}_i$). Formula (2) is recommended for hand calculations.

PRACTICE PROBLEM

Let $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Use the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $\text{proj}_W\,\mathbf{y}$.

Note: above u_1 ,and u_2 are orthogonal, but their norm is not 1.
To use
    Theorem 10 (orthonormality is needed).
    You need to convert u_1 and u_2 to normalized col
Else
    Use theorem 8 (only orthogonality is needed).

Example

PRACTICE PROBLEM

Let $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Use the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $\text{proj}_W\, \mathbf{y}$.

example4_Walsh.m | example5_orthogonalSet.m | example6_Lay_pg352.m

```
% Lay Pg 352, orthogonality, projecting y onto U
U = [- 7 -1; 1 1; 4 -2];
y = [-9 1 6]'

U  % sanity check
y  % to see the values

orthogonalityCheck = U'*U   % quick way to find transpose column of U
    % and dot producting its other column

u1 = U(:,1) % extracting u1 and u2 from the column of U
u2 = U(:,2)

u1'*u2   % checking u1 is orthogonal to u2 using dot product

Proj_y_onto_u1 = y'*u1/(u1'*u1)   % remember to normalize using denorminator
Proj_y_onto_u2 = y'*u2/(u2'*u2)   % to make it unit vector

est_y = Proj_y_onto_u1*u1 + Proj_y_onto_u2*u2
est_y - y


% Using projection matrix to find est_y2
U_norm(:,1) = u1/sqrt((u1'*u1))
U_norm(:,2) = u2/sqrt((u2'*u2))

est_y = U_norm*U_norm'*y


y2 = [-8.5, 1,6]'
est_y2 = U_norm*U_norm'*y2
error = y2 - est_y2
```

Screen shot showing the answer:

```
U =

    -7    -1
     1     1
     4    -2


y =

    -9
     1
     6


orthogonalityCheck =

    66     0
     0     6
```

```
Proj_y_onto_u1 =

    1.3333


Proj_y_onto_u2 =

   -0.3333


est_y =

   -9.0000
    1.0000
    6.0000

ans =

   1.0e-14 *

    0.1776
         0
         0
```

```
U_norm =

   -0.8616   -0.4082
    0.1231    0.4082
    0.4924   -0.8165


est_y =

   -9.0000
    1.0000
    6.0000
```

```
y2 =

   -8.5000
    1.0000
    6.0000


est_y2 =

   -8.5455
    0.8636
    5.9545


error =

    0.0455
    0.1364
    0.0455
```

Lay's Linear Algebra and Applications

350 CHAPTER 6 Orthogonality and Least Squares