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

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.1 (mod by Tay Kian Boon)**

Lecture : **Orthogonality**

Topic : **Orthogonality**

Concept : **Definition of Orthogonality and Orthogonal Complements**

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

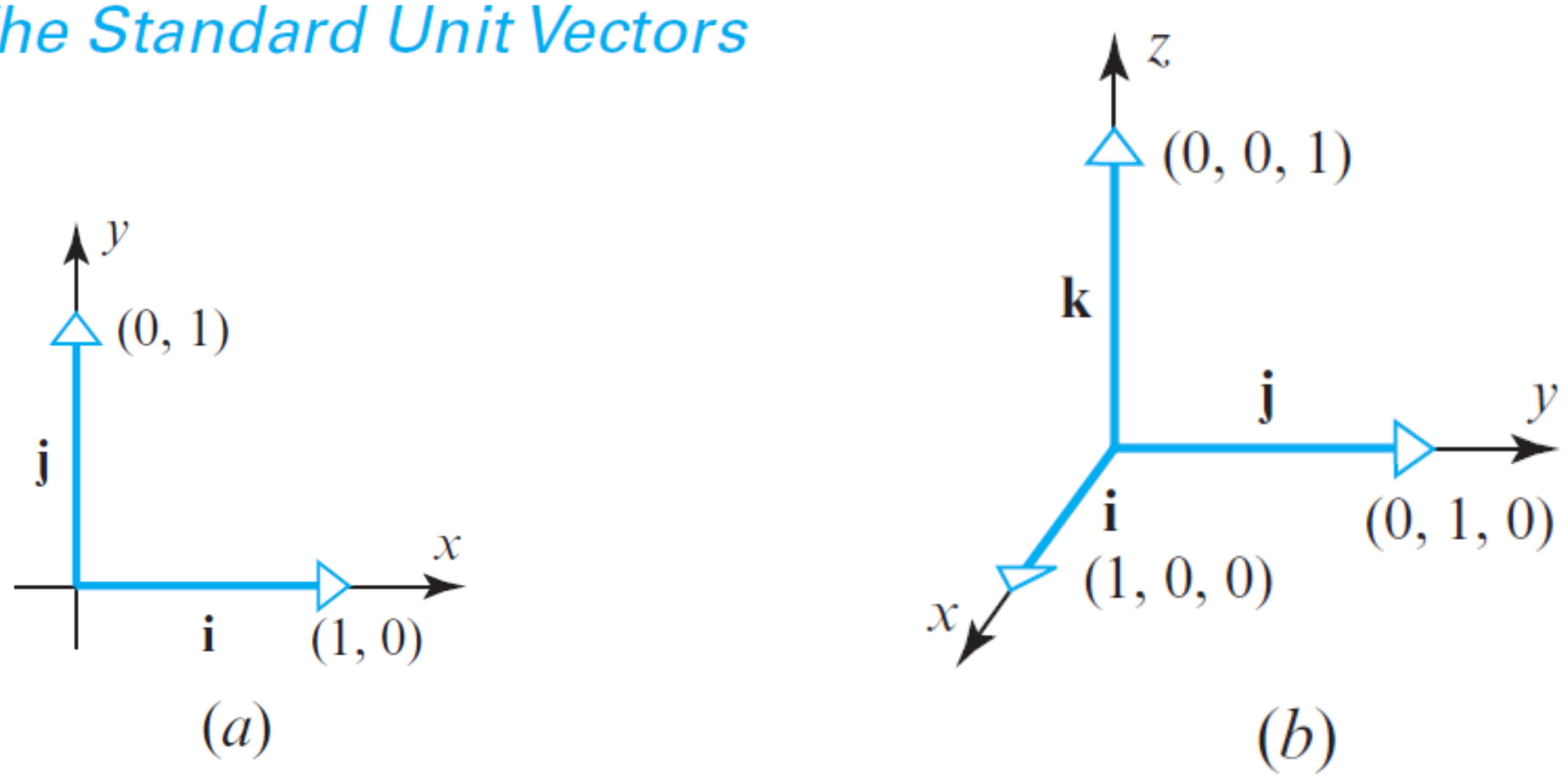
TAs: **Zhang Su, Vishal Choudhari**

LAST REVISED: 25 Sep 2021

Rev: 26th June 2020

Orthogonality Definition

The Standard Unit Vectors



▲ Figure 3.2.2

When a rectangular coordinate system is introduced in R^2 or R^3, the unit vectors in the positive directions of the coordinate axes are called the ***standard unit vectors***. In R^2 these vectors are denoted by

$$\mathbf{i} = (1, 0) \quad \text{and} \quad \mathbf{j} = (0, 1)$$

and in R^3 by

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \text{and} \quad \mathbf{k} = (0, 0, 1)$$

(Figure 3.2.2). Every vector $\mathbf{v} = (v_1, v_2)$ in R^2 and every vector $\mathbf{v} = (v_1, v_2, v_3)$ in R^3 can be expressed as a linear combination of standard unit vectors by writing

$$\mathbf{v} = (v_1, v_2) = v_1(1, 0) + v_2(0, 1) = v_1\mathbf{i} + v_2\mathbf{j}$$

$$\mathbf{v} = (v_1, v_2, v_3) = v_1(1, 0, 0) + v_2(0, 1, 0) + v_3(0, 0, 1) = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

DEFINITION 1 Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n are said to be ***orthogonal*** (or ***perpendicular***) if $\mathbf{u} \cdot \mathbf{v} = 0$. We will also agree that the zero vector in R^n is orthogonal to *every* vector in R^n.

Recall from Formula (20) in the previous section that the angle θ between two *nonzero* vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n is defined by the formula

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$$

It follows from this that $\theta = \pi/2$ if and only if $\mathbf{u} \cdot \mathbf{v} = 0$. Thus, we make the following definition.

▶ **EXAMPLE 1 Orthogonal Vectors**

(a) Show that $\mathbf{u} = (-2, 3, 1, 4)$ and $\mathbf{v} = (1, 2, 0, -1)$ are orthogonal vectors in R^4.

(b) Let $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be the set of standard unit vectors in R^3. Show that each ordered pair of vectors in S is orthogonal.

Solution (a) The vectors are orthogonal since

$$\mathbf{u} \cdot \mathbf{v} = (-2)(1) + (3)(2) + (1)(0) + (4)(-1) = 0$$

Solution (b) It suffices to show that

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k} = \mathbf{j} \cdot \mathbf{k} = 0$$

because it will follow automatically from the symmetry property of the dot product that

$$\mathbf{j} \cdot \mathbf{i} = \mathbf{k} \cdot \mathbf{i} = \mathbf{k} \cdot \mathbf{j} = 0$$

Although the orthogonality of the vectors in S is evident geometrically from Figure 3.2.2, it is confirmed algebraically by the computations

$$\mathbf{i} \cdot \mathbf{j} = (1, 0, 0) \cdot (0, 1, 0) = 0$$

$$\mathbf{i} \cdot \mathbf{k} = (1, 0, 0) \cdot (0, 0, 1) = 0$$

$$\mathbf{j} \cdot \mathbf{k} = (0, 1, 0) \cdot (0, 0, 1) = 0 \blacktriangleleft$$

Lines and Planes Determined by Points and Normals

One learns in analytic geometry that a line in R^2 is determined uniquely by its slope and one of its points, and that a plane in R^3 is determined uniquely by its "inclination" and one of its points. One way of specifying slope and inclination is to use a *nonzero* vector $\mathbf{n}$, called a ***normal***, that is orthogonal to the line or plane in question. For example, Figure 3.3.1 shows the line through the point $P_0(x_0, y_0)$ that has normal $\mathbf{n} = (a, b)$ and the plane through the point $P_0(x_0, y_0, z_0)$ that has normal $\mathbf{n} = (a, b, c)$. Both the line and the plane are represented by the vector equation

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0 \qquad (1)$$

where P is either an arbitrary point (x, y) on the line or an arbitrary point (x, y, z) in the plane. The vector $\overrightarrow{P_0P}$ can be expressed in terms of components as

$$\overrightarrow{P_0P} = (x - x_0,\ y - y_0) \quad \textbf{[line]}$$

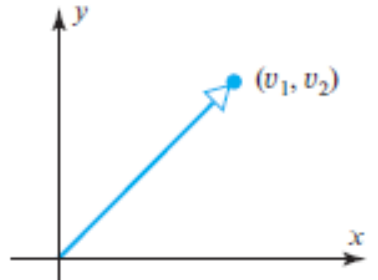
$$\overrightarrow{P_0P} = (x - x_0,\ y - y_0,\ z - z_0) \quad \textbf{[plane]}$$

Thus, Equation (1) can be written as

$$a(x - x_0) + b(y - y_0) = 0 \quad \textbf{[line]} \qquad (2)$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \textbf{[plane]} \qquad (3)$$

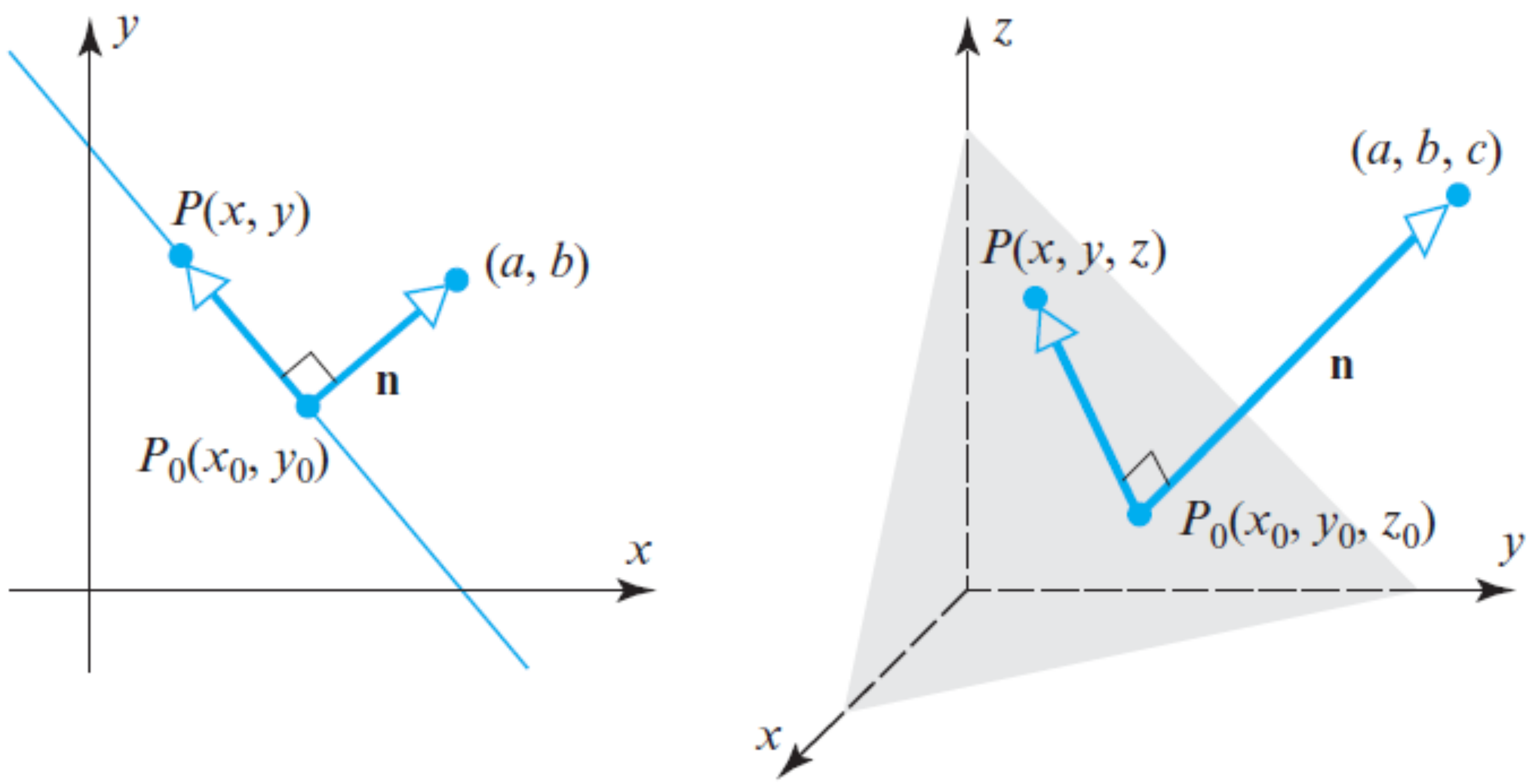
These are called the ***point-normal*** equations of the line and plane.



▲ Figure 3.1.11 The ordered pair (v_1, v_2) can represent a point or a vector.

Note, **n** above represents components of the normal vector and not coordinates.

Remark It may have occurred to you that an ordered pair (v_1, v_2) can represent either a vector with *components* v_1 and v_2 or a point with *coordinates* v_1 and v_2 (and similarly for ordered triples). Both are valid geometric interpretations, so the appropriate choice will depend on the geometric viewpoint that we want to emphasize (Figure 3.1.11).



▶ Figure 3.3.1

▶ **EXAMPLE 2 Point-Normal Equations**

It follows from (2) that in R^2 the equation

$$6(x - 3) + (y + 7) = 0$$

represents the line through the point $(3, -7)$ with normal $\mathbf{n} = (6, 1)$; and it follows from (3) that in R^3 the equation

$$4(x - 3) + 2y - 5(z - 7) = 0$$

represents the plane through the point $(3, 0, 7)$ with normal $\mathbf{n} = (4, 2, -5)$. ◀

Lines and Planes Determined by Points and Normals

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0 \tag{1}$$

Thus, Equation (1) can be written as

$$a(x - x_0) + b(y - y_0) = 0 \quad \textbf{[line]} \tag{2}$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \textbf{[plane]} \tag{3}$$

These are called the ***point-normal*** equations of the line and plane.

THEOREM 3.3.1

(a) *If a and b are constants that are not both zero, then an equation of the form*

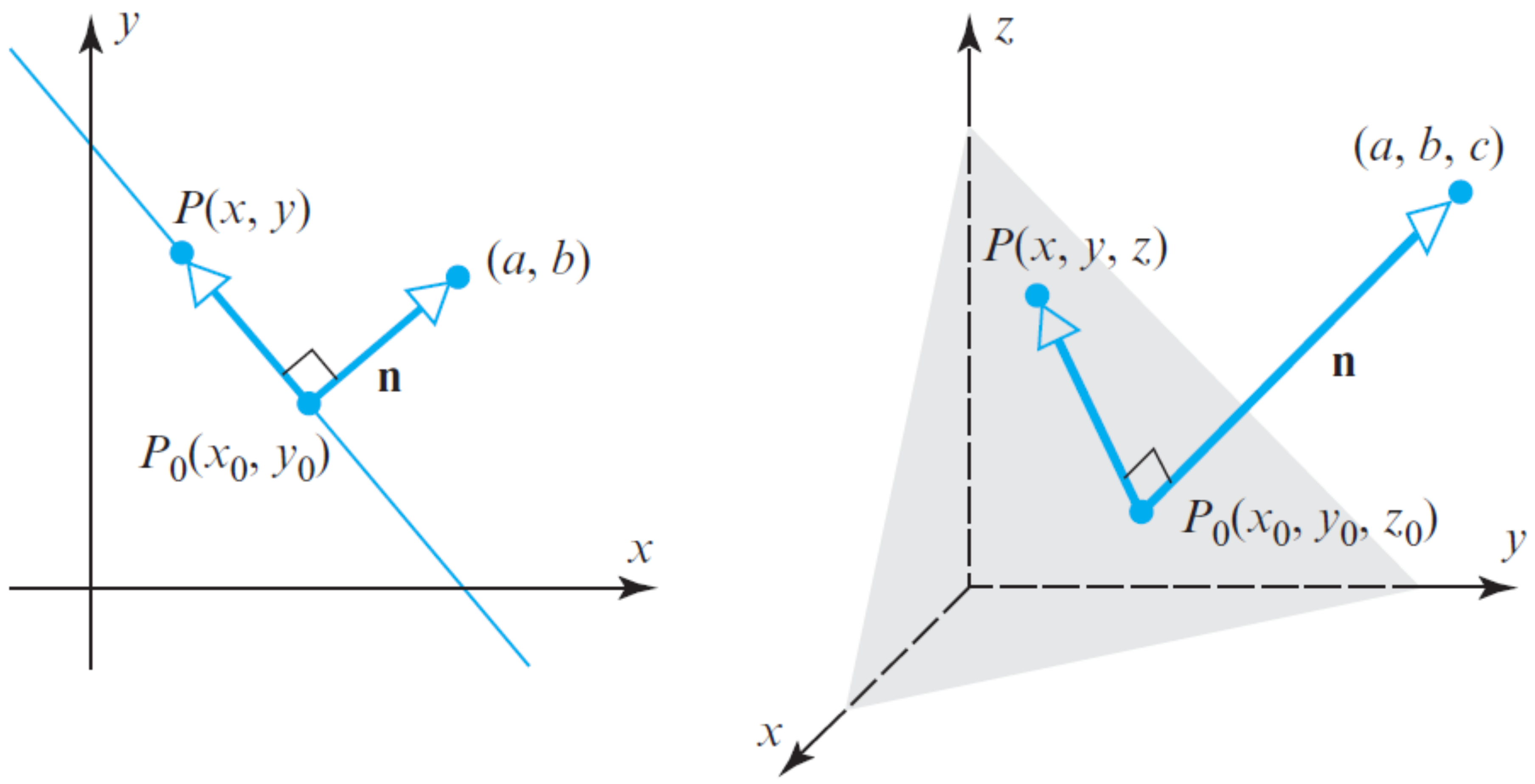
$$ax + by + c = 0 \tag{4}$$

represents a line in R^2 with normal $\mathbf{n} = (a, b)$.

(b) *If a, b, and c are constants that are not all zero, then an equation of the form*

$$ax + by + cz + d = 0 \tag{5}$$

represents a plane in R^3 with normal $\mathbf{n} = (a, b, c)$.



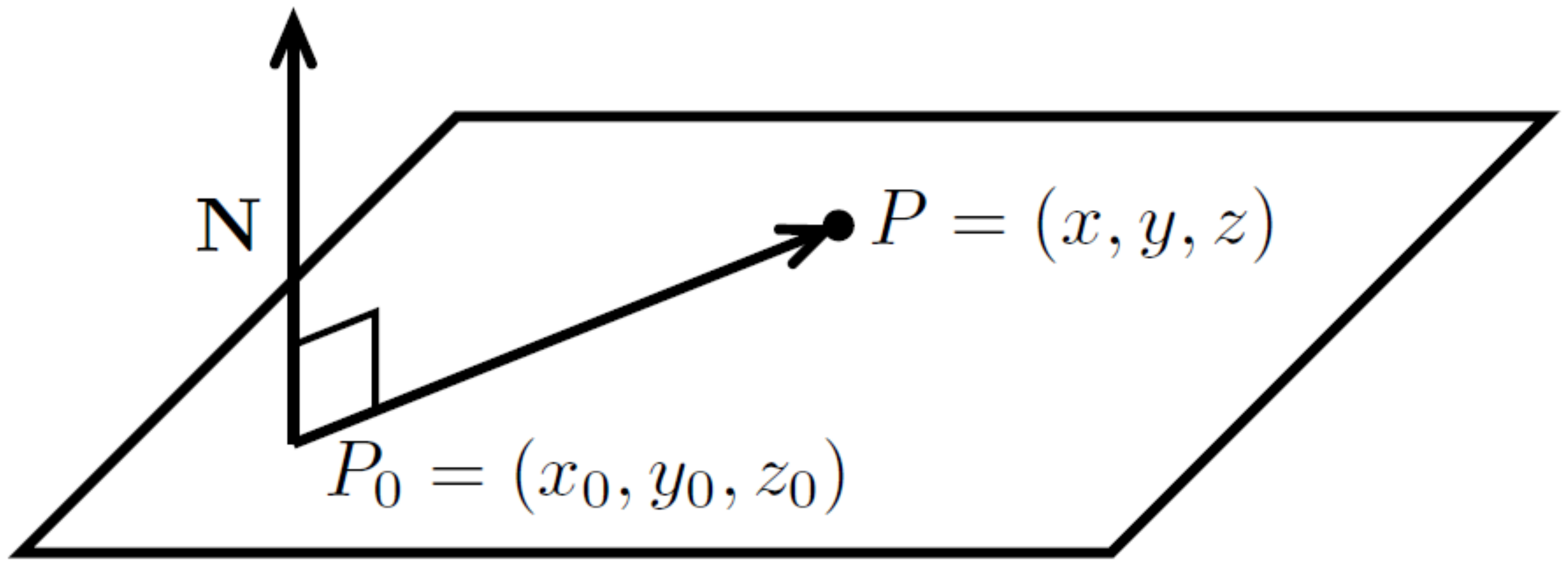
▶ Figure 3.3.1

Lines and Planes Determined by Points and Normals

Let $P = (x, y, z)$ be an arbitrary point in the plane. Then the vector $\overrightarrow{\mathbf{P_0P}}$ is in the plane and therefore orthogonal to $\mathbf{N}$. This means

$$
\begin{aligned}
& \mathbf{N} \cdot \overrightarrow{\mathbf{P_0P}} = 0 \\
\Leftrightarrow \quad & \langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0 \\
\Leftrightarrow \quad & a(x - x_0) + b(y - y_0) + c(z - z_0) = 0
\end{aligned}
$$

We call this last equation the point-normal form for the plane.



Example 1: Find the plane through the point (1,4,9) with normal $\langle 2, 3, 4\rangle$.

Answer: Point-normal form of the plane is $2(x-1)+3(y-4)+4(z-9) = 0$. We can also write this as $2x + 3y + 4z = 50$.

Orthogonal Complements

[redacted] If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (and read as "W perpendicular" or simply "W perp").

1. A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^\perp$ is a subspace of $\mathbb{R}^n$.

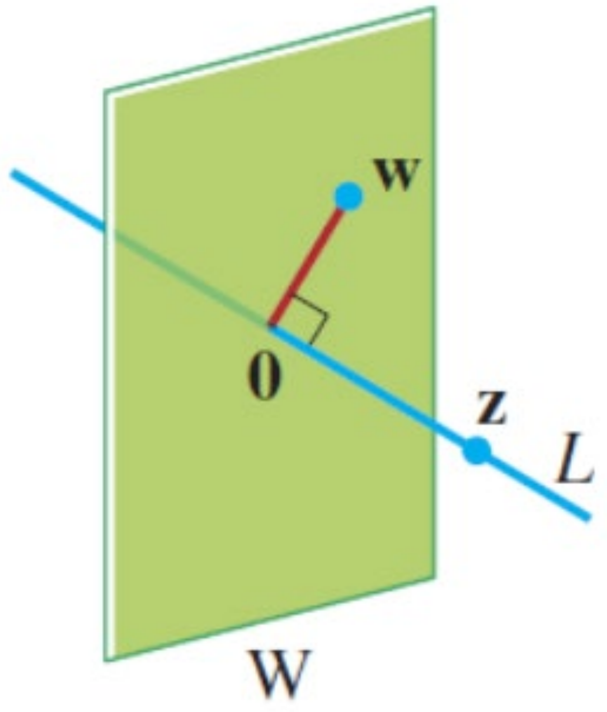


FIGURE 7

A plane and line through $\mathbf{0}$ as orthogonal complements.

EXAMPLE 6 Let W be a plane through the origin in $\mathbb{R}^3$, and let L be the line through the origin and perpendicular to W. If $\mathbf{z}$ and $\mathbf{w}$ are nonzero, $\mathbf{z}$ is on L, and $\mathbf{w}$ is in W, then the line segment from $\mathbf{0}$ to $\mathbf{z}$ is perpendicular to the line segment from $\mathbf{0}$ to $\mathbf{w}$; that is, $\mathbf{z} \cdot \mathbf{w} = 0$. See Figure 7. So each vector on L is orthogonal to every $\mathbf{w}$ in W. In fact, L consists of *all* vectors that are orthogonal to the $\mathbf{w}$'s in W, and W consists of all vectors orthogonal to the $\mathbf{z}$'s in L. That is,

$$L = W^\perp \quad \text{and} \quad W = L^\perp$$

Ref: Lay 5e, pg 336

Orthogonal Complements

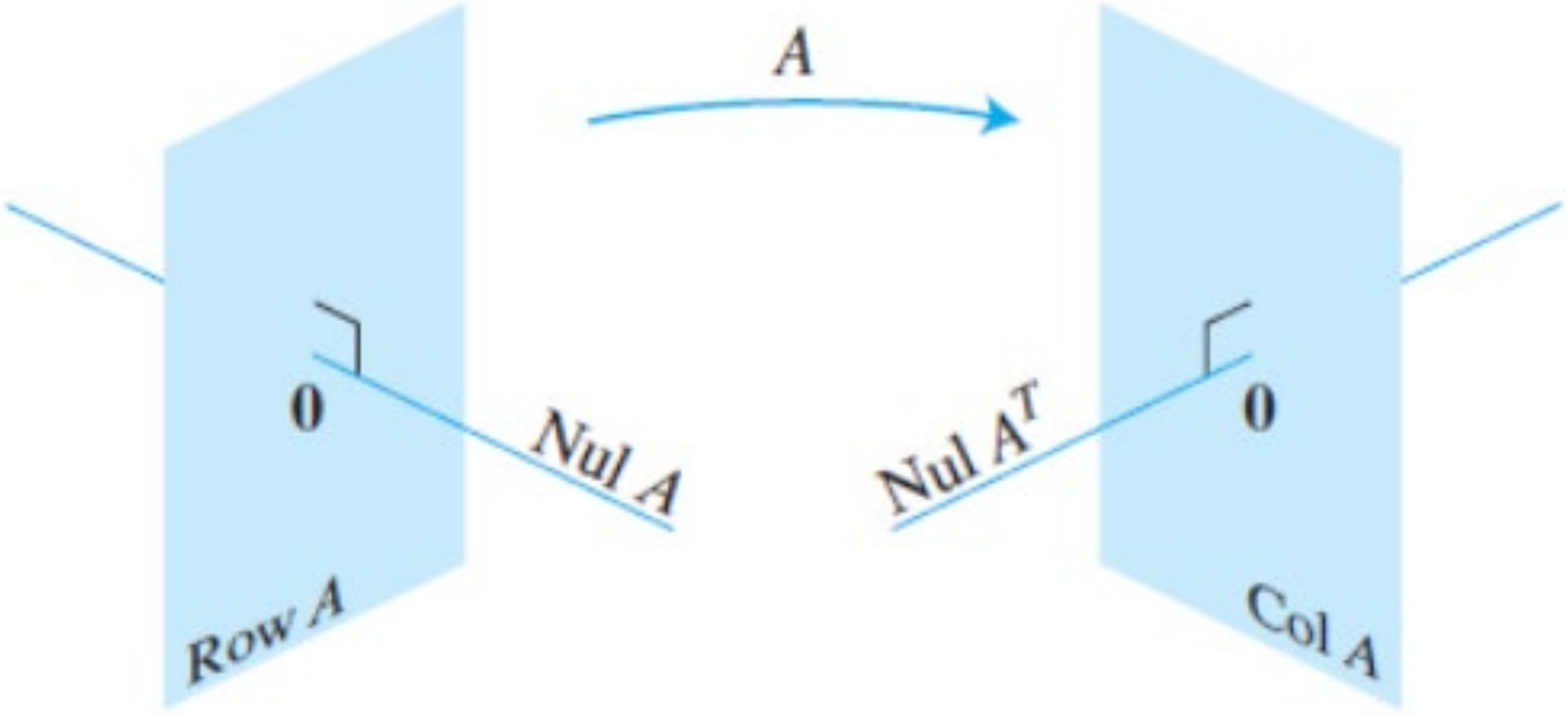


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

THEOREM 3

Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:

$$(\text{Row } A)^\perp = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^\perp = \text{Nul } A^T$$

PROOF The row–column rule for computing $A\mathbf{x}$ shows that if $\mathbf{x}$ is in Nul A, then $\mathbf{x}$ is orthogonal to each row of A (with the rows treated as vectors in $\mathbb{R}^n$). Since the rows of A span the row space, $\mathbf{x}$ is orthogonal to Row A. Conversely, if $\mathbf{x}$ is orthogonal to Row A, then $\mathbf{x}$ is certainly orthogonal to each row of A, and hence $A\mathbf{x} = \mathbf{0}$. This proves the first statement of the theorem. Since this statement is true for any matrix, it is true for A^T. That is, the orthogonal complement of the row space of A^T is the null space of A^T. This proves the second statement, because Row A^T = Col A. ■

Remark: A common way to prove that two sets, say S and T, are equal is to show that S is a subset of T and T is a subset of S. The proof of the next theorem that Nul $A = (\text{Row } A)^\perp$ is established by showing that Nul A is a subset of $(\text{Row } A)^\perp$ and $(\text{Row } A)^\perp$ is a subset of Nul A. That is, an arbitrary element $\mathbf{x}$ in Nul A is shown to be in $(\text{Row } A)^\perp$, and then an arbitrary element $\mathbf{x}$ in $(\text{Row } A)^\perp$ is shown to be in Nul A.

Ref: Lay 5e, pg 337

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

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.2 (mod by Tay Kian Boon)**
Lecture : **Orthogonality**
Topic : **Orthogonality**
Concept : **Orthogonal Projections**

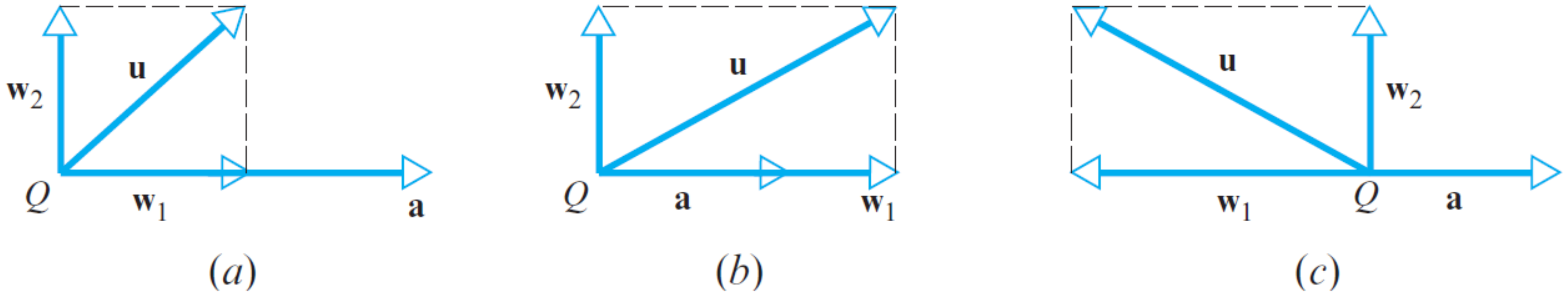
Instructor: **A/P Chng Eng Siong**
TAs: **Zhang Su, Vishal Choudhari**

Rev: 26th June 2020

Orthogonal Projections

Decomposition in R^2 *space*

In many applications it is necessary to "decompose" a vector $\mathbf{u}$ into a sum of two terms, one term being a scalar multiple of a specified nonzero vector $\mathbf{a}$ and the other term being orthogonal to $\mathbf{a}$. For example, if $\mathbf{u}$ and $\mathbf{a}$ are vectors in R^2 that are positioned so their initial points coincide at a point Q, then we can create such a decomposition as follows (Figure 3.3.2):



▲ Figure 3.3.2 Three possible cases.

- Drop a perpendicular from the tip of $\mathbf{u}$ to the line through $\mathbf{a}$.
- Construct the vector $\mathbf{w}_1$ from Q to the foot of the perpendicular.
- Construct the vector $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1$.

Since

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u}$$

we have decomposed $\mathbf{u}$ into a sum of two orthogonal vectors, the first term being a scalar multiple of $\mathbf{a}$ and the second being orthogonal to $\mathbf{a}$.

Decomposing vectors into orthogonal basis has many advantages!
The standard basis is an orthogonal basis.
Orthogonal basis is explained in 6.2.3.

▶ EXAMPLE 1 **The Standard Basis for R^n**

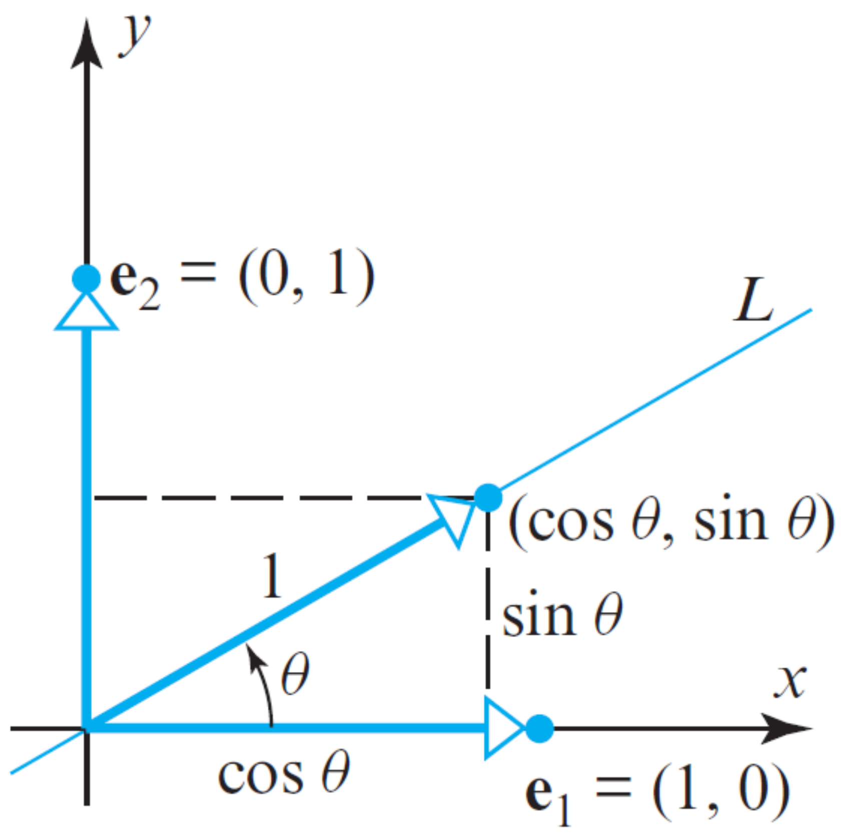
Recall from Example 11 of Section 4.2 that the standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

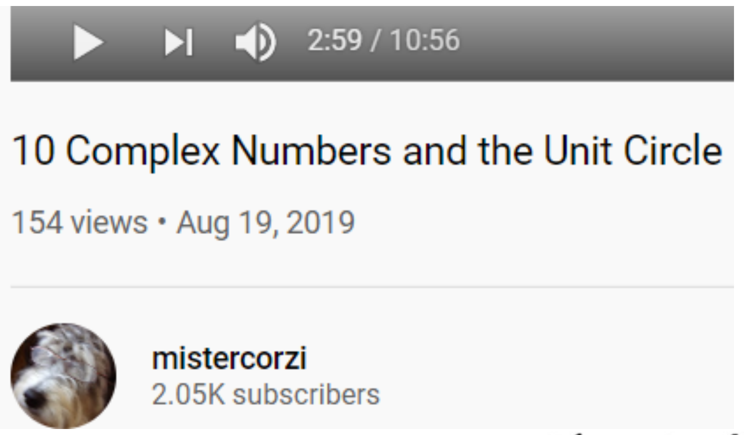
span R^n and from Example 1 of Section 4.3 that they are linearly independent. Thus, they form a basis for R^n that we call the ***standard basis for R^n***. In particular,

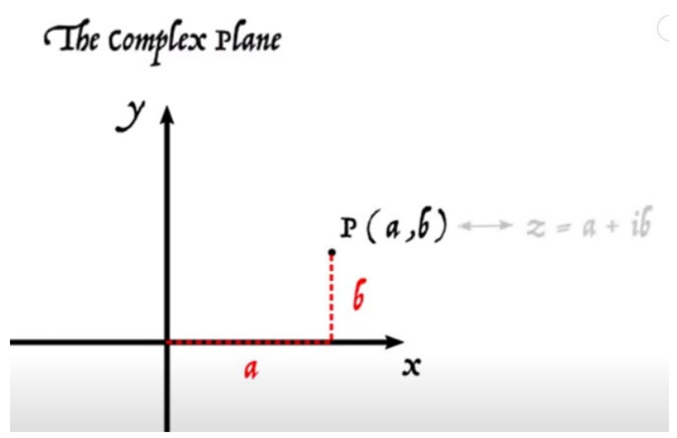
$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

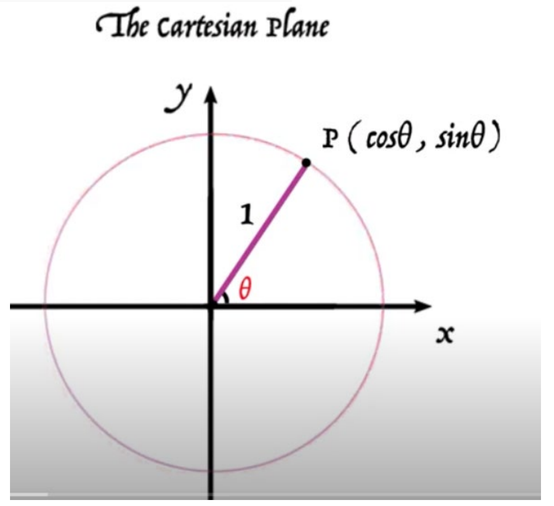
is the standard basis for R^3.



▲ Figure 3.3.3







Mister Corzi: https://www.youtube.com/watch?v=UF172DZOiWo

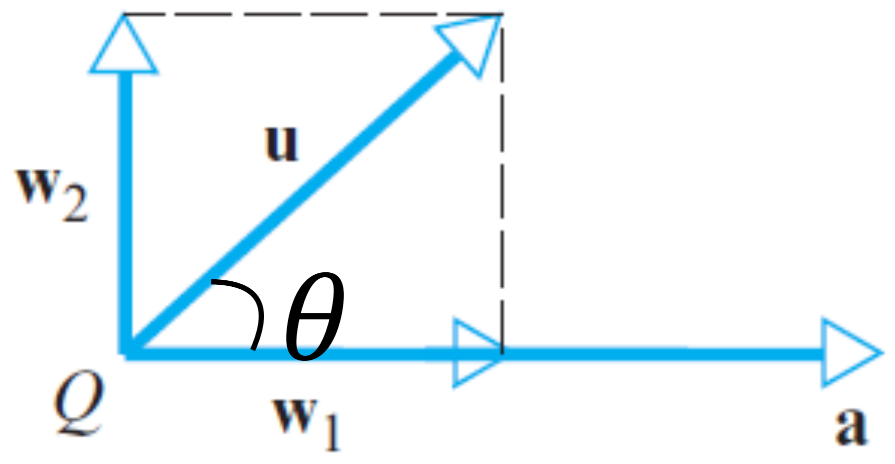
Example above:
Commonly found in complex number of unit circle.
Where e_1 is the real line, and e_2 is the imaginary line

Khan Academy:
https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:complex/x9e81a4f98389efdf:complex-polar/v/polar-form-complex-number
https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:complex/x9e81a4f98389efdf:complex-mul-div-polar/a/complex-number-polar-form-review

Orthogonal Projections

THEOREM 3.3.2 Projection Theorem

If $\mathbf{u}$ *and* $\mathbf{a}$ *are vectors in* R^n, *and if* $\mathbf{a} \neq 0$, *then* $\mathbf{u}$ *can be expressed in exactly one way in the form* $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, *where* $\mathbf{w}_1$ *is a scalar multiple of* $\mathbf{a}$ *and* $\mathbf{w}_2$ *is orthogonal to* $\mathbf{a}$.



(*a*)

The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in the Projection Theorem have associated names—the vector $\mathbf{w}_1$ is called the ***orthogonal projection of u on a*** or sometimes ***the vector component of u along a***, and the vector $\mathbf{w}_2$ is called the ***vector component of u orthogonal to a***. The vector $\mathbf{w}_1$ is commonly denoted by the symbol $\text{proj}_{\mathbf{a}}\mathbf{u}$, in which case it follows from (8) that $\mathbf{w}_2 = \mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u}$. In summary,

$$w_1 = \text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad (\textit{vector component of } \mathbf{u} \textit{ along } \mathbf{a}) \quad = ||u||cos\theta \tag{10}$$

$$w_2 = \mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad (\textit{vector component of } \mathbf{u} \textit{ orthogonal to } \mathbf{a}) \tag{11}$$

EXAMPLE 5 Vector Component of u Along a

Let $\mathbf{u} = (2, -1, 3)$ and $\mathbf{a} = (4, -1, 2)$. Find the vector component of $\mathbf{u}$ along $\mathbf{a}$ and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.

Solution

$$\mathbf{u} \cdot \mathbf{a} = (2)(4) + (-1)(-1) + (3)(2) = 15$$
$$\|\mathbf{a}\|^2 = 4^2 + (-1)^2 + 2^2 = 21$$

Thus the vector component of $\mathbf{u}$ along $\mathbf{a}$ is

$$\text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \tfrac{15}{21}(4, -1, 2) = \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right)$$

and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$ is

$$\mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = (2, -1, 3) - \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right) = \left(-\tfrac{6}{7}, -\tfrac{2}{7}, \tfrac{11}{7}\right)$$

As a check, you may wish to verify that the vectors $\mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u}$ and $\mathbf{a}$ are perpendicular by showing that their dot product is zero.

Linear Algebra: Projection onto a Subspace

12,031 views • Jun 28, 2014

Worldwide Center of Mathematics
26.5K subscribers

Watch these:
Projection onto Line: https://www.youtube.com/watch?v=GnvYEbaSBoY
Projection onto Subspace: https://www.youtube.com/watch?v=zZW6JV4yA54

Orthogonal Projections

Fig. 2.5 shows two vectors $\mathbf{u}, \mathbf{v}$ in R^2 space. The orthogonal projection of $\mathbf{u}$ on $\mathbf{v}$, i.e. $Proj_{\mathbf{v}}\mathbf{u}$, can be thought of as the shadow formed by vector $\mathbf{u}$ onto $\mathbf{v}$ when an imaginary light is directed to $\mathbf{u}$ along the normal of $\mathbf{v}$; The vector $Proj_{\mathbf{v}}\mathbf{u}$ is the best approximation of $\mathbf{u}$ using $\mathbf{v}$. The following are the definitions of orthogonal projection and the equation to find the orthogonal residual.

$$Proj_{\mathbf{v}}\mathbf{u} = \mathbf{v}\left(\frac{\mathbf{u}^T\mathbf{v}}{||\mathbf{v}||^2}\right), \tag{2.34}$$

$$\mathbf{w}_{\mathbf{v}} = \mathbf{u} - Proj_{\mathbf{v}}\mathbf{u}, \tag{2.35}$$

where $||\mathbf{v}|| = \sqrt{\mathbf{v}^T\mathbf{v}}$ is the Euclidean norm of $\mathbf{v}$ and the vector $\mathbf{w}_{\mathbf{v}}$ is the residual when $\mathbf{v}$ is used to approximate $\mathbf{u}$.

Why is the denominator $||v||^2$? (eqn 2.34).

Ans: it is to normalize vector v to have unit length.

Notice that v occurs twice, to get a vector to have unit length,

We need to divide v by $||v||$

$$\left(\frac{v}{||v||}\right) * u^T\left(\frac{v}{||v||}\right) = v * \frac{u^T v}{(||v||^2)}$$

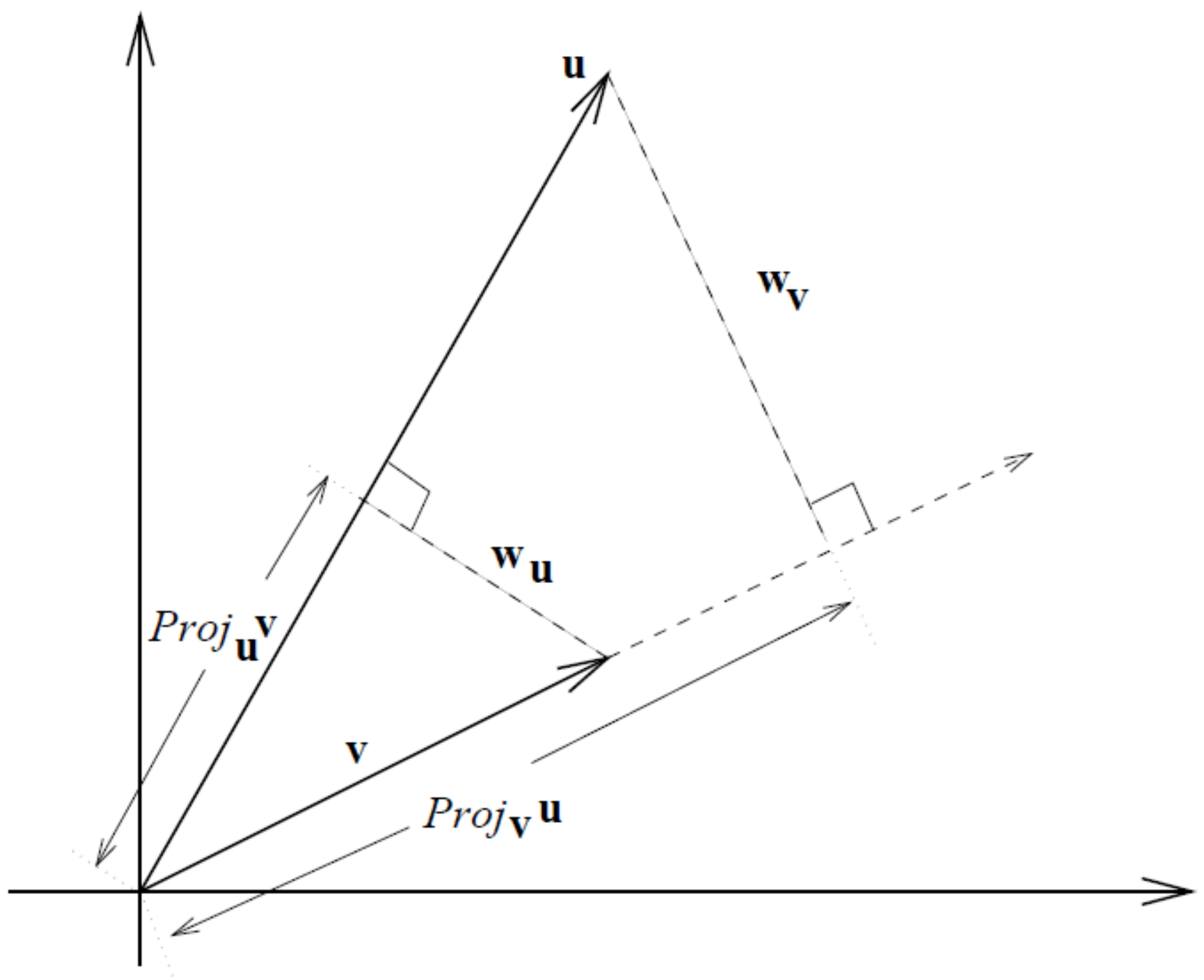
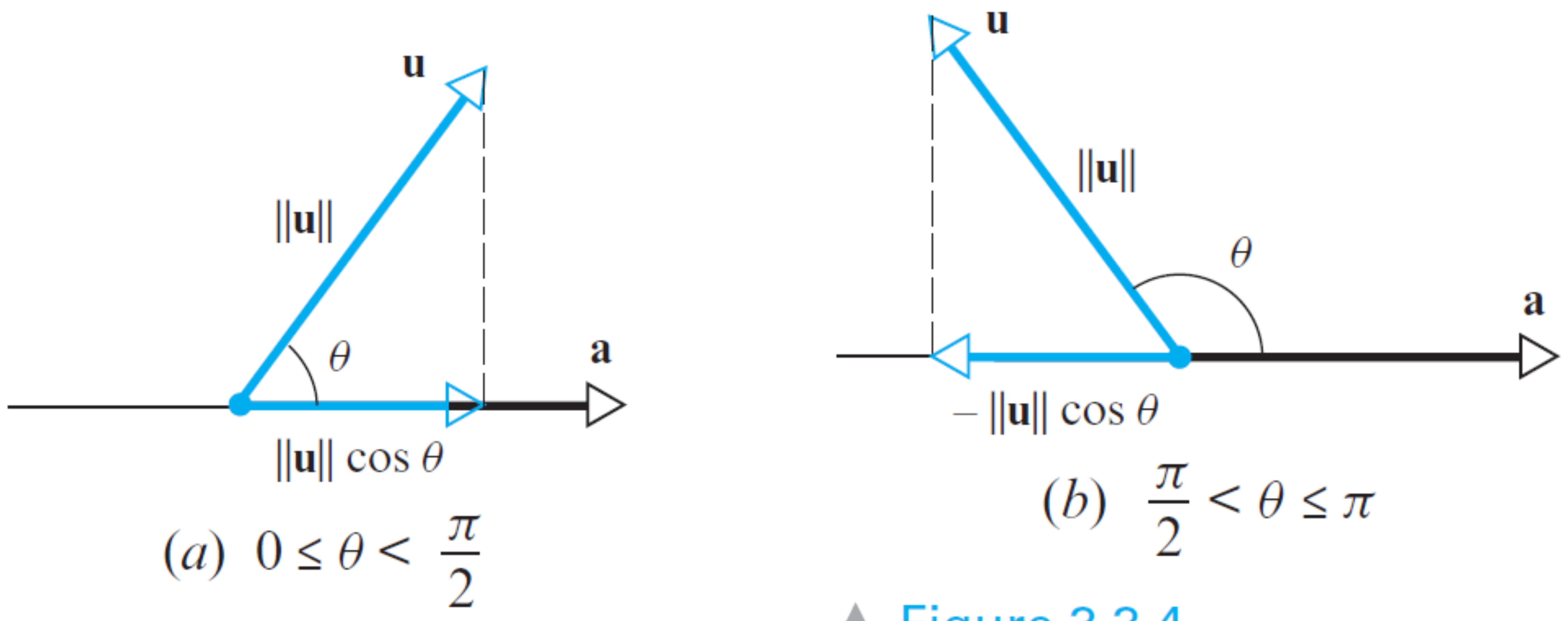


Figure 2.5: Orthogonal projections

Orthogonal Projections

Finding the norm of the projected component.



▲ Figure 3.3.4

Sometimes we will be more interested in the *norm* of the vector component of $\mathbf{u}$ along $\mathbf{a}$ than in the vector component itself. A formula for this norm can be derived as follows:

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \left\|\frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}\right\| = \left|\frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\right|\|\mathbf{a}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|^2}\|\mathbf{a}\|$$

where the second equality follows from part (c) of Theorem 3.2.1 and the third from the fact that $\|\mathbf{a}\|^2 > 0$. Thus,

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|} \tag{12}$$

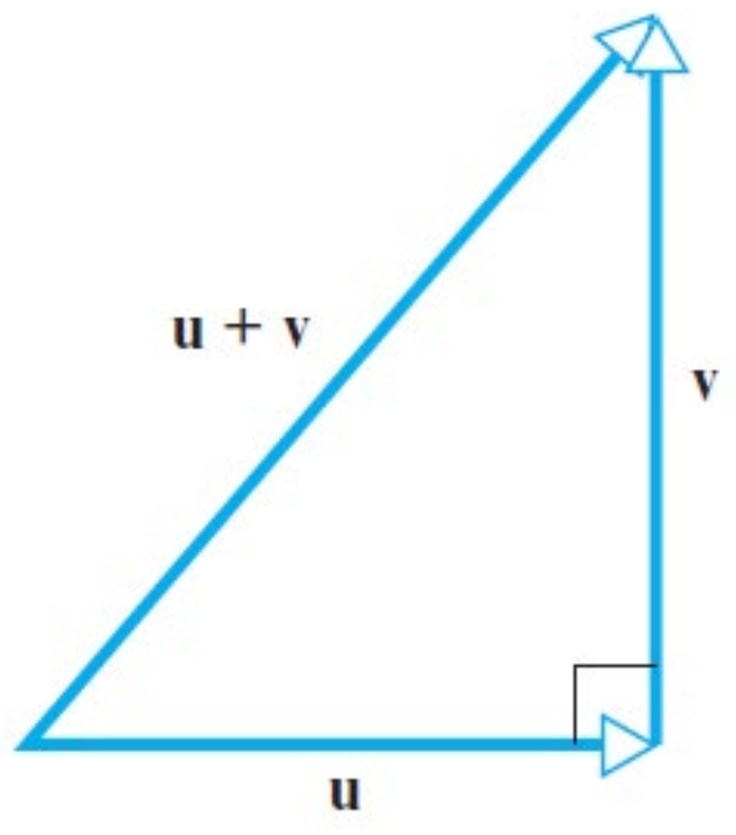
If θ denotes the angle between $\mathbf{u}$ and $\mathbf{a}$, then $\mathbf{u}\cdot\mathbf{a} = \|\mathbf{u}\|\|\mathbf{a}\|\cos\theta$, so (12) can also be written as

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \|\mathbf{u}\||\cos\theta| \tag{13}$$

(Verify.) A geometric interpretation of this result is given in Figure 3.3.4.

Pythagoras Theorem

The Theorem of Pythagoras



▲ Figure 3.3.5

THEOREM 3.3.3 Theorem of Pythagoras in R^n

If **u** *and* **v** *are orthogonal vectors in* R^n *with the Euclidean inner product, then*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \tag{14}$$

Proof Since **u** and **v** are orthogonal, we have $\mathbf{u}\cdot\mathbf{v} = 0$, from which it follows that

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad ◀$$

▶ EXAMPLE 6 **Theorem of Pythagoras in R^4**

We showed in Example 1 that the vectors

$$\mathbf{u} = (-2, 3, 1, 4) \quad \text{and} \quad \mathbf{v} = (1, 2, 0, -1)$$

are orthogonal. Verify the Theorem of Pythagoras for these vectors.

Solution We leave it for you to confirm that

$$\begin{aligned} \mathbf{u}+\mathbf{v} &= (-1, 5, 1, 3) \\ \|\mathbf{u}+\mathbf{v}\|^2 &= 36 \\ \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 &= 30 + 6 \end{aligned}$$

Thus, $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ ◀

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

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.3 (mod By Tay KB)**

Lecture : **Orthogonality**

Topic : **Orthogonality**

Concept : **Orthogonal Sets & Orthogonal Basis**

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

TAs: **Zhang Su, Vishal Choudhari**

Rev: 26th June 2020

Orthogonal Sets

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$
$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0$$
$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0$$

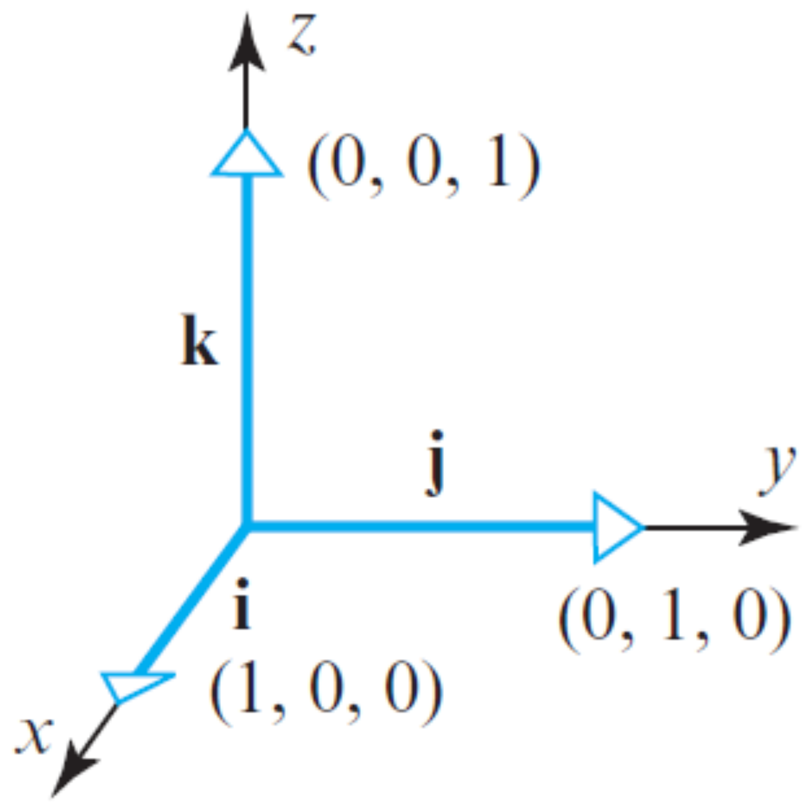
Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Fig. 1; the three line segments there are mutually perpendicular. ■

Lay's Linear Algebra and Applications

Standard basis for R^n is an orthogonal set

Standard Basis

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



(b)

▲ Figure 3.2.2

Ref: https://mathworld.wolfram.com/StandardBasis.html

Orthogonal Sets and Orthogonal Basis

THEOREM 4 $\quad Note{:}\, p \leq n$

> If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \ldots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \ldots, c_p$ must be zero. Thus S is linearly independent. ■

Lay's Linear Algebra and Applications

338 CHAPTER 6 Orthogonality and Least Squares

Examples of orthogonal set of vectors in R^3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & -1 & -1/2 \\ 1 & 2 & -2 \\ 1 & 1 & 7/2 \end{bmatrix}$$

```
y = -8/6;
S = [ 1 0 ;
      2 4 ;
      6 y ];
S
checkOrthognality = S'*S
```

$\mathbf{u}_1 \quad \mathbf{u}_2$

```
S =

    1.0000         0
    2.0000    4.0000
    6.0000   -1.3333

checkOrthognality =

   41.0000    0.0000
    0.0000   17.7778
```

$$S = \begin{bmatrix} \uparrow & \uparrow \\ u_1 & u_2 \\ \downarrow & \downarrow \end{bmatrix}_{3\times 2} \quad S^T = \begin{bmatrix} \leftarrow & u_1^T & \rightarrow \\ \leftarrow & u_2^T & \rightarrow \end{bmatrix}_{2\times 3} \quad S^T \times S = \begin{bmatrix} ||u_1||^2 & u_1^T u_2 \\ u_2^T u_1 & ||u_2||^2 \end{bmatrix}_{2\times 2}$$

Dot product between vectors u_1 and u_2

Orthogonal Sets and Orthogonal Basis

THEOREM 4 $Note: p \leq n$

If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \dots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \dots, c_p$ must be zero. Thus S is linearly independent. ■

Lay's Linear Algebra and Applications

An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

THEOREM 5

The next theorem suggests why an orthogonal basis is much nicer than other bases. The weights in a linear combination can be computed easily.

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \dots, p)$$

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ shows that

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Since $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero, the equation above can be solved for c_1. To find c_j for $j = 2, \dots, p$, compute $\mathbf{y}\cdot\mathbf{u}_j$ and solve for c_j. ■

Projecting a vector onto a subspace (span by an orthogonal basis)

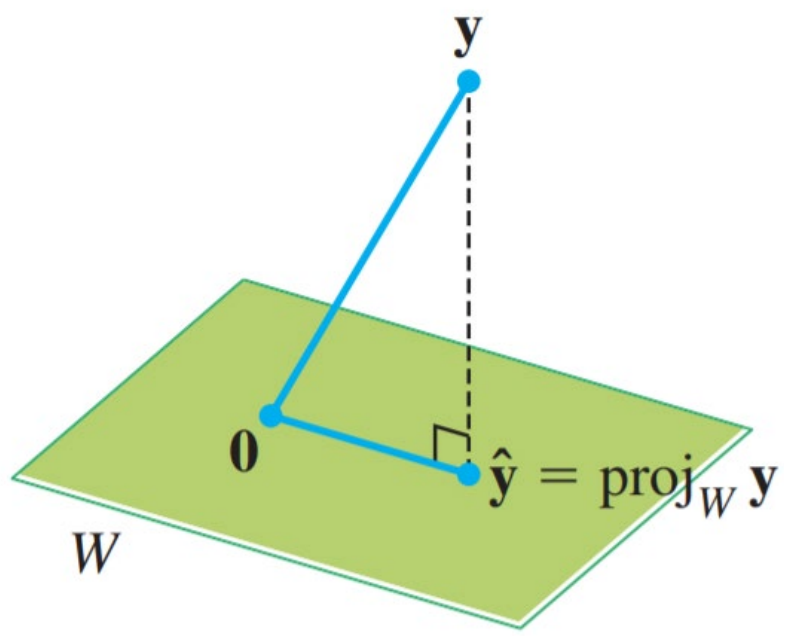


FIGURE 1

Given a vector $\mathbf{y}$ and a subspace W in $\mathbb{R}^n$, there is a vector $\hat{\mathbf{y}}$ in W such that (1) $\hat{\mathbf{y}}$ is the unique vector in W for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W, and (2) $\hat{\mathbf{y}}$ is the unique vector in W closest to $\mathbf{y}$. See Figure 1. These two properties of $\hat{\mathbf{y}}$ provide the key to finding least-squares solutions of linear systems, mentioned in the introductory example for this chapter.

To prepare for the first theorem, observe that whenever a vector $\mathbf{y}$ is written as a linear combination of vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ in $\mathbb{R}^n$, the terms in the sum for $\mathbf{y}$ can be grouped into two parts so that $\mathbf{y}$ can be written as

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$

where $\mathbf{z}_1$ is a linear combination of some of the $\mathbf{u}_i$ and $\mathbf{z}_2$ is a linear combination of the rest of the $\mathbf{u}_i$. This idea is particularly useful when $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthogonal basis.

Lay 5e, pg 349

EXAMPLE 1 Let $\{\mathbf{u}_1, \dots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$ and let

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Consider the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, and write $\mathbf{y}$ as the sum of a vector $\mathbf{z}_1$ in W and a vector $\mathbf{z}_2$ in $W^\perp$.

SOLUTION Write

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

where $\quad \mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \quad$ is in $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$

and $\quad \mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5 \quad$ is in $\text{Span}\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

To show that $\mathbf{z}_2$ is in $W^\perp$, it suffices to show that $\mathbf{z}_2$ is orthogonal to the vectors in the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. (See Section 6.1.) Using properties of the inner product, compute

$$\begin{aligned} \mathbf{z}_2 \cdot \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5) \cdot \mathbf{u}_1 \\ &= c_3\mathbf{u}_3 \cdot \mathbf{u}_1 + c_4\mathbf{u}_4 \cdot \mathbf{u}_1 + c_5\mathbf{u}_5 \cdot \mathbf{u}_1 \\ &= 0 \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_3$, $\mathbf{u}_4$, and $\mathbf{u}_5$. A similar calculation shows that $\mathbf{z}_2 \cdot \mathbf{u}_2 = 0$. Thus $\mathbf{z}_2$ is in $W^\perp$. ■

Orthogonal Sets and Orthogonal Basis

The next theorem suggests why an orthogonal basis is much nicer than other bases. The weights in a linear combination can be computed easily.

THEOREM 5

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \ldots, p)$$

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ shows that

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Since $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero, the equation above can be solved for c_1. To find c_j for $j = 2, \ldots, p$, compute $\mathbf{y}\cdot\mathbf{u}_j$ and solve for c_j. ■

Lay's Linear Algebra and Applications

THEOREM 8

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 $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ 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 $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

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 Figure 2. When W is a one-dimensional subspace, the formula for $\hat{\mathbf{y}}$ matches the formula given in Section 6.2.

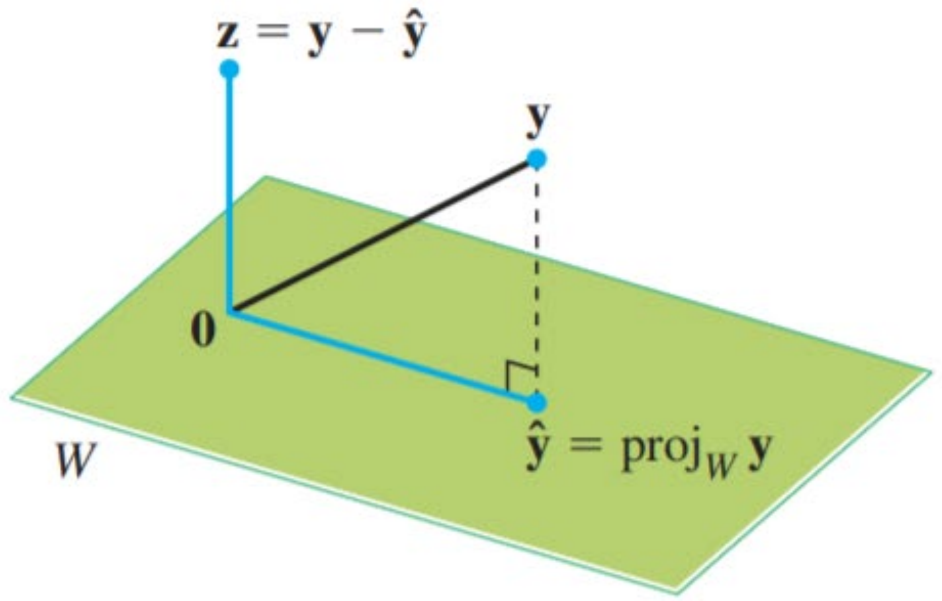


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

Proof later: ch 6.2.5

Example

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Decompose $y = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ using the standard basis.

$$\begin{bmatrix} 6 \\ 1 \\ 8 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + (-8)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Lay’s Linear Algebra and Applications

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

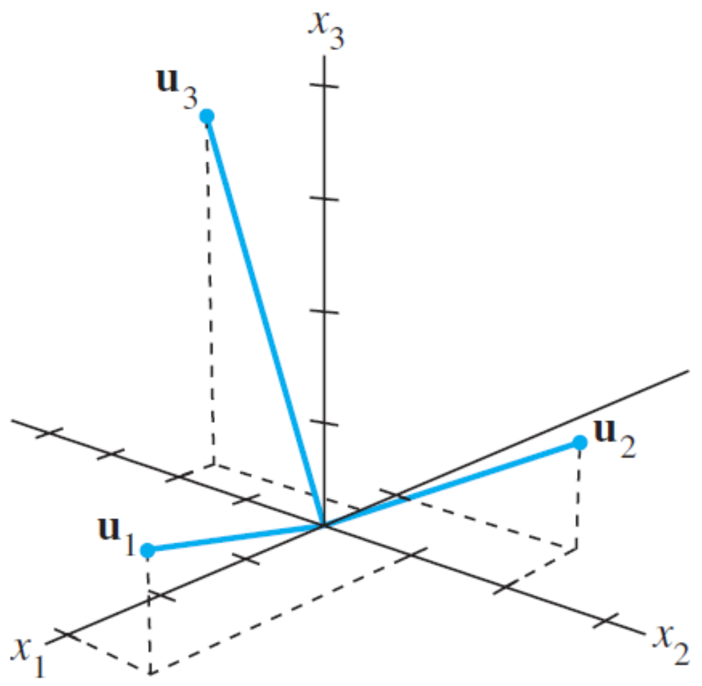


FIGURE 1

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$
$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0$$
$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Figure 1; the three line segments there are mutually perpendicular. ■

EXAMPLE 2 The set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in Example 1 is an orthogonal basis for $\mathbb{R}^3$. Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.

SOLUTION Compute

$$\mathbf{y} \cdot \mathbf{u}_1 = 11, \qquad \mathbf{y} \cdot \mathbf{u}_2 = -12, \qquad \mathbf{y} \cdot \mathbf{u}_3 = -33$$
$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 11, \qquad \mathbf{u}_2 \cdot \mathbf{u}_2 = 6, \qquad \mathbf{u}_3 \cdot \mathbf{u}_3 = 33/2$$

By Theorem 5,

$$\begin{aligned} \mathbf{y} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3 \end{aligned}$$ ■

Notice how easy it is to compute the weights needed to build $\mathbf{y}$ from an orthogonal basis. If the basis were not orthogonal, it would be necessary to solve a system of linear equations in order to find the weights, as in Chapter 1.

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

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.4**

Lecture : **Orthogonality (adapted TKB)**

Topic : **Orthogonality**

Concept : **Orthonormal Sets & Orthogonal Matrices**

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

TAs: **Zhang Su, Vishal Choudhari**

Rev: 26th June 2020

Orthonormal Sets-vectors within have length 1!

Orthonormal Sets

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$. Any nonempty subset of $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is orthonormal, too. Here is a more complicated example.

THEOREM 4 If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

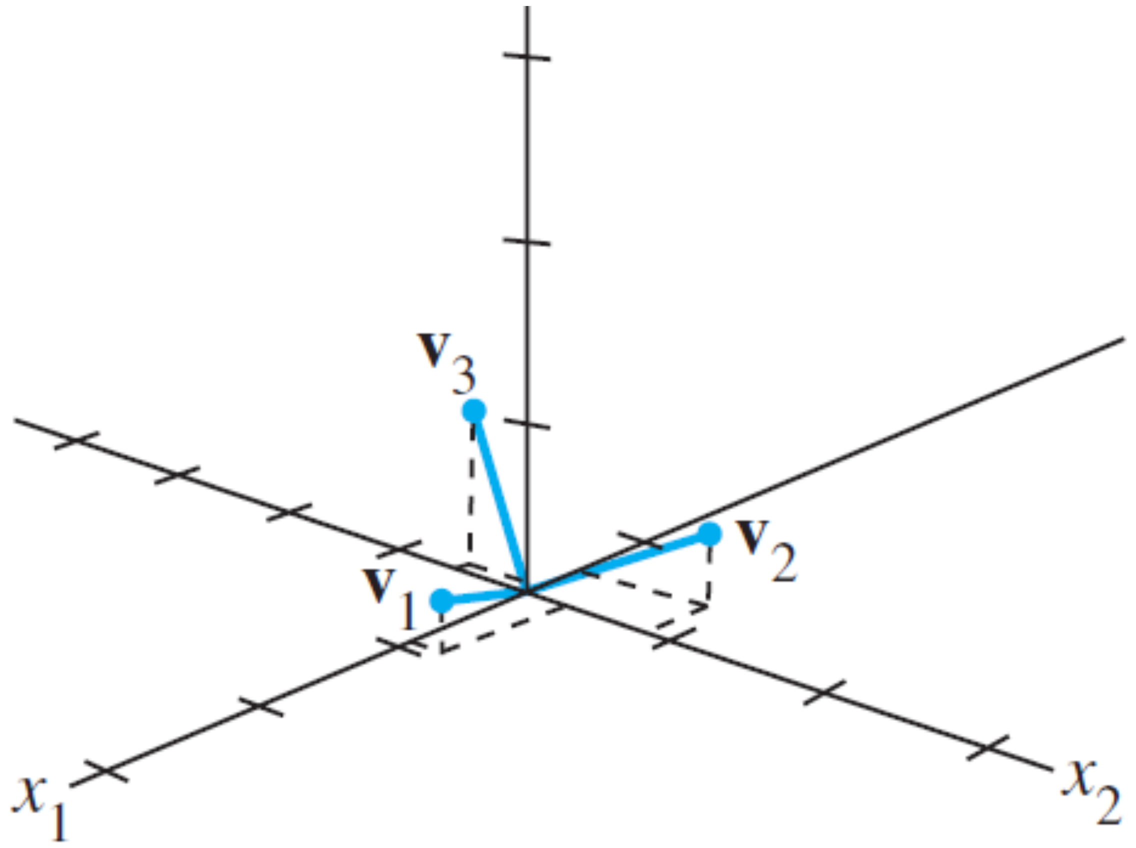


FIGURE 6

Ref: https://www.youtube.com/watch?v=ZJu26chXEiw

Lay's Linear Algebra and Applications

EXAMPLE 5 Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUTION Compute

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0 \end{aligned}$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set. Also,

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_1 &= 9/11 + 1/11 + 1/11 = 1 \\ \mathbf{v}_2 \cdot \mathbf{v}_2 &= 1/6 + 4/6 + 1/6 = 1 \\ \mathbf{v}_3 \cdot \mathbf{v}_3 &= 1/66 + 16/66 + 49/66 = 1 \end{aligned}$$

which shows that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors. Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set. Since the set is linearly independent, its three vectors form a basis for $\mathbb{R}^3$. See Fig. 6. ■

```
S3 = [ 3/sqrt(11)   -1/sqrt(6)   -1/sqrt(66);
       1/sqrt(11)   2/sqrt(6)    -4/sqrt(66);
       1/sqrt(11)   1/sqrt(6)     7/sqrt(66)];

S3
checkOrthognality = S3'*S3
```

```
S3 =

    0.9045   -0.4082   -0.1231
    0.3015    0.8165   -0.4924
    0.3015    0.4082    0.8616


checkOrthognality =

    1.0000    0.0000    0.0000
    0.0000    1.0000    0.0000
    0.0000    0.0000    1.0000
```

NOTE: The 0's correspond to dot products of orthogonal vectors. See next slide for explanation of result!

Orthonormal Sets and Orthogonal Matrices

THEOREM 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

PROOF To simplify notation, we suppose that U has only three columns, each a vector in $\mathbb{R}^m$. The proof of the general case is essentially the same. Let $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

The entries in the matrix at the right are inner products, using transpose notation. The columns of U are orthogonal if and only if

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

The columns of U all have unit length if and only if

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

The theorem follows immediately from (4)–(6). ■

Lay's Linear Algebra and Applications

6.2 Orthogonal Sets 343

Orthogonal matrices.

- A matrix $Q \in \mathbb{R}^{m \times n}$ is called orthogonal if $Q^T Q = I_n$, i.e., if its columns are orthogonal and have 2-norm one.
- If $Q \in \mathbb{R}^{n \times n}$ is orthogonal, then $Q^T Q = I$ implies that $Q^{-1} = Q^T$.
- If $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, then Q^T is an orthogonal matrix.

Ref: https://en.wikipedia.org/wiki/Orthogonal_matrix **Important!**

Matlab Example:

```
S3 = [ 3/sqrt(11)   -1/sqrt(6)   -1/sqrt(66);
       1/sqrt(11)    2/sqrt(6)   -4/sqrt(66);
       1/sqrt(11)    1/sqrt(6)    7/sqrt(66)];

S3
checkOrthognality = S3'*S3
```

```
S3 =

    0.9045   -0.4082   -0.1231
    0.3015    0.8165   -0.4924
    0.3015    0.4082    0.8616


checkOrthognality =

    1.0000    0.0000    0.0000
    0.0000    1.0000    0.0000
    0.0000    0.0000    1.0000
```

Orthonormal Sets-Nice Properties

THEOREM 7

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$

b. $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$

c. $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$

Properties (a) and (c) say that the linear mapping $\mathbf{x} \mapsto U\mathbf{x}$ preserves lengths and orthogonality. These properties are crucial for many computer algorithms.

Lay's Linear Algebra and Applications

6.2 Orthogonal Sets 343

EXAMPLE 6 Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Notice that U has orthonormal columns and

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verify that $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUTION

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

Theorems 6 and 7 are particularly useful when applied to *square* matrices. An **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$. By Theorem 6, such a matrix has orthonormal columns.[1] It is easy to see that any *square* matrix with orthonormal columns is an orthogonal matrix. Surprisingly, such a matrix must have orthonormal *rows*, too.

Important: Definition of Orthogonal Matrix

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

CX1104: **Linear Algebra for Computing**

Chap. No : **6.2.5 (mod Tay KB)**

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

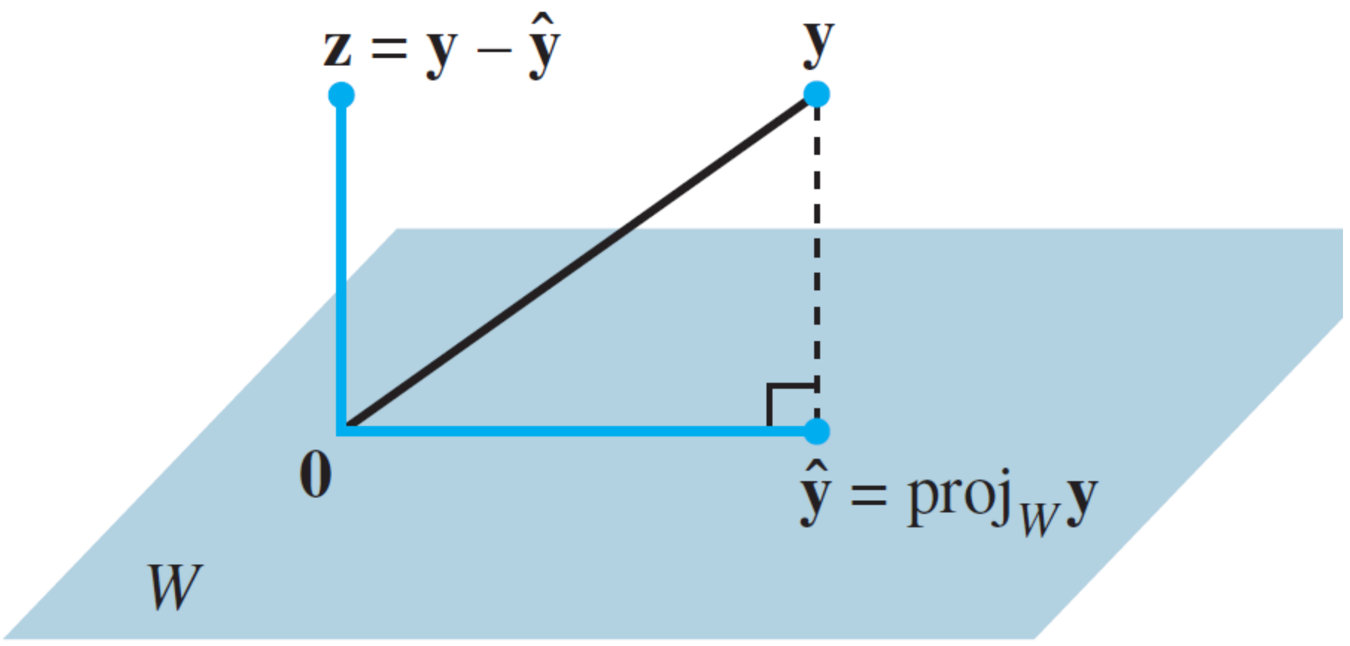


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

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

[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

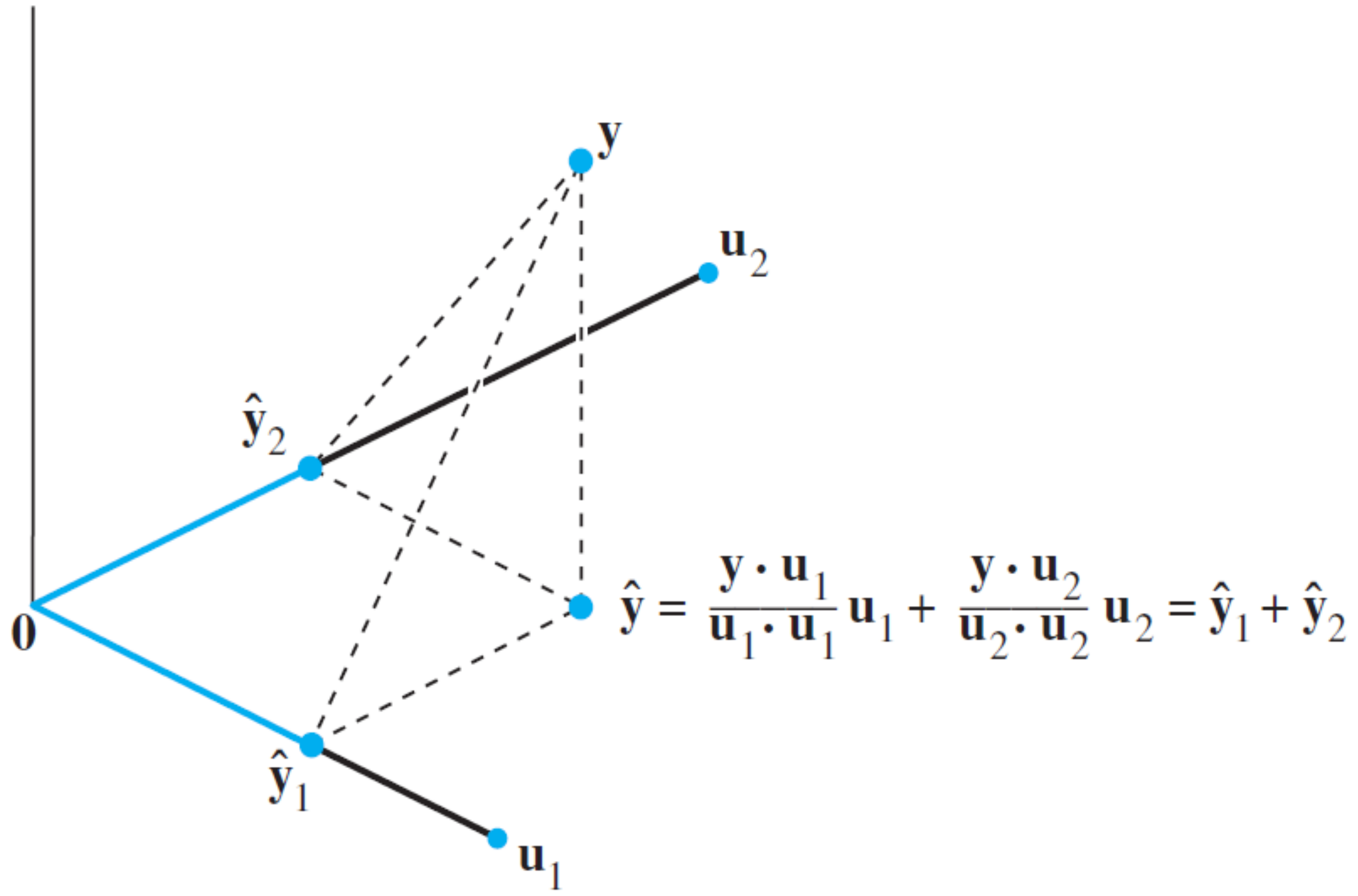


FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

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

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 $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ 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 $W = \text{Span}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, then $\text{proj}_W\, \mathbf{y} = \mathbf{y}$.

Lay's Linear Algebra and Applications

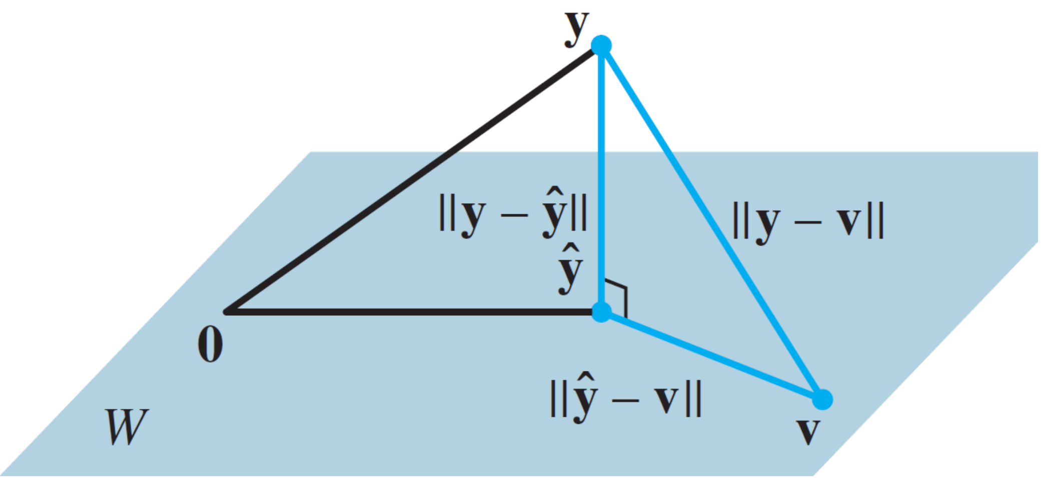


FIGURE 4 The orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$.

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. ■

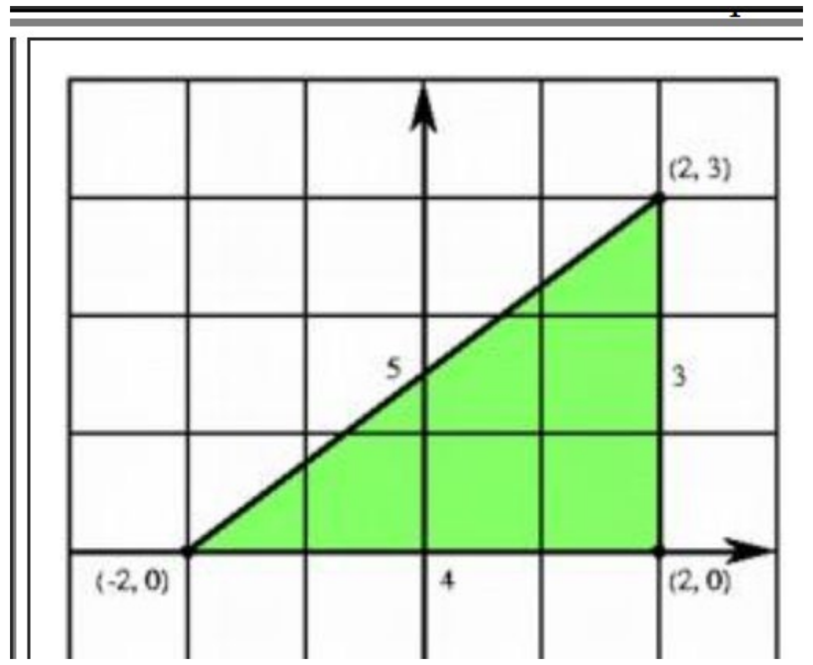
Conclusion: if y is in the space of W, then we can show $\hat{y}$ to be equal to y using theorem 8. Else, it's the best approximation according to theorem 9.

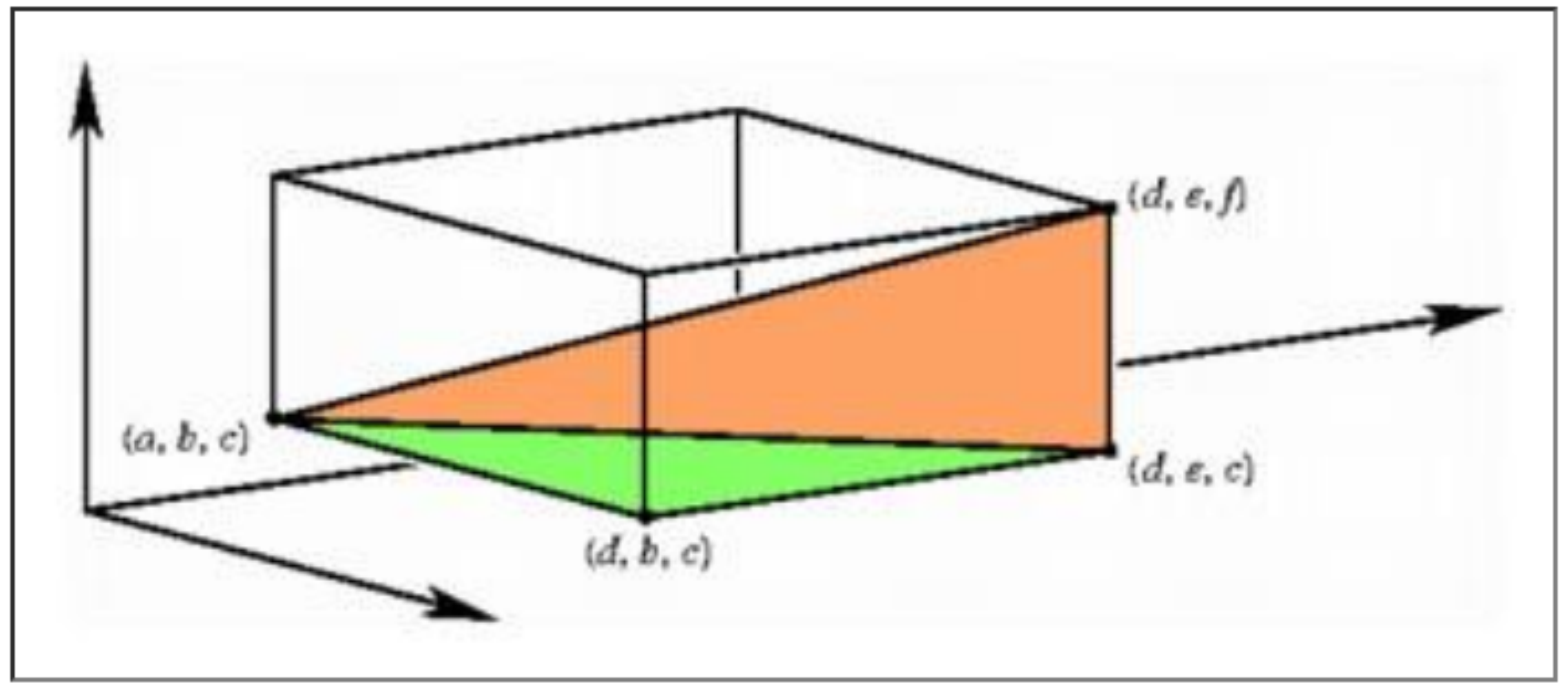
Using Pythagoras Theorem in Higher Dimensions

Ref:

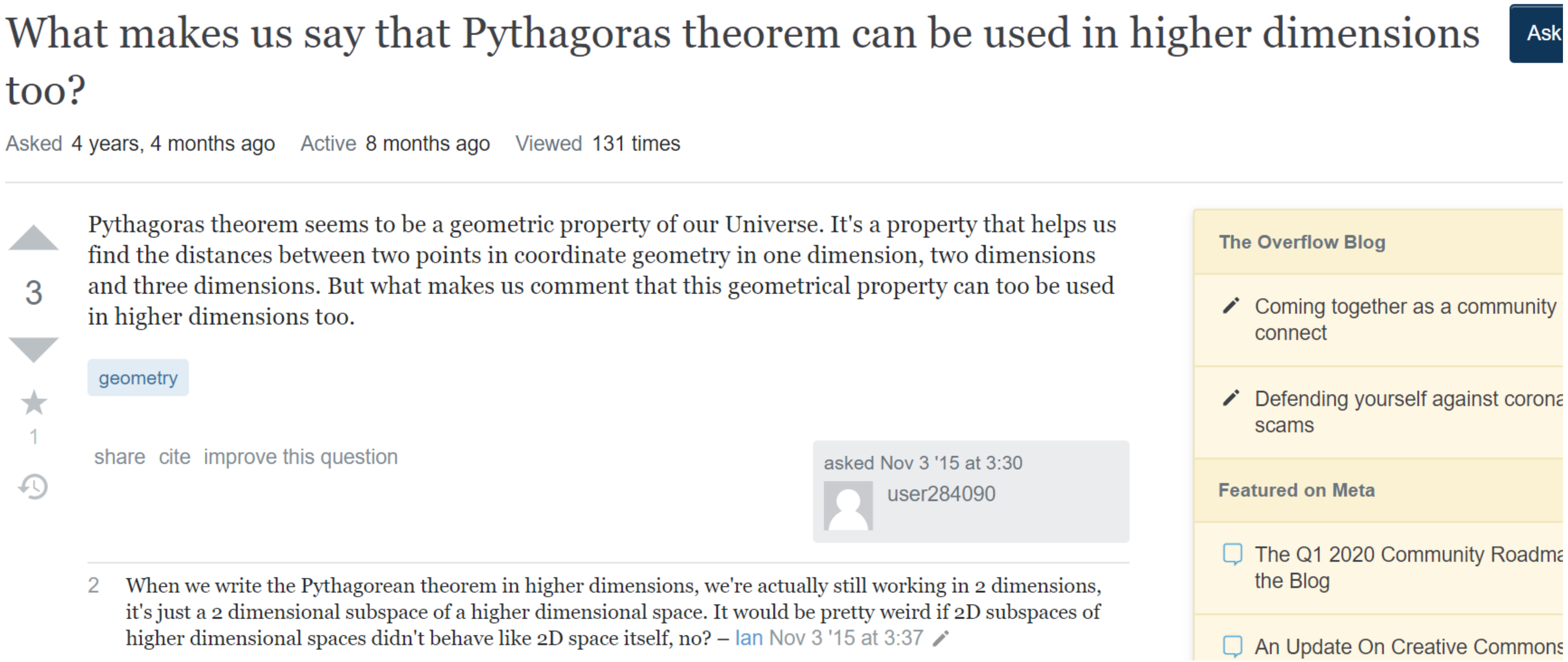
i) https://math.stackexchange.com/questions/1510549/what-makes-us-say-that-pythagoras-theorem-can-be-used-in-higher-dimensions-too

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

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

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 $\sqrt{45} = 3\sqrt{5}$. $\blacksquare$

Best Approximation Theorem

THEOREM 10

If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ 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 $\mathbf{y}\cdot\mathbf{u}_1$, $\mathbf{y}\cdot\mathbf{u}_2, \dots, \mathbf{y}\cdot\mathbf{u}_p$. The weights can be written as $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \dots, \mathbf{u}_p^T\mathbf{y}$, 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. You need to convert them to norm 1 and use Theorem 10 (orthonormality is needed). 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