

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.1.2 (modified by Tay Kian Boon)**

Lecture : **Orthogonality**

Topic : **Dot Product**

Concept : **Norm of a Vector and Unit Vectors**

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

Norm of Vector

DEFINITION 1 If $\mathbf{v} = (v_1, v_2, \dots, v_n)$ is a vector in R^n , then the *norm* of \mathbf{v} (also called the *length* of \mathbf{v} or the *magnitude* of \mathbf{v}) is denoted by $\|\mathbf{v}\|$, and is defined by the formula

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \quad (3)$$

Unit Vector

For any non-zero vector v in Euclidean space, $v/\|v\|$ is a unit vector pointing in same direction as v .

Unit Vector Example

► EXAMPLE 2 Normalizing a Vector

Find the unit vector \mathbf{u} that has the same direction as $\mathbf{v} = (2, 2, -1)$.

Solution The vector \mathbf{v} has length

$$\|\mathbf{v}\| = \sqrt{2^2 + 2^2 + (-1)^2} = 3$$

Thus, from (4)

$$\mathbf{u} = \frac{1}{3}(2, 2, -1) = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$$

As a check, you may want to confirm that $\|\mathbf{u}\| = 1$. ◀

Norm of a Vector: properties

THEOREM 3.2.1 *If \mathbf{v} is a vector in R^n , and if k is any scalar, then:*

- (a) $\|\mathbf{v}\| \geq 0$
- (b) $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$
- (c) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$

We will prove part (c) and leave (a) and (b) as exercises.

Proof (c) If $\mathbf{v} = (v_1, v_2, \dots, v_n)$, then $k\mathbf{v} = (kv_1, kv_2, \dots, kv_n)$, so

$$\begin{aligned}\|k\mathbf{v}\| &= \sqrt{(kv_1)^2 + (kv_2)^2 + \cdots + (kv_n)^2} \\ &= \sqrt{(k^2)(v_1^2 + v_2^2 + \cdots + v_n^2)} \\ &= |k|\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \\ &= |k|\|\mathbf{v}\| \quad \blacktriangleleft\end{aligned}$$

Examples

A vector whose length is 1 is called a **unit vector**. If we *divide* a nonzero vector \mathbf{v} by its length—that is, multiply by $1/\|\mathbf{v}\|$ —we obtain a unit vector \mathbf{u} because the length of \mathbf{u} is $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. The process of creating \mathbf{u} from \mathbf{v} is sometimes called **normalizing** \mathbf{v} , and we say that \mathbf{u} is *in the same direction* as \mathbf{v} .

Several examples that follow use the space-saving notation for (column) vectors.

EXAMPLE 2 Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector \mathbf{u} in the same direction as \mathbf{v} .

SOLUTION First, compute the length of \mathbf{v} :

$$\begin{aligned}\|\mathbf{v}\|^2 &= \mathbf{v} \cdot \mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9 \\ \|\mathbf{v}\| &= \sqrt{9} = 3\end{aligned}$$

Then, multiply \mathbf{v} by $1/\|\mathbf{v}\|$ to obtain

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|} \mathbf{v} = \frac{1}{3} \mathbf{v} = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

To check that $\|\mathbf{u}\| = 1$, it suffices to show that $\|\mathbf{u}\|^2 = 1$.

$$\begin{aligned}\|\mathbf{u}\|^2 &= \mathbf{u} \cdot \mathbf{u} = \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + (0)^2 \\ &= \frac{1}{9} + \frac{4}{9} + \frac{4}{9} + 0 = 1\end{aligned}$$

Note: $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$

Derived in Slide 6 of Lecture 6.1.3 on Dot Product

DEFINITION

For \mathbf{u} and \mathbf{v} in \mathbb{R}^n , the **distance between \mathbf{u} and \mathbf{v}** , written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

In \mathbb{R}^2 and \mathbb{R}^3 , this definition of distance coincides with the usual formulas for the Euclidean distance between two points, as the next two examples show.

EXAMPLE 4 Compute the distance between the vectors $\mathbf{u} = (7, 1)$ and $\mathbf{v} = (3, 2)$.

SOLUTION Calculate

$$\begin{aligned}\mathbf{u} - \mathbf{v} &= \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix} \\ \|\mathbf{u} - \mathbf{v}\| &= \sqrt{4^2 + (-1)^2} = \sqrt{17}\end{aligned}$$

The vectors \mathbf{u} , \mathbf{v} , and $\mathbf{u} - \mathbf{v}$ are shown in Fig. 4. When the vector $\mathbf{u} - \mathbf{v}$ is added to \mathbf{v} , the result is \mathbf{u} . Notice that the parallelogram in Fig. 4 shows that the distance from \mathbf{u} to \mathbf{v} is the same as the distance from $\mathbf{u} - \mathbf{v}$ to $\mathbf{0}$. ■

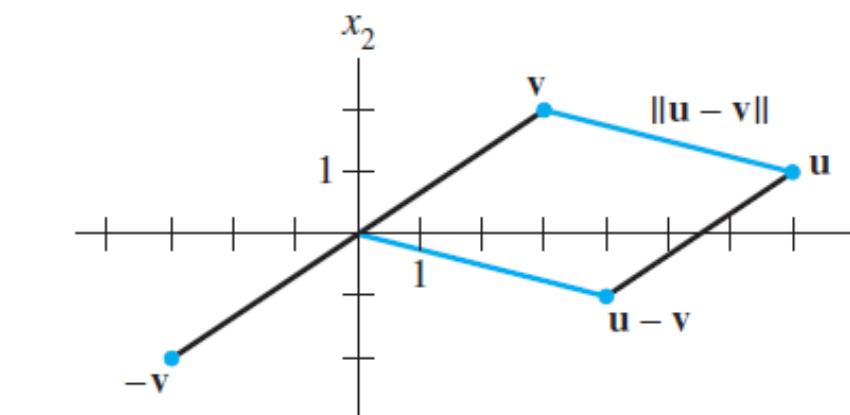


FIGURE 4 The distance between \mathbf{u} and \mathbf{v} is the length of $\mathbf{u} - \mathbf{v}$.

EXAMPLE 5 If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$, then

$$\begin{aligned}\text{dist}(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}\end{aligned}$$

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.1.3**

Lecture : **Orthogonality**

Topic : **Dot Product**

Concept : **The Dot Product**

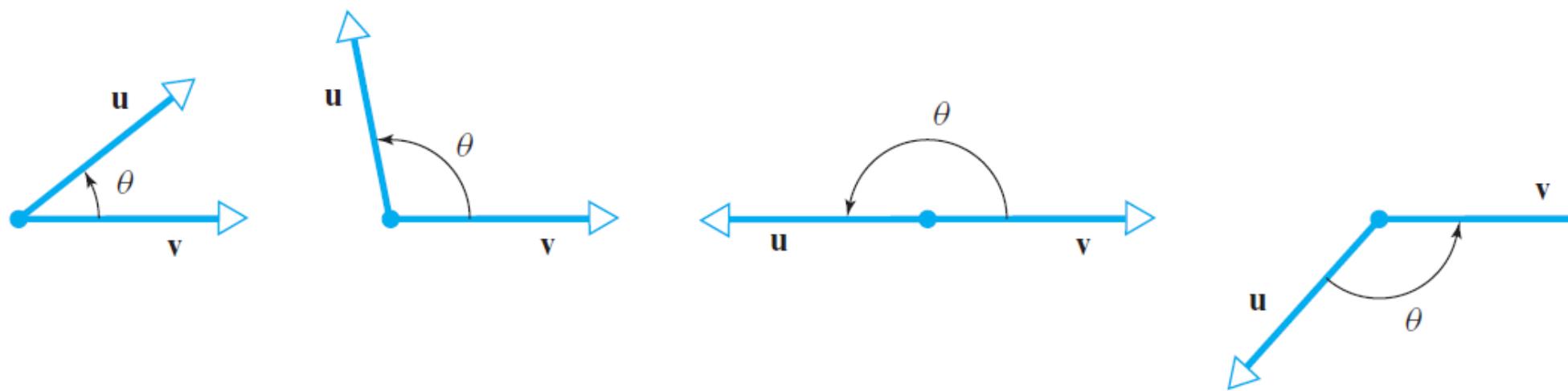
Instructor: A/P Chng Eng Siong
TAs: Zhang Su, Vishal Choudhari

Dot Product

DEFINITION 3 If \mathbf{u} and \mathbf{v} are nonzero vectors in R^2 or R^3 , and if θ is the angle between \mathbf{u} and \mathbf{v} , then the *dot product* (also called the *Euclidean inner product*) of \mathbf{u} and \mathbf{v} is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta \quad (12)$$

If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, then we define $\mathbf{u} \cdot \mathbf{v}$ to be 0.



The angle θ between \mathbf{u} and \mathbf{v} satisfies $0 \leq \theta \leq \pi$.

How to define “angle” between two vectors in R^2 or R^3 ? For this purpose, let \mathbf{u} and \mathbf{v} be nonzero vectors in R^2 or R^3 that have been positioned so that their initial points coincide. We define the *angle between \mathbf{u} and \mathbf{v}* to be the angle θ determined by \mathbf{u} and \mathbf{v} that satisfies the inequalities $0 \leq \theta \leq \pi$ (Figure 3.2.4).

The sign of the dot product reveals information about the angle θ that we can obtain by rewriting Formula (12) as

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \quad (13)$$

Since $0 \leq \theta \leq \pi$, it follows from Formula (13) and properties of the cosine function studied in trigonometry that

- θ is acute if $\mathbf{u} \cdot \mathbf{v} > 0$.
- θ is obtuse if $\mathbf{u} \cdot \mathbf{v} < 0$.
- $\theta = \pi/2$ if $\mathbf{u} \cdot \mathbf{v} = 0$.

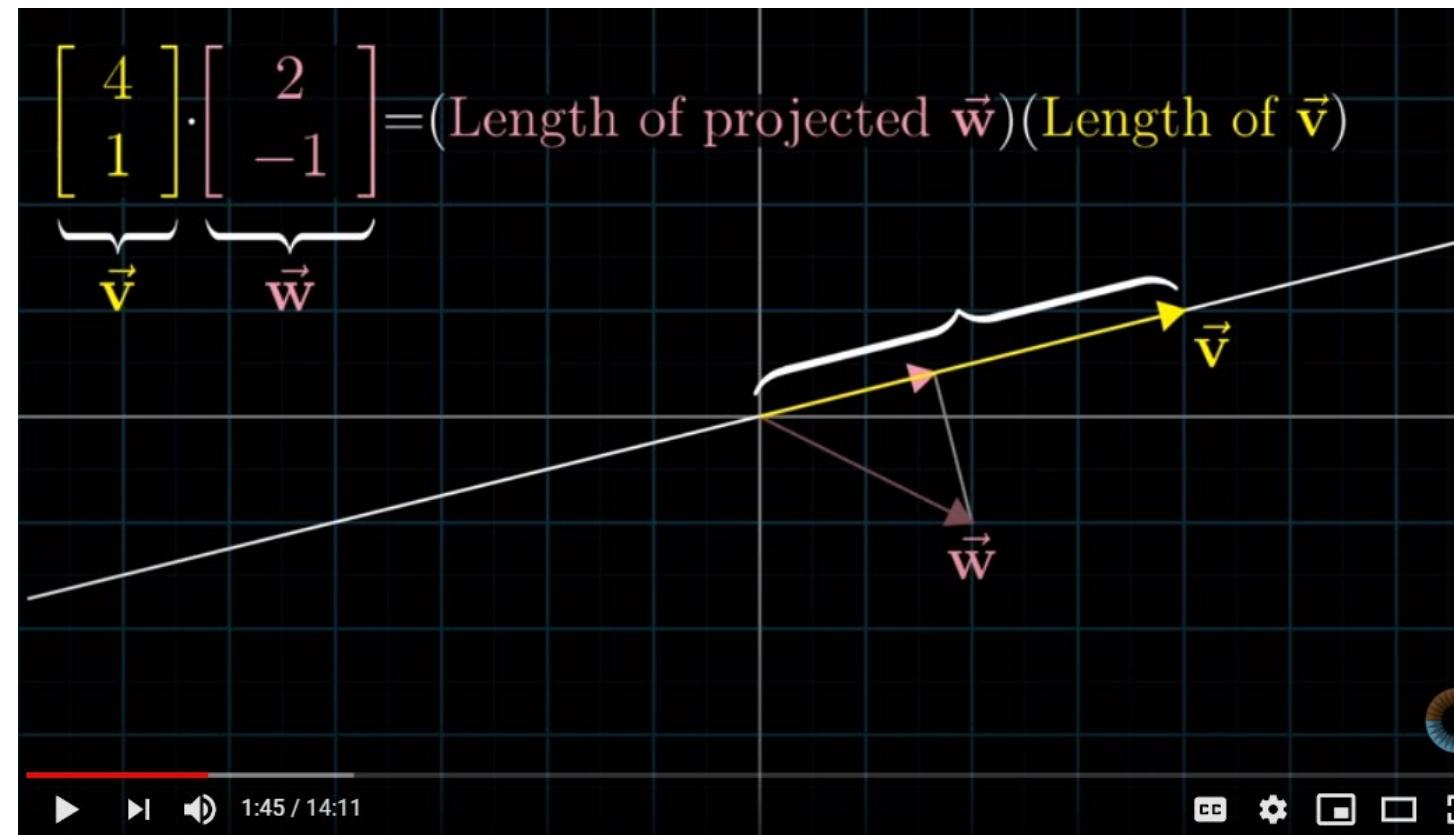
Dot product relates the length of two vectors and the angle (θ) between them.

If vectors \mathbf{u} and \mathbf{v} are unit vectors, i.e., $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$, the dot product is $\mathbf{u} \cdot \mathbf{v} = \cos \theta$.

Ref:

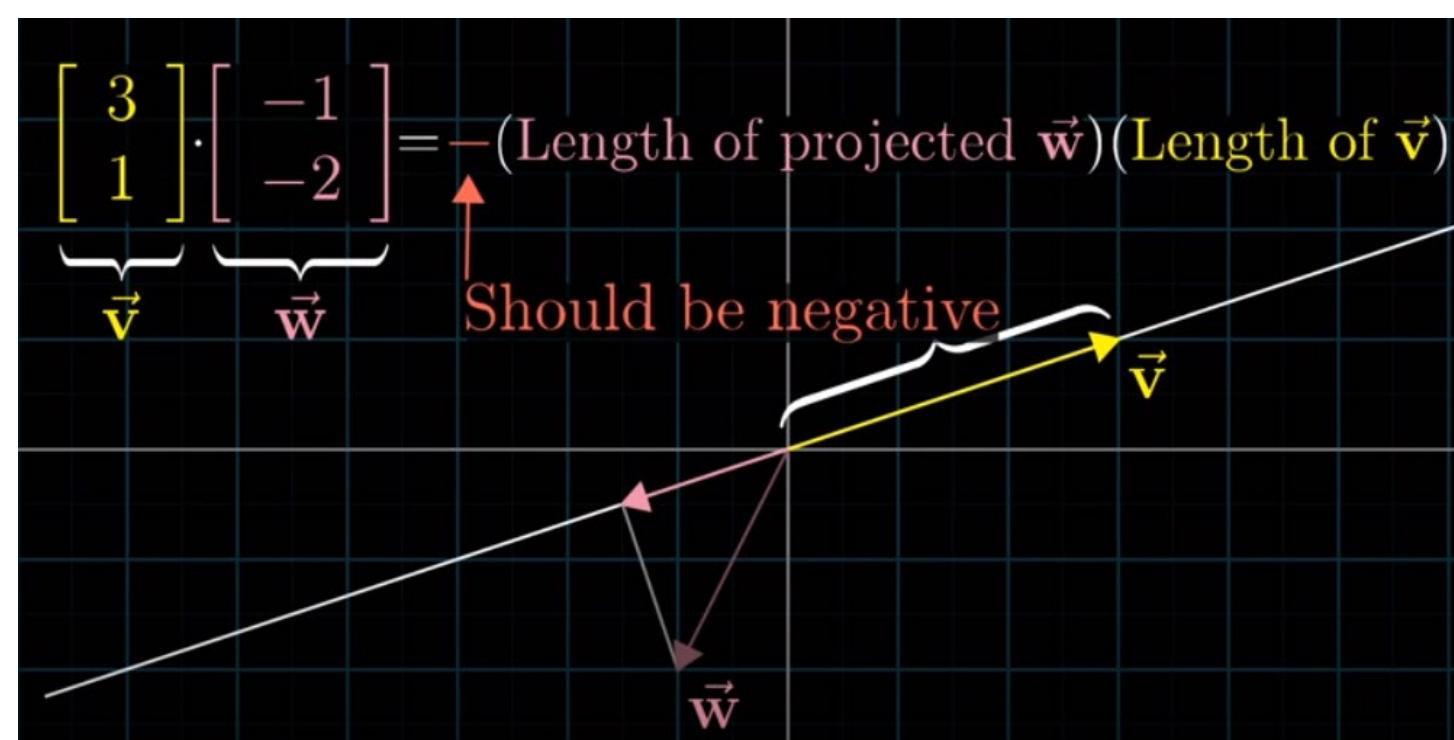
1. Stack Exchange
2. Khan Academy: <https://www.youtube.com/watch?v=KDHuWxy53uM>
3. 3Blue1Brown, Dot Product and Duality: <https://www.youtube.com/watch?v=LyGKycYT2v0>
4. MathsTheBeautiful: https://www.youtube.com/watch?v=QPkKWGq_VOU

Dot Product Interpretation



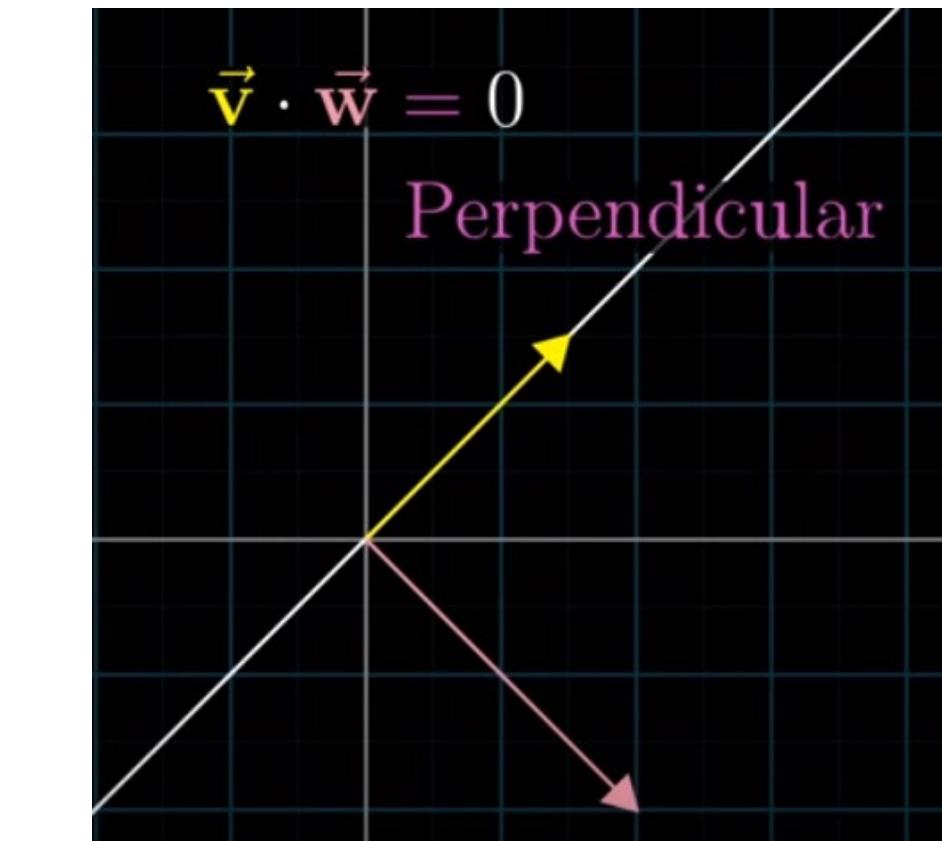
Case 1: When two vectors are pointing nearly towards the same direction, their dot product is +ve.

Angle between vectors: acute.



Case 2: When two vectors are pointing away from one another, their dot product is -ve.

Angle between vectors: obtuse.



Case 3: When two vectors are perpendicular, their dot product is zero.

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= \sum_{i=1}^n v_i w_i = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \\ &= \|\mathbf{v}\| \times \|\mathbf{w}\| \times \cos\theta \end{aligned}$$

Derived in Slide 5

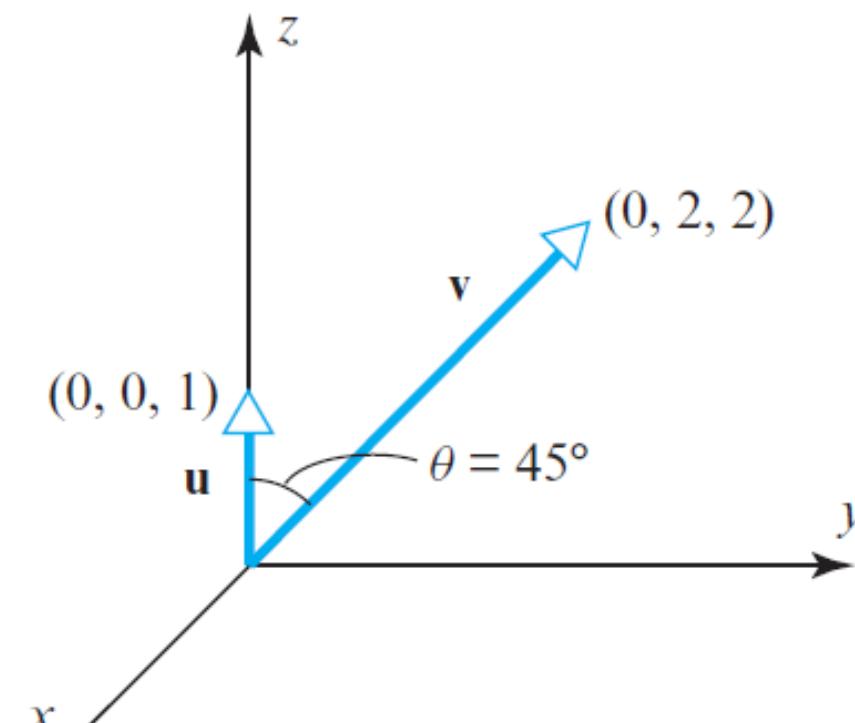
$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \times \|\mathbf{w}\| \times \cos\theta$$

$$\mathbf{v} \cdot \mathbf{w} = (\underbrace{\|\mathbf{w}\| \times \cos\theta}_{\text{Length of projection of } \mathbf{w} \text{ onto } \mathbf{v}}) \times \underbrace{\|\mathbf{v}\|}_{\text{Length of vector } \mathbf{v}}$$

$$\mathbf{v} \cdot \mathbf{w} = (\underbrace{\|\mathbf{v}\| \times \cos\theta}_{\text{Length of projection of } \mathbf{v} \text{ onto } \mathbf{w}}) \times \underbrace{\|\mathbf{w}\|}_{\text{Length of vector } \mathbf{w}}$$

Example

146 Chapter 3 Euclidean Vector Spaces



▲ Figure 3.2.5

► EXAMPLE 5 Dot Product

Find the dot product of the vectors shown in Figure 3.2.5.

Solution The lengths of the vectors are

$$\|\mathbf{u}\| = 1 \text{ and } \|\mathbf{v}\| = \sqrt{8} = 2\sqrt{2}$$

and the cosine of the angle θ between them is

$$\cos(45^\circ) = 1/\sqrt{2}$$

Thus, it follows from Formula (12) that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = (1)(2\sqrt{2})(1/\sqrt{2}) = 2$$

Component Form of the Dot Product

Component Form of the Dot Product

For computational purposes it is desirable to have a formula that expresses the dot product of two vectors in terms of components. We will derive such a formula for vectors in 3-space; the derivation for vectors in 2-space is similar.

Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be two nonzero vectors. If, as shown in Figure 3.2.6, θ is the angle between \mathbf{u} and \mathbf{v} , then the law of cosines yields

$$\|\overrightarrow{PQ}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\| \cos \theta \quad (14)$$

Since $\overrightarrow{PQ} = \mathbf{v} - \mathbf{u}$, we can rewrite (14) as

$$\|\mathbf{u}\|\|\mathbf{v}\| \cos \theta = \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{v} - \mathbf{u}\|^2)$$

or

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{v} - \mathbf{u}\|^2)$$

Substituting

$$\|\mathbf{u}\|^2 = u_1^2 + u_2^2 + u_3^2, \quad \|\mathbf{v}\|^2 = v_1^2 + v_2^2 + v_3^2$$

and

$$\|\mathbf{v} - \mathbf{u}\|^2 = (v_1 - u_1)^2 + (v_2 - u_2)^2 + (v_3 - u_3)^2$$

we obtain, after simplifying,

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 \quad (15)$$

The companion formula for vectors in 2-space is

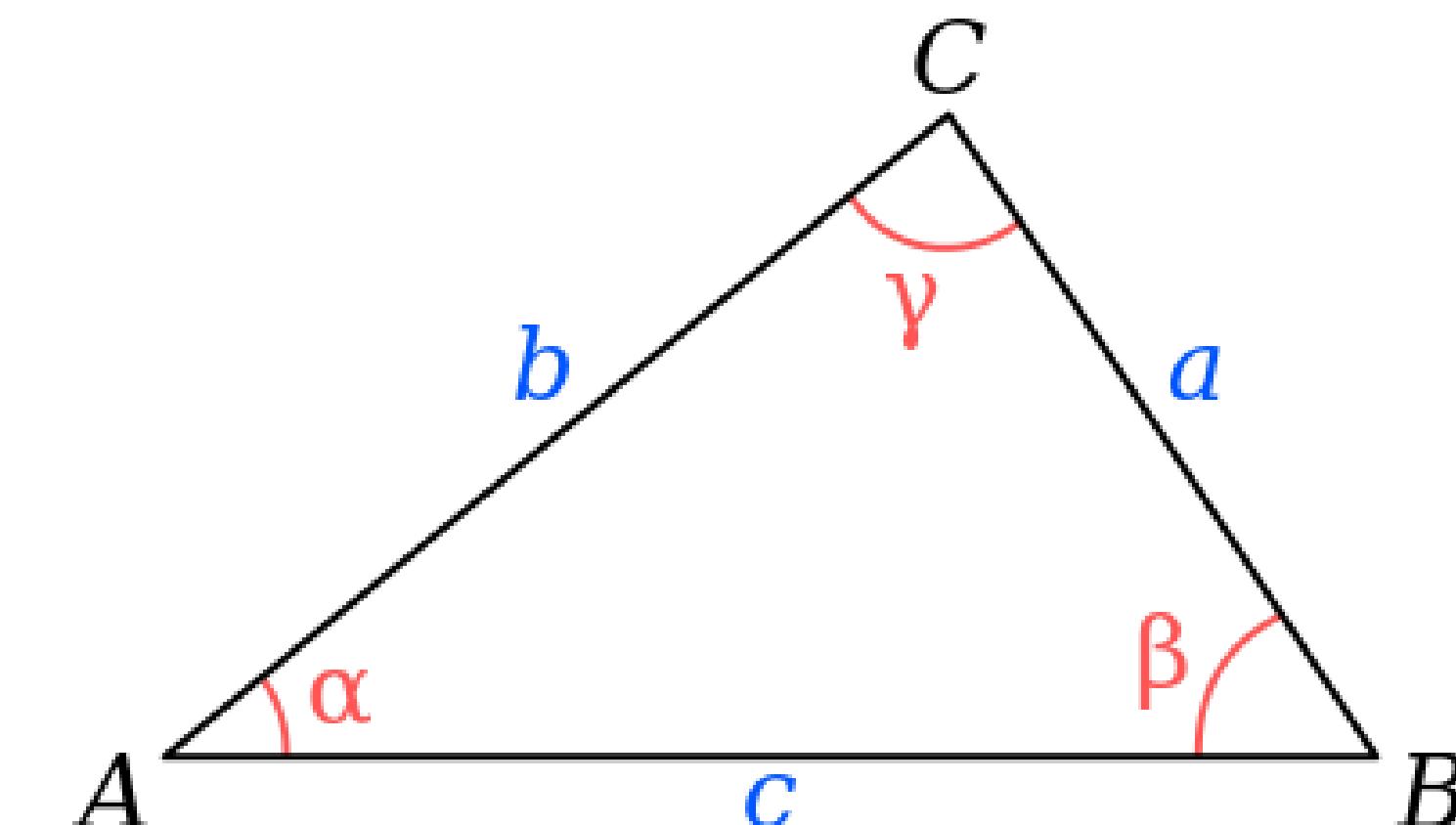
$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 \quad (16)$$

Motivated by the pattern in Formulas (15) and (16), we make the following definition.

DEFINITION 4 If $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ are vectors in R^n , then the *dot product* (also called the *Euclidean inner product*) of \mathbf{u} and \mathbf{v} is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \quad (17)$$

Reviewing Law of Cosines



$$c^2 = a^2 + b^2 - 2ab \cos \gamma,$$

Example

► EXAMPLE 6 Calculating Dot Products Using Components

- (a) Use Formula (15) to compute the dot product of the vectors \mathbf{u} and \mathbf{v} in Example 5.
(b) Calculate $\mathbf{u} \cdot \mathbf{v}$ for the following vectors in R^4 :

$$\mathbf{u} = (-1, 3, 5, 7), \quad \mathbf{v} = (-3, -4, 1, 0)$$

Solution (a) The component forms of the vectors are $\mathbf{u} = (0, 0, 1)$ and $\mathbf{v} = (0, 2, 2)$. Thus,

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

which agrees with the result obtained geometrically in Example 5.

Solution (b)

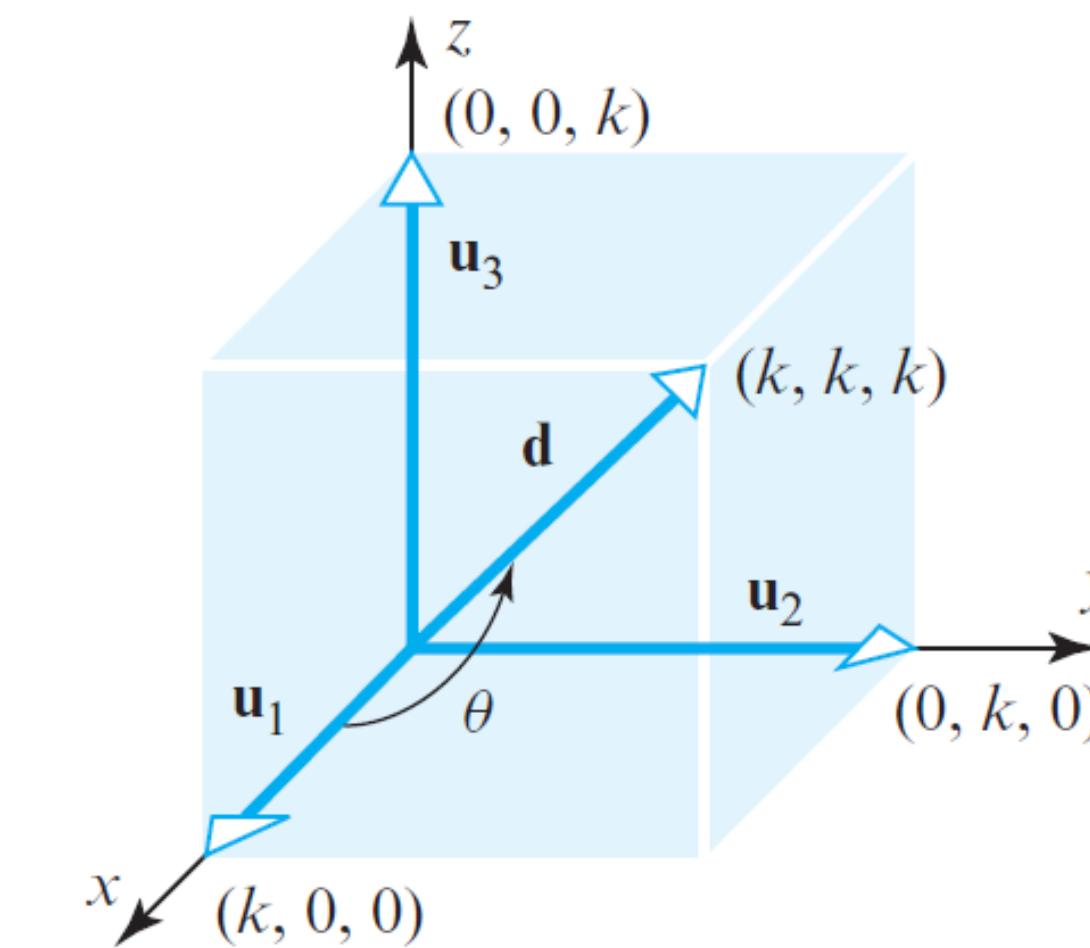
$$\mathbf{u} \cdot \mathbf{v} = (-1)(-3) + (3)(-4) + (5)(1) + (7)(0) = -4$$

In the special case where $\mathbf{u} = \mathbf{v}$ in Definition 4, we obtain the relationship

$$\mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + \cdots + v_n^2 = \|\mathbf{v}\|^2 \quad (18)$$

This yields the following formula for expressing the length of a vector in terms of a dot product:

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} \quad (19)$$



▲ Figure 3.2.7

► EXAMPLE 7 A Geometry Problem Solved Using Dot Product

Find the angle between a diagonal of a cube and one of its edges.

Solution Let k be the length of an edge and introduce a coordinate system as shown in Figure 3.2.7. If we let $\mathbf{u}_1 = (k, 0, 0)$, $\mathbf{u}_2 = (0, k, 0)$, and $\mathbf{u}_3 = (0, 0, k)$, then the vector $\mathbf{d} = (k, k, k) = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$

is a diagonal of the cube. It follows from Formula (13) that the angle θ between \mathbf{d} and the edge \mathbf{u}_1 satisfies

$$\cos \theta = \frac{\mathbf{u}_1 \cdot \mathbf{d}}{\|\mathbf{u}_1\| \|\mathbf{d}\|} = \frac{k^2}{(k)(\sqrt{3k^2})} = \frac{1}{\sqrt{3}}$$

With the help of a calculator we obtain

$$\theta = \cos^{-1} \left(\frac{1}{\sqrt{3}} \right) \approx 54.74^\circ \quad \blacktriangleleft$$

Properties of Dot Product

Dot products have many of the same algebraic properties as products of real numbers.

THEOREM 3.2.2 *If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors in R^n , and if k is a scalar, then:*

- (a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ [Symmetry property]
- (b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ [Distributive property]
- (c) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ [Homogeneity property]
- (d) $\mathbf{v} \cdot \mathbf{v} \geq 0$ and $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$ [Positivity property]

Proof (c) Let $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$. Then

$$\begin{aligned} k(\mathbf{u} \cdot \mathbf{v}) &= k(u_1v_1 + u_2v_2 + \dots + u_nv_n) \\ &= (ku_1)v_1 + (ku_2)v_2 + \dots + (ku_n)v_n = (k\mathbf{u}) \cdot \mathbf{v} \end{aligned}$$

Proof (d) The result follows from parts (a) and (b) of Theorem 3.2.1 and the fact that

$$\mathbf{v} \cdot \mathbf{v} = v_1v_1 + v_2v_2 + \dots + v_nv_n = v_1^2 + v_2^2 + \dots + v_n^2 = \|\mathbf{v}\|^2 \quad \blacktriangleleft$$

THEOREM 3.2.3 *If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors in R^n , and if k is a scalar, then:*

- (a) $\mathbf{0} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{0} = 0$
- (b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- (c) $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$
- (d) $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w}$
- (e) $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (kv)$

Proof (b)

$$\begin{aligned} (\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} &= \mathbf{w} \cdot (\mathbf{u} + \mathbf{v}) && [\text{By symmetry}] \\ &= \mathbf{w} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{v} && [\text{By distributivity}] \\ &= \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w} && [\text{By symmetry}] \quad \blacktriangleleft \end{aligned}$$

► EXAMPLE 8 Calculating with Dot Products

$$\begin{aligned} (\mathbf{u} - 2\mathbf{v}) \cdot (3\mathbf{u} + 4\mathbf{v}) &= \mathbf{u} \cdot (3\mathbf{u} + 4\mathbf{v}) - 2\mathbf{v} \cdot (3\mathbf{u} + 4\mathbf{v}) \\ &= 3(\mathbf{u} \cdot \mathbf{u}) + 4(\mathbf{u} \cdot \mathbf{v}) - 6(\mathbf{v} \cdot \mathbf{u}) - 8(\mathbf{v} \cdot \mathbf{v}) \\ &= 3\|\mathbf{u}\|^2 - 2(\mathbf{u} \cdot \mathbf{v}) - 8\|\mathbf{v}\|^2 \quad \blacktriangleleft \end{aligned}$$

Property and Example

Table 1

Form	Dot Product	Example	
\mathbf{u} a column matrix and \mathbf{v} a column matrix	$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \mathbf{v}^T \mathbf{u}$	$\mathbf{u} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}$	$\mathbf{u}^T \mathbf{v} = [1 \quad -3 \quad 5] \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix} = -7$

If A is an $n \times n$ matrix and \mathbf{u} and \mathbf{v} are $n \times 1$ matrices, then it follows from the first row in Table 1 and properties of the transpose that

$$\begin{aligned} A\mathbf{u} \cdot \mathbf{v} &= \mathbf{v}^T(A\mathbf{u}) = (\mathbf{v}^T A)\mathbf{u} = (A^T \mathbf{v})^T \mathbf{u} = \mathbf{u} \cdot A^T \mathbf{v} \\ \mathbf{u} \cdot A\mathbf{v} &= (A\mathbf{v})^T \mathbf{u} = (\mathbf{v}^T A^T)\mathbf{u} = \mathbf{v}^T(A^T \mathbf{u}) = A^T \mathbf{u} \cdot \mathbf{v} \end{aligned}$$

The resulting formulas

$$A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T \mathbf{v} \tag{26}$$

$$\mathbf{u} \cdot A\mathbf{v} = A^T \mathbf{u} \cdot \mathbf{v} \tag{27}$$

provide an important link between multiplication by an $n \times n$ matrix A and multiplication by A^T .

► EXAMPLE 9 Verifying that $A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T \mathbf{v}$

Suppose that

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix}$$

Then

$$A\mathbf{u} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \\ 5 \end{bmatrix}$$

$$A^T \mathbf{v} = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 4 & 0 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} -7 \\ 4 \\ -1 \end{bmatrix}$$

from which we obtain

$$A\mathbf{u} \cdot \mathbf{v} = 7(-2) + 10(0) + 5(5) = 11$$

$$\mathbf{u} \cdot A^T \mathbf{v} = (-1)(-7) + 2(4) + 4(-1) = 11$$

Thus, $A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T \mathbf{v}$ as guaranteed by Formula (26). We leave it for you to verify that Formula (27) also holds. ◀

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.1.4**

Lecture : **Orthogonality**

Topic : **Dot Product**

Concept : **Important Inequalities**

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

Cauchy-Schwarz Inequality

DEFINITION 3 If \mathbf{u} and \mathbf{v} are nonzero vectors in R^2 or R^3 , and if θ is the angle between \mathbf{u} and \mathbf{v} , then the *dot product* (also called the *Euclidean inner product*) of \mathbf{u} and \mathbf{v} is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta \quad (12)$$

If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, then we define $\mathbf{u} \cdot \mathbf{v}$ to be 0.

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

Formula (20) is not defined unless its argument satisfies the inequalities

$$-1 \leq \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1 \quad (21)$$

Fortunately, these inequalities *do* hold for all nonzero vectors in R^n as a result of the following fundamental result known as the *Cauchy-Schwarz inequality*.

THEOREM 3.2.4 Cauchy-Schwarz Inequality

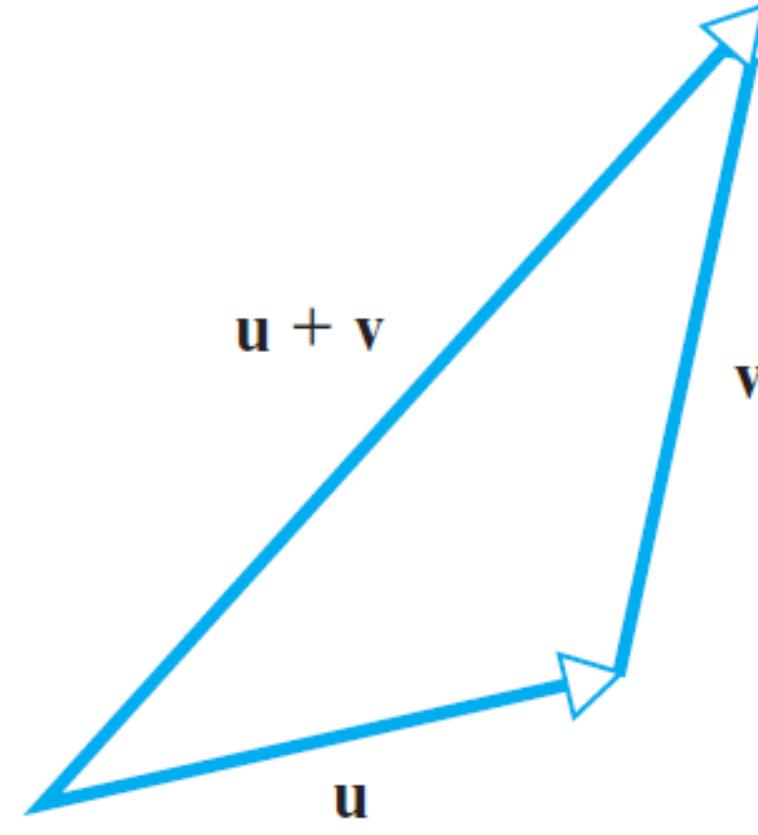
If $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ are vectors in R^n , then

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\| \quad (22)$$

or in terms of components

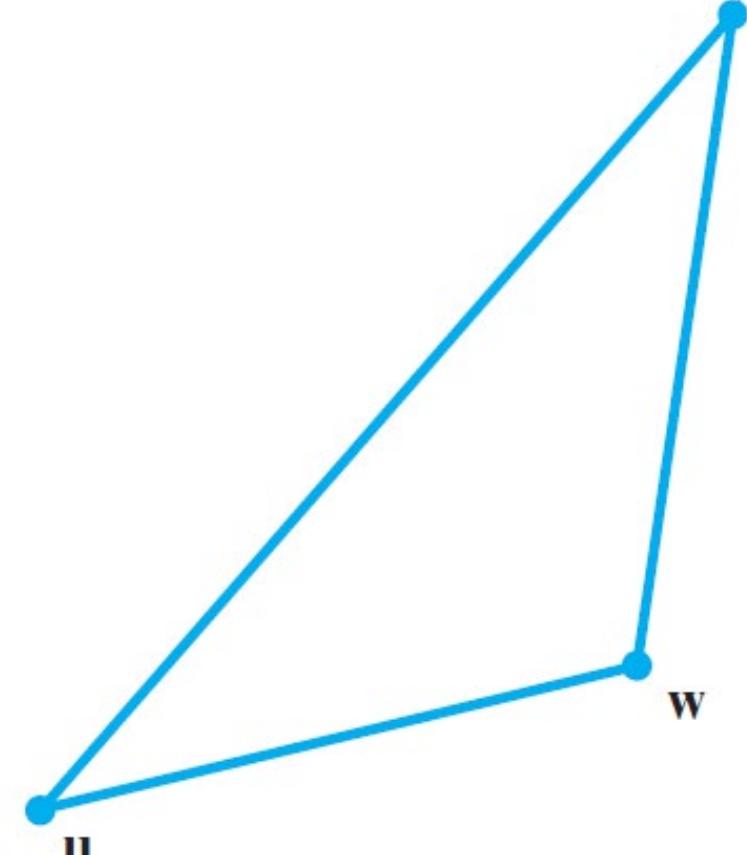
$$|u_1 v_1 + u_2 v_2 + \cdots + u_n v_n| \leq (u_1^2 + u_2^2 + \cdots + u_n^2)^{1/2} (v_1^2 + v_2^2 + \cdots + v_n^2)^{1/2} \quad (23)$$

Triangle Inequality



$$\|u + v\| \leq \|u\| + \|v\|$$

▲ Figure 3.2.8



$$d(u, v) \leq d(u, w) + d(w, v)$$

▲ Figure 3.2.9

THEOREM 3.2.5 If u , v , and w are vectors in R^n , then:

- (a) $\|u + v\| \leq \|u\| + \|v\|$ [Triangle inequality for vectors]
- (b) $d(u, v) \leq d(u, w) + d(w, v)$ [Triangle inequality for distances]

Proof:

Proof (a)

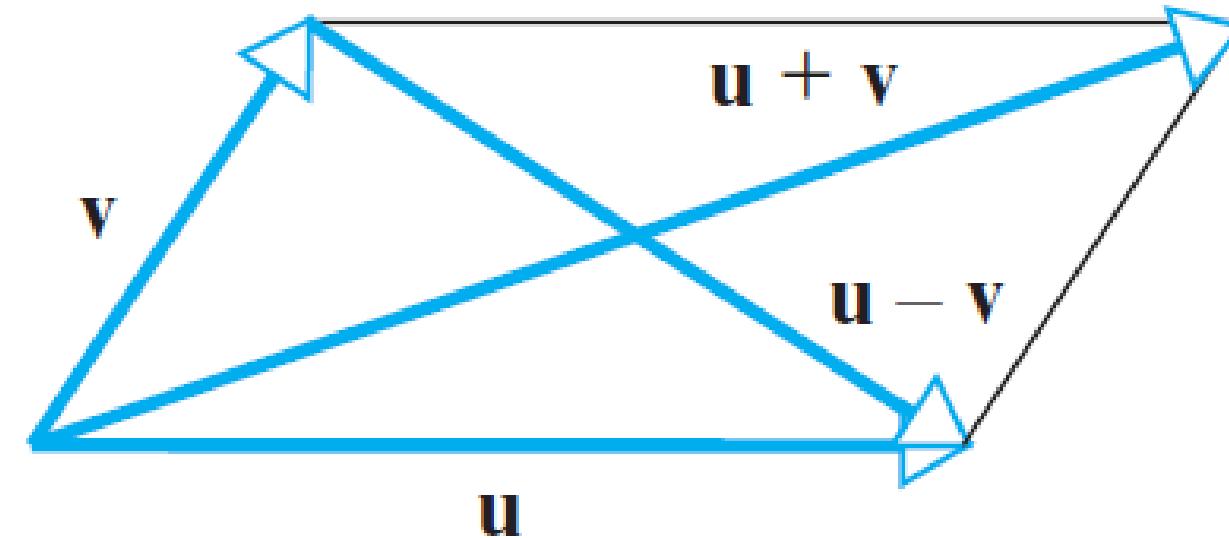
$$\begin{aligned}\|u + v\|^2 &= (u + v) \cdot (u + v) = (u \cdot u) + 2(u \cdot v) + (v \cdot v) \\&= \|u\|^2 + 2(u \cdot v) + \|v\|^2 \\&\leq \|u\|^2 + 2|u \cdot v| + \|v\|^2 \quad \leftarrow \text{Property of absolute value} \\&\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \quad \leftarrow \text{Cauchy-Schwarz inequality} \\&= (\|u\| + \|v\|)^2\end{aligned}$$

This completes the proof since both sides of the inequality in part (a) are nonnegative.

Proof (b) It follows from part (a) and Formula (11) that

$$\begin{aligned}d(u, v) &= \|u - v\| = \|(u - w) + (w - v)\| \\&\leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v) \quad \blacktriangleleft\end{aligned}$$

Parallelogram Equation for Vectors



▲ Figure 3.2.10

THEOREM 3.2.6 Parallelogram Equation for Vectors

If \mathbf{u} and \mathbf{v} are vectors in R^n , then

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2) \quad (24)$$

Proof:

THEOREM 3.2.6 Parallelogram Equation for Vectors

If \mathbf{u} and \mathbf{v} are vectors in R^n , then

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2) \quad (24)$$

Proof

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) + (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) \\ &= 2(\mathbf{u} \cdot \mathbf{u}) + 2(\mathbf{v} \cdot \mathbf{v}) \\ &= 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2) \quad \blacktriangleleft \end{aligned}$$

THEOREM 3.2.7 If \mathbf{u} and \mathbf{v} are vectors in R^n with the Euclidean inner product, then

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2 \quad (25)$$

Proof

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \\ \|\mathbf{u} - \mathbf{v}\|^2 &= (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 - 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \end{aligned}$$

from which (25) follows by simple algebra. \blacktriangleleft

Reference

Materials in these slides have been taken from:

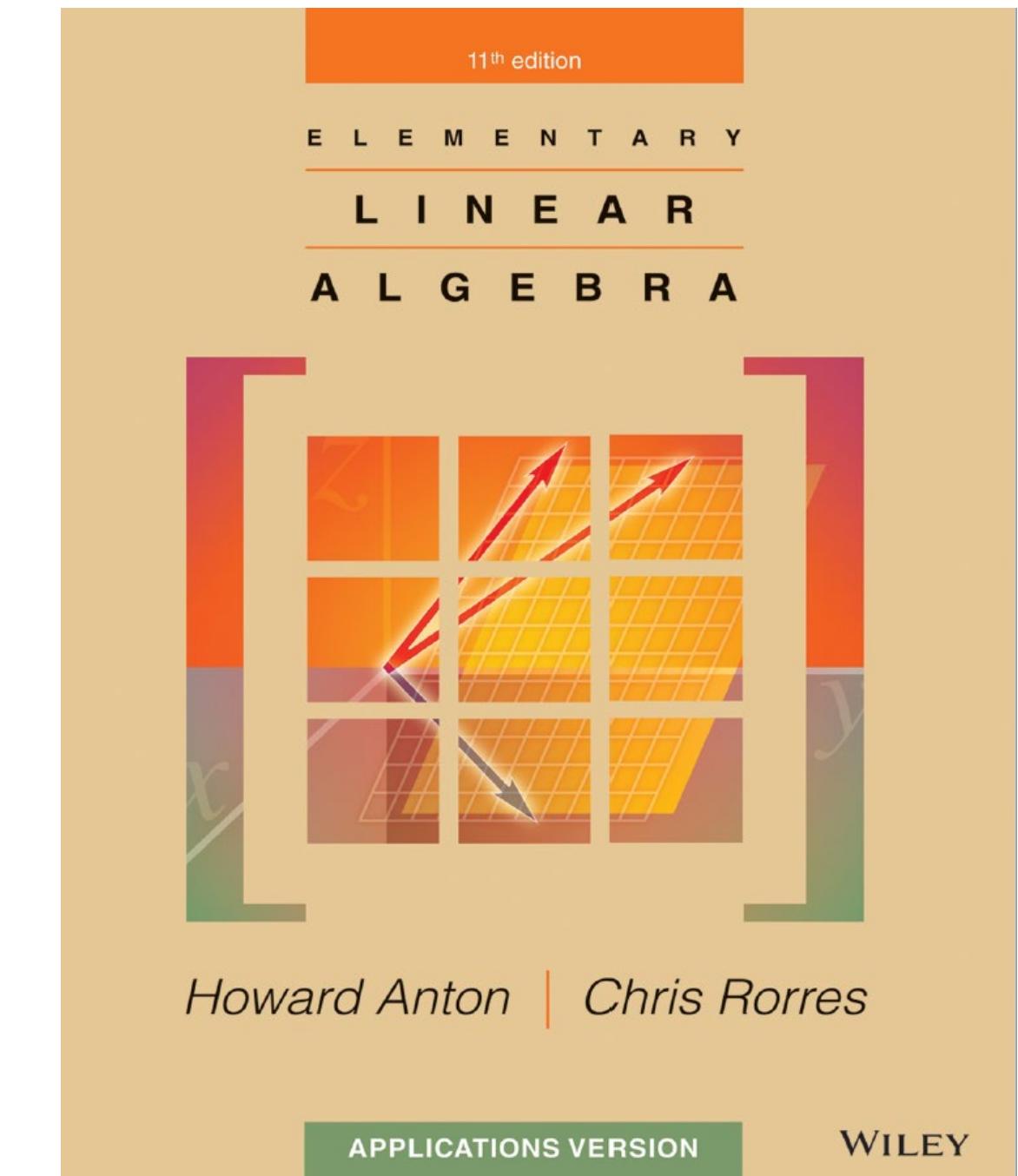
Anton and Rorres, “Linear Algebra”, 11th edition , Wiley.

Chapter: 3.1, 3.2

Euclidean Vector Spaces

3.1 Vectors in 2-Space, 3-Space, and n -Space 131

3.2 Norm, Dot Product, and Distance in R^n 142



CX1104: Linear Algebra for Computing

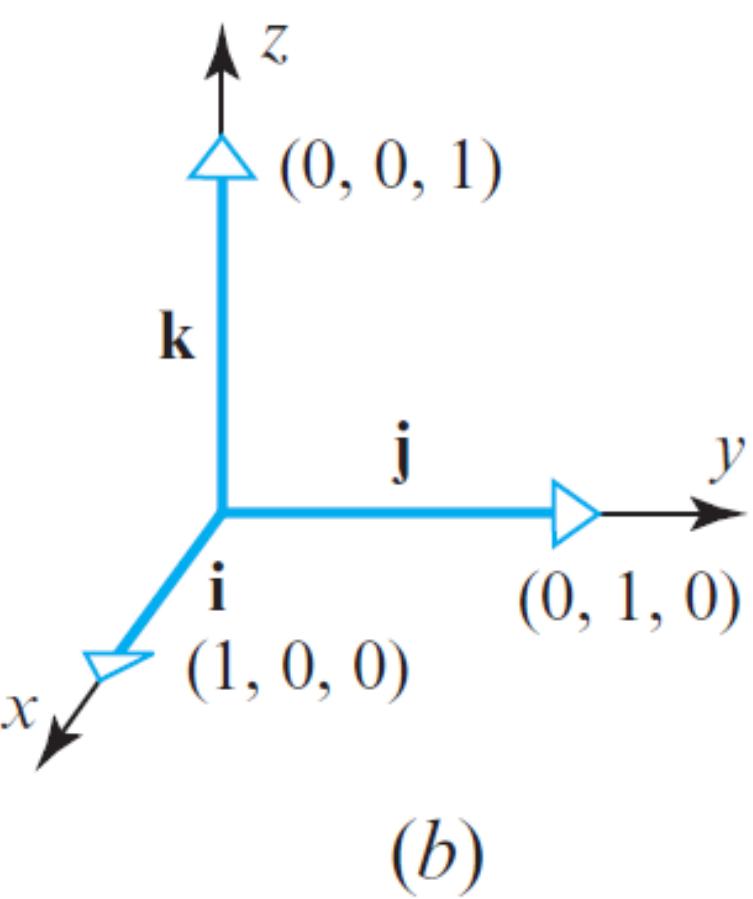
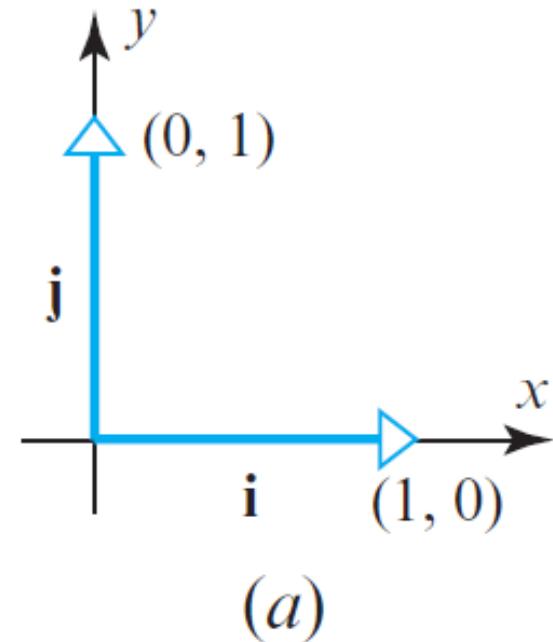
$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

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

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

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

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

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

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



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

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

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

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

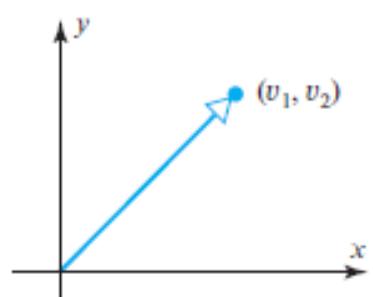
$$\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0) \quad [\text{plane}]$$

Thus, Equation (1) can be written as

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

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

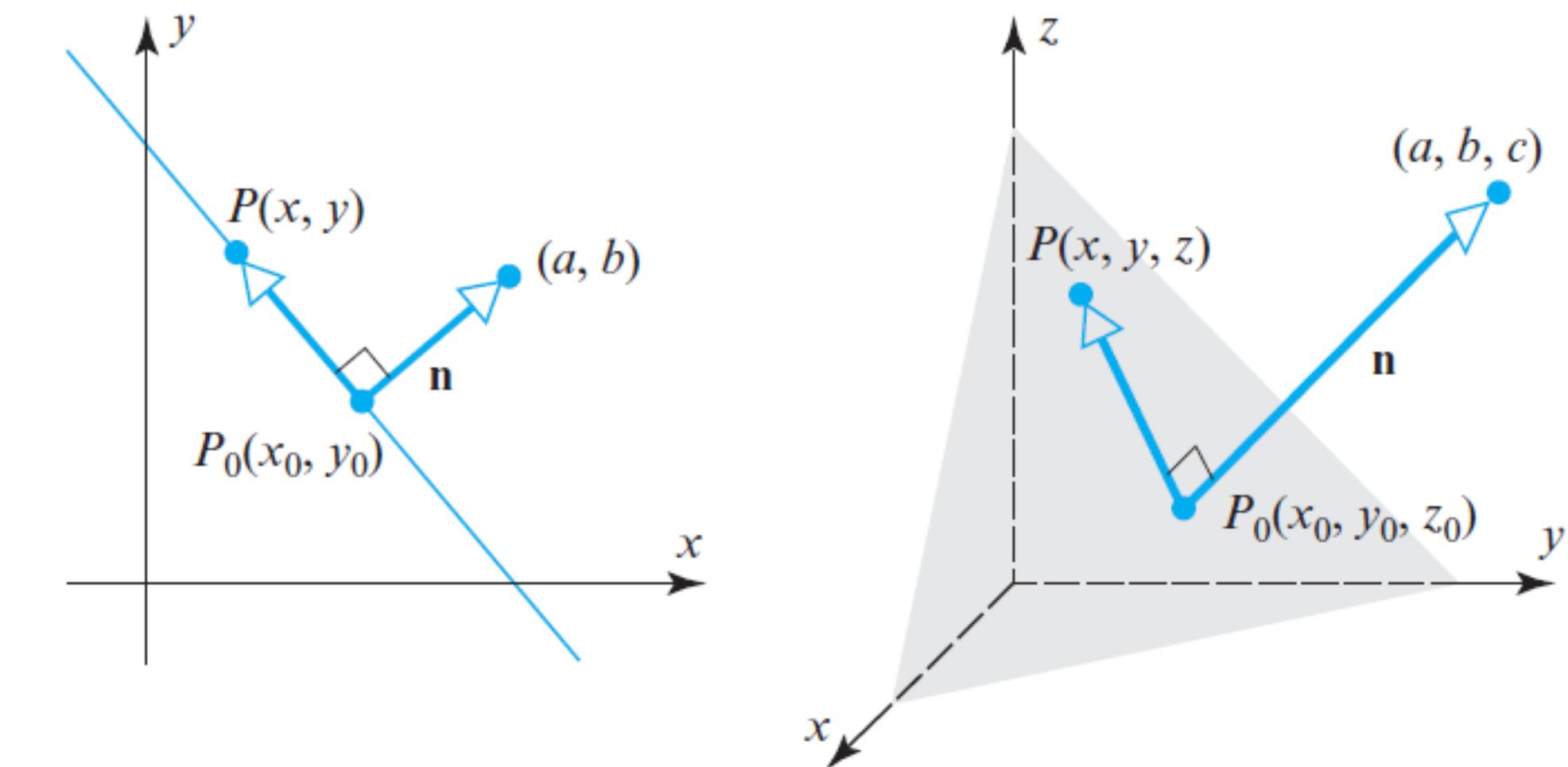
These are called the **point-normal** equations of the line and plane.



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

Note, \mathbf{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

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

Thus, Equation (1) can be written as

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

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

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

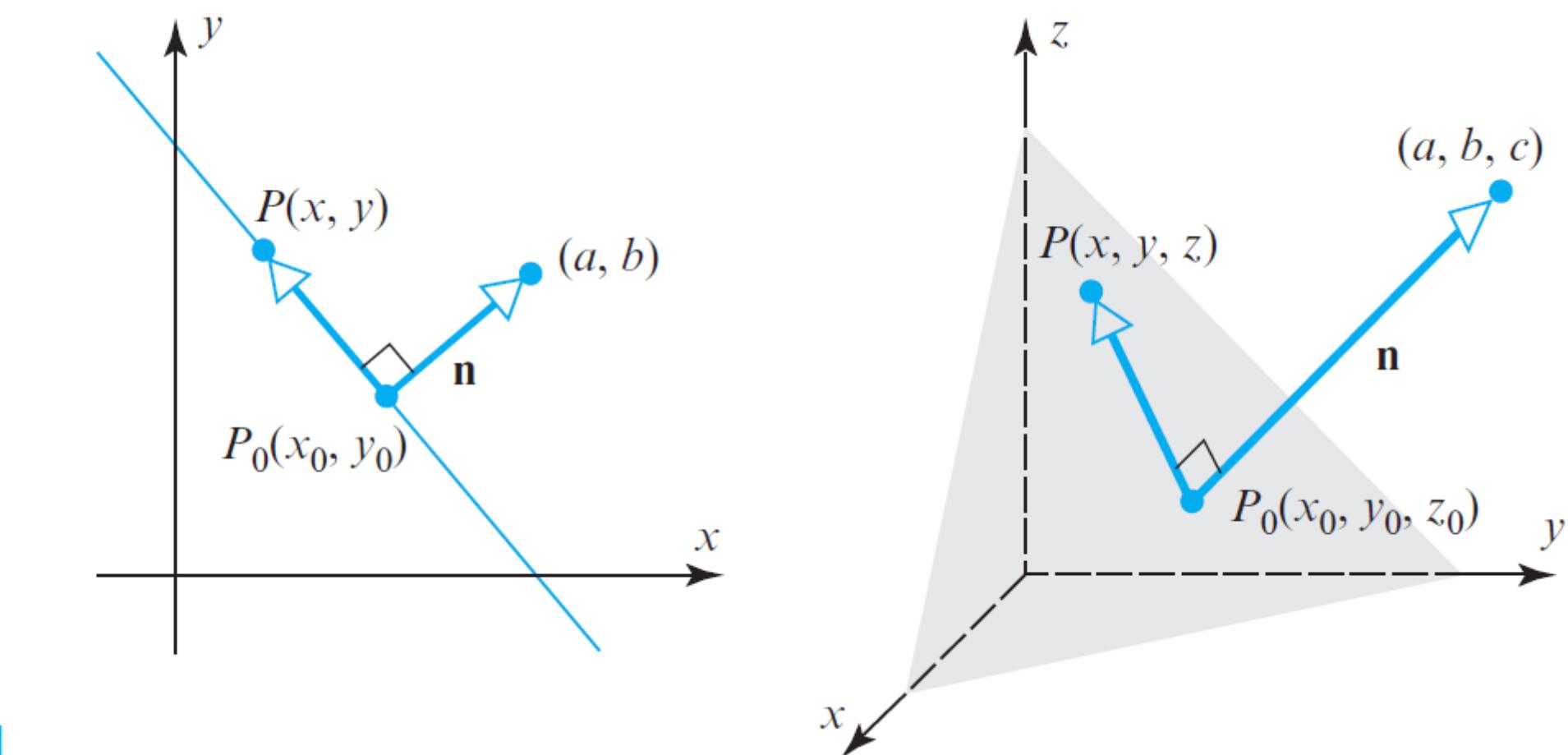
$$ax + by + c = 0 \quad (4)$$

represents a line in \mathbb{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

$$ax + by + cz + d = 0 \quad (5)$$

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



► Figure 3.3.1

Lines and Planes Determined by Points and Normals

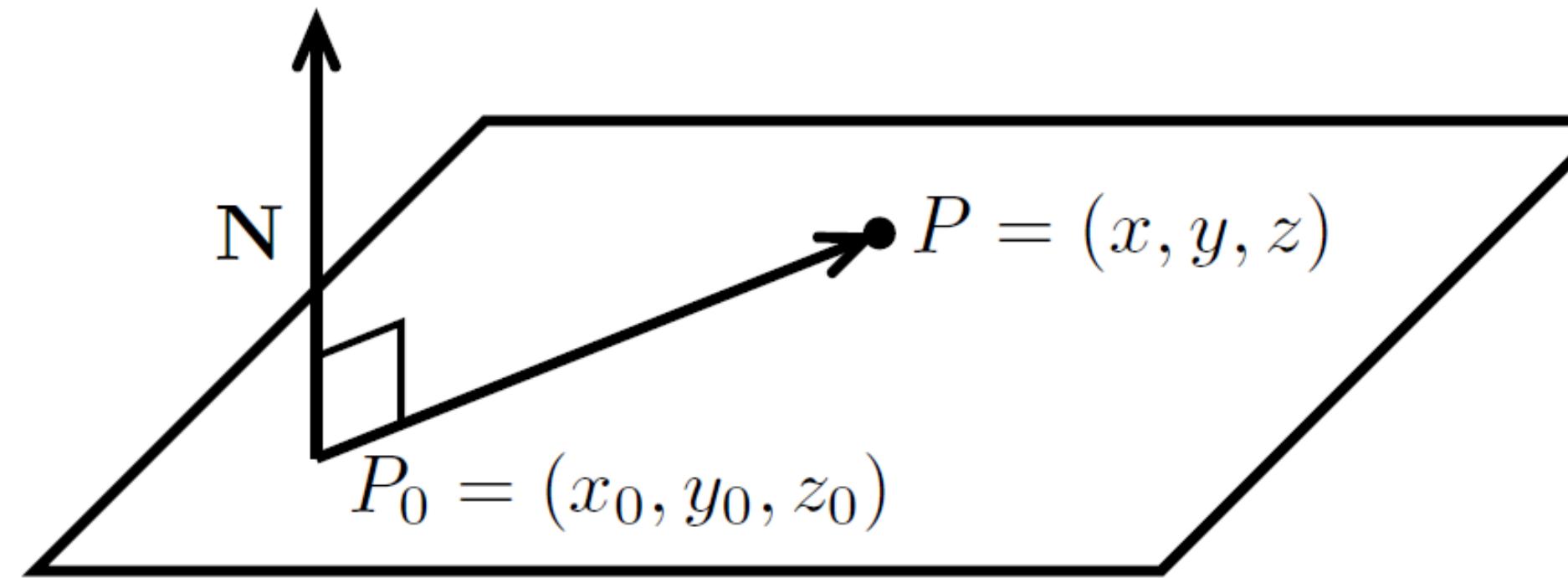
Let $P = (x, y, z)$ be an arbitrary point in the plane. Then the vector $\overrightarrow{P_0P}$ is in the plane and therefore orthogonal to \mathbf{N} . This means

$$\mathbf{N} \cdot \overrightarrow{P_0P} = 0$$

$$\Leftrightarrow \langle a, b, c \rangle \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

$$\Leftrightarrow a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

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

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 .

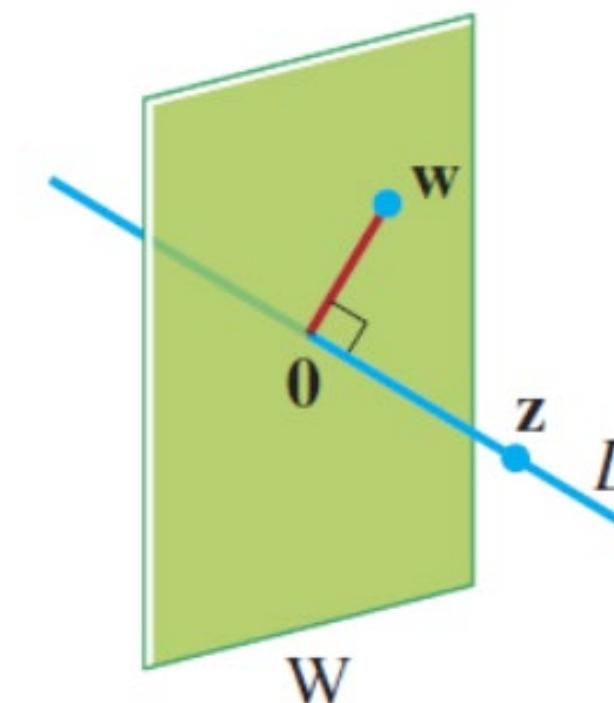


FIGURE 7

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,

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

Orthogonal Complements

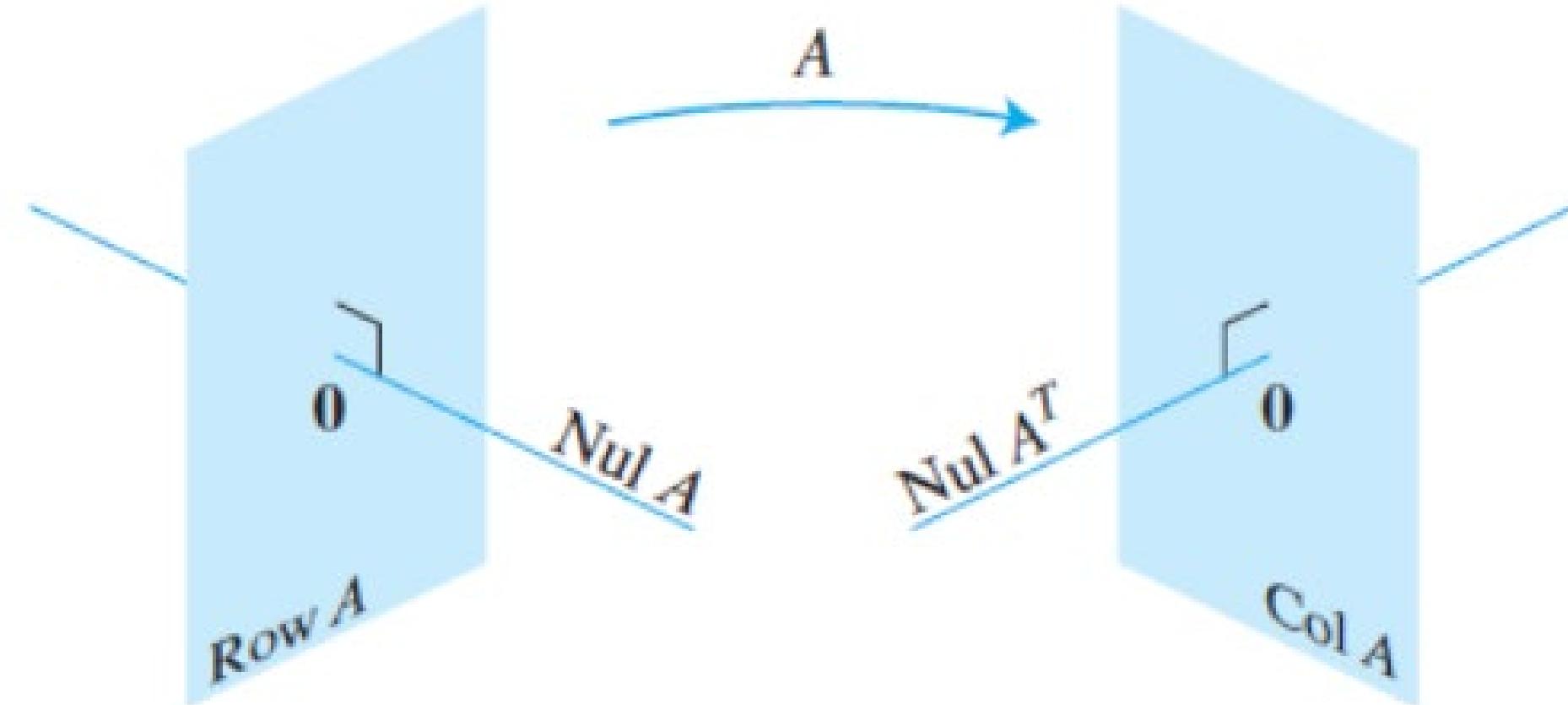


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A .

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 $\text{Nul } A$, 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 $\text{Row } A$. Conversely, if \mathbf{x} is orthogonal to $\text{Row } A$, 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 $\text{Row } A^T = \text{Col } A$. ■

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 $\text{Nul } A = (\text{Row } A)^\perp$ is established by showing that $\text{Nul } A$ is a subset of $(\text{Row } A)^\perp$ and $(\text{Row } A)^\perp$ is a subset of $\text{Nul } A$. That is, an arbitrary element \mathbf{x} in $\text{Nul } A$ 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 $\text{Nul } A$.

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.2.2 (mod by Tay Kian Boon)**

Lecture : **Orthogonality**

Topic : **Orthogonality**

Concept : **Orthogonal Projections**

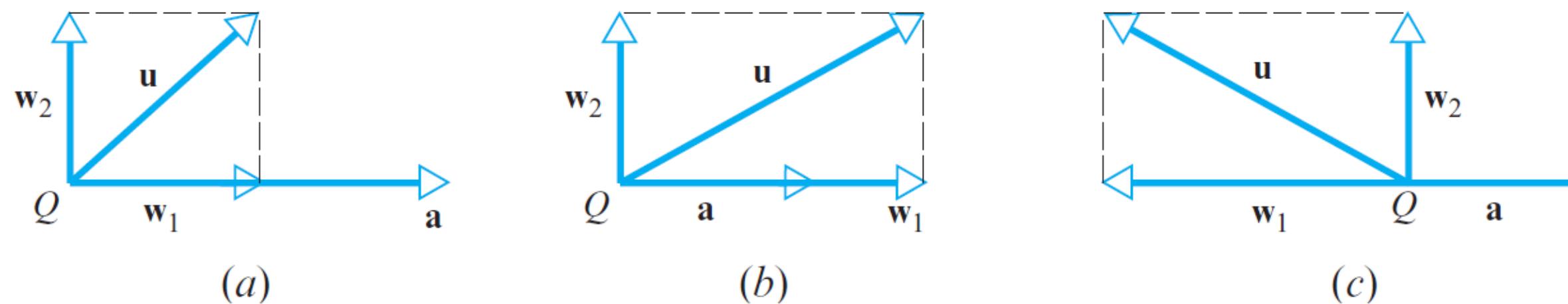
Instructor: A/P Chng Eng Siong

TAs: Zhang Su, Vishal Choudhari

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

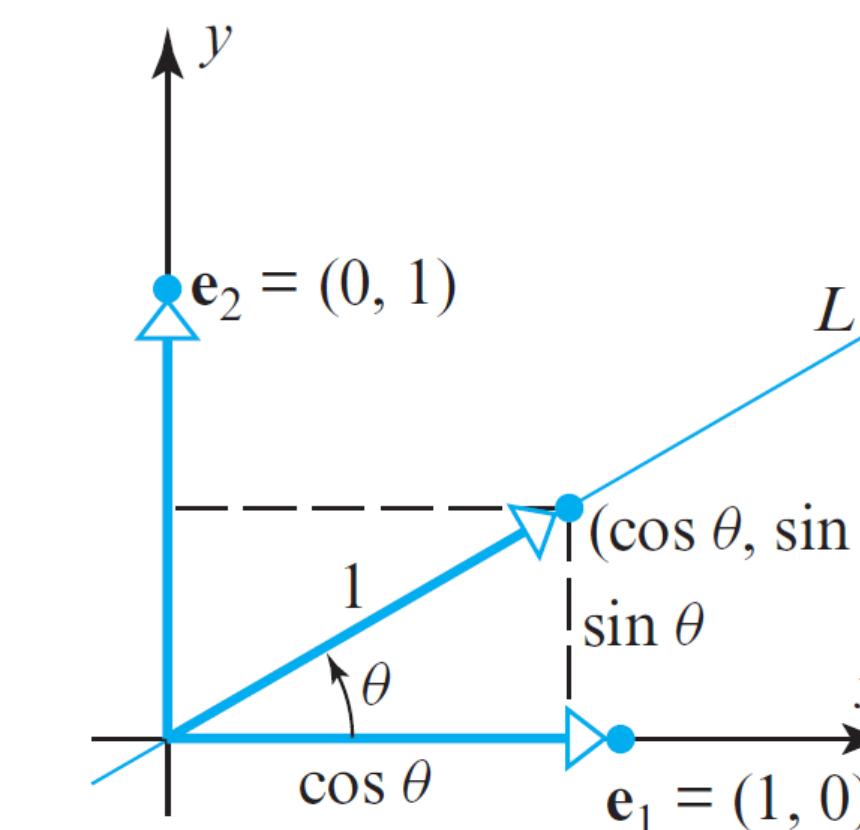
Recall from Example 11 of Section 4.2 that the standard unit vectors

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

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

A diagram illustrating the decomposition of a complex number $z = a + bi$ in the complex plane. The real part a is represented by a horizontal arrow along the positive x-axis, and the imaginary part b is represented by a vertical arrow along the positive y-axis. The vector z is shown originating from the origin. A right-angle symbol indicates the perpendicularity of the projections onto the axes. The resulting vectors are labeled \mathbf{w}_1 and \mathbf{w}_2 . The angle θ is shown between the positive x-axis and the vector z . The unit circle is also shown.

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:

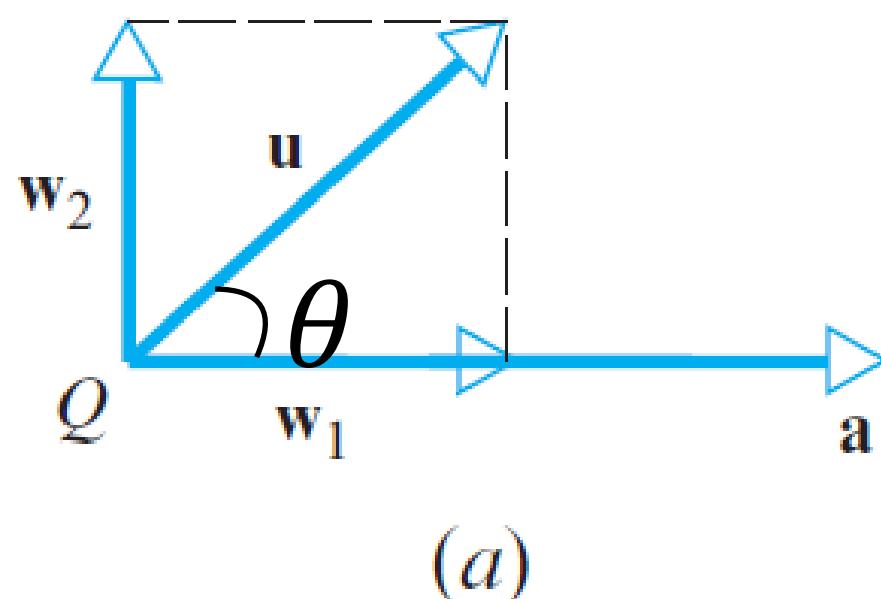
<https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:complex/x9e81a4f98389efdf:complex-polar/v/polar-form-complex-number>

<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 \mathbf{u} on \mathbf{a}* or sometimes *the vector component of \mathbf{u} along \mathbf{a}* , and the vector \mathbf{w}_2 is called the vector *component of \mathbf{u} orthogonal to \mathbf{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,

$$\mathbf{w}_1 = \text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (\text{vector component of } \mathbf{u} \text{ along } \mathbf{a}) = \|\mathbf{u}\| \cos \theta \quad (10)$$

$$\mathbf{w}_2 = \mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \quad (\text{vector component of } \mathbf{u} \text{ orthogonal to } \mathbf{a}) \quad (11)$$

► EXAMPLE 5 Vector Component of \mathbf{u} Along \mathbf{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

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

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

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

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. $\text{Proj}_{\mathbf{v}}\mathbf{u}$, 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 $\text{Proj}_{\mathbf{v}}\mathbf{u}$ 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.

$$\text{Proj}_{\mathbf{v}}\mathbf{u} = \mathbf{v} \left(\frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{v}\|^2} \right), \quad (2.34)$$

$$\mathbf{w}_{\mathbf{v}} = \mathbf{u} - \text{Proj}_{\mathbf{v}}\mathbf{u}, \quad (2.35)$$

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 $\|\mathbf{v}\|^2$? (eqn 2.34).

Ans: it is to normalize vector \mathbf{v} to have unit length.

Notice that \mathbf{v} occurs twice, to get a vector to have unit length,

We need to divide \mathbf{v} by $\|\mathbf{v}\|$

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

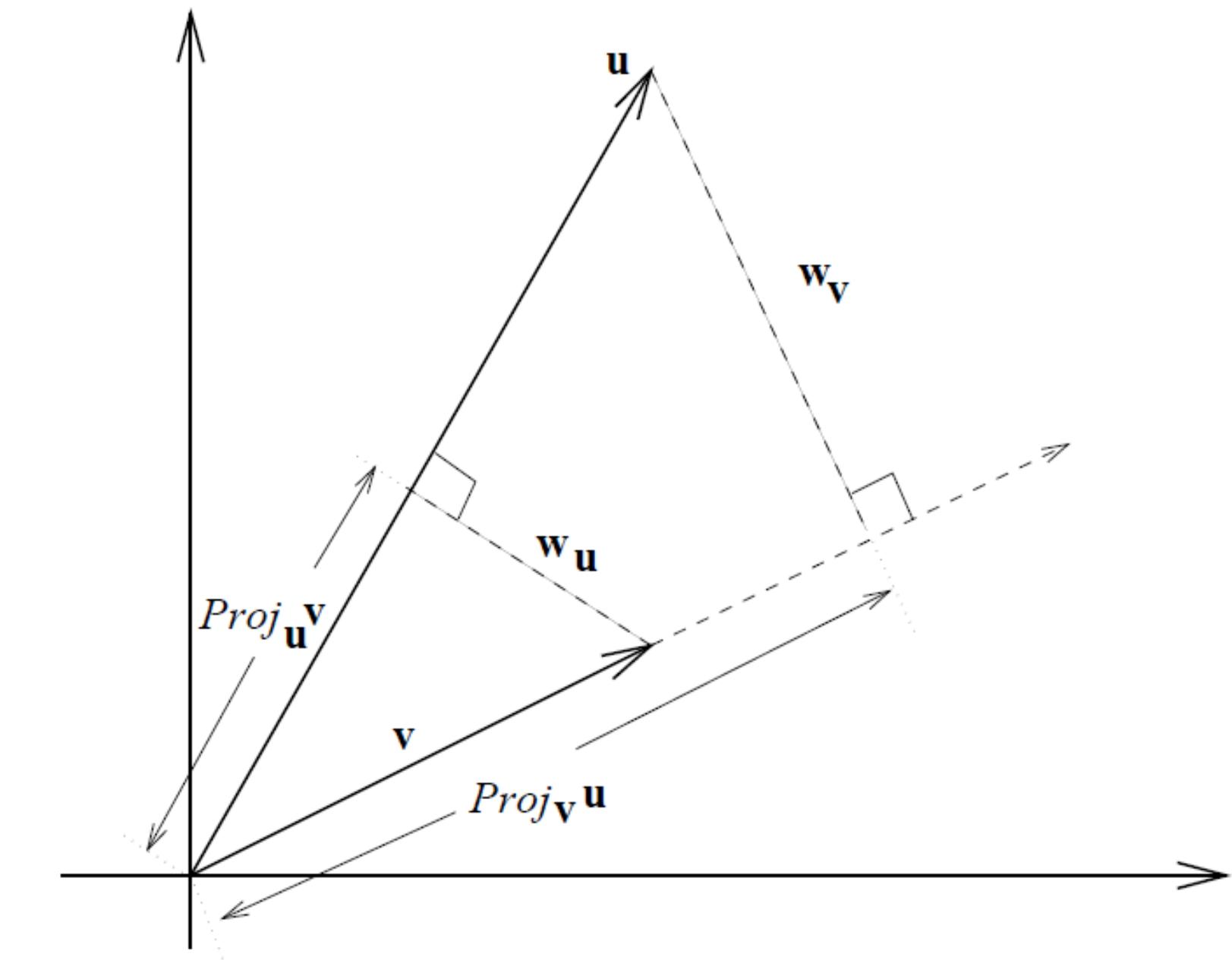
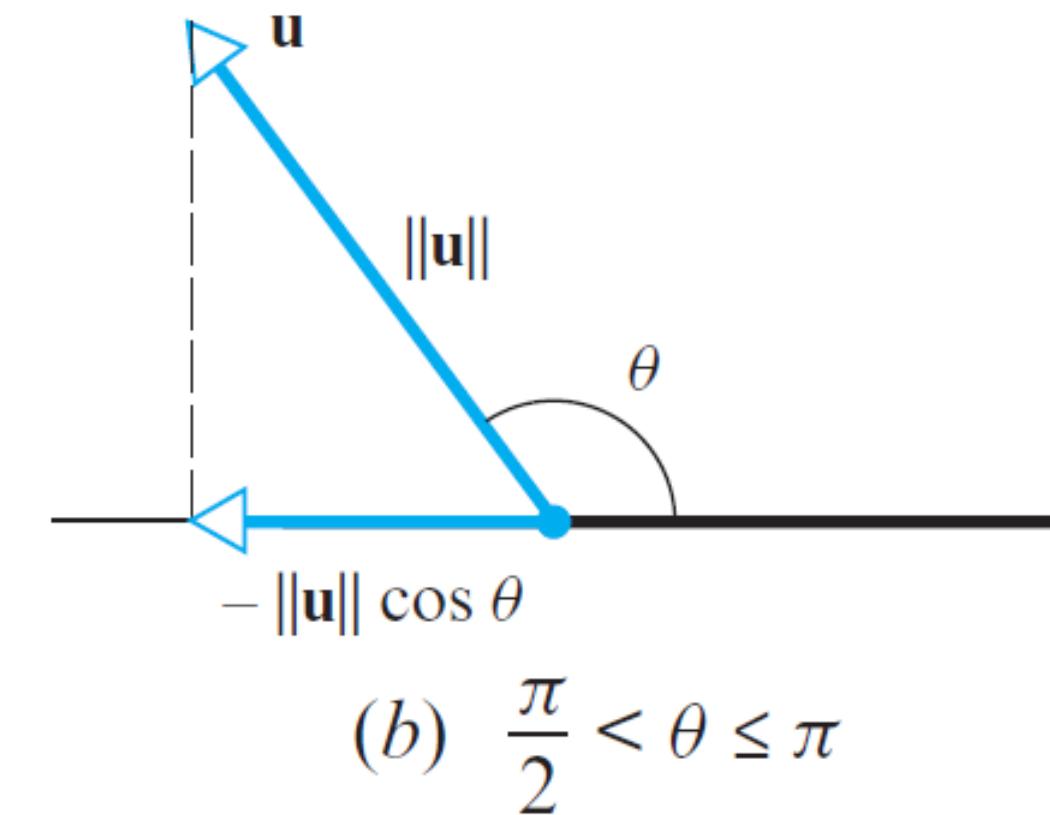
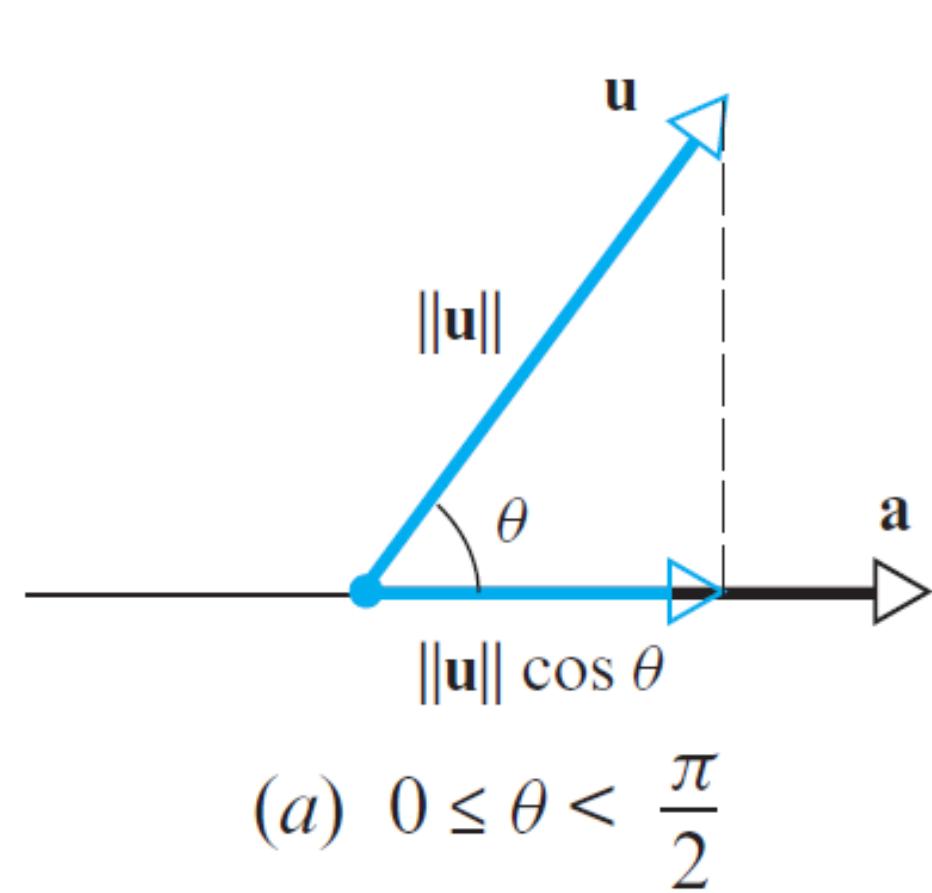


Figure 2.5: Orthogonal projections

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,

$$\|\text{proj}_{\mathbf{a}} \mathbf{u}\| = \frac{|\mathbf{u} \cdot \mathbf{a}|}{\|\mathbf{a}\|} \quad (12)$$

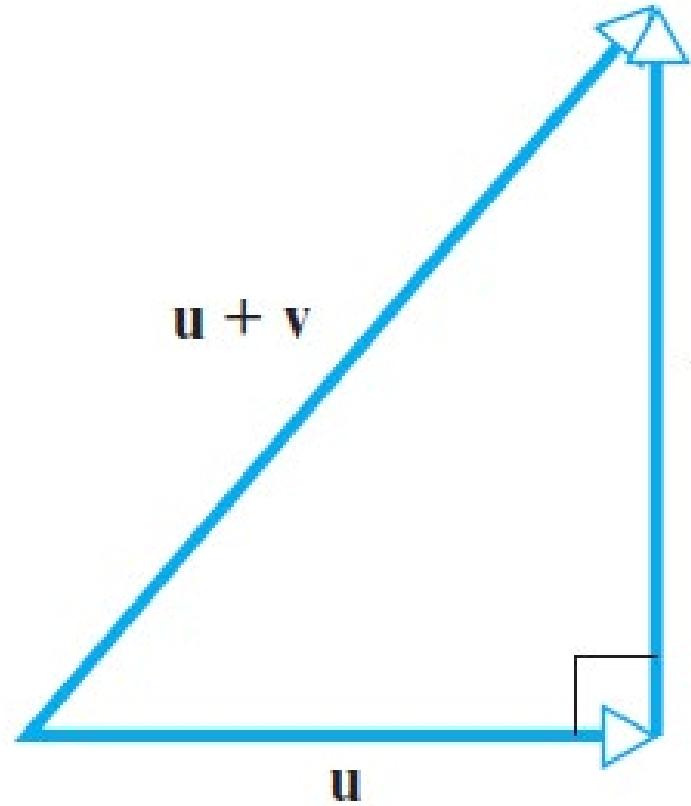
If θ denotes the angle between \mathbf{u} and \mathbf{a} , then $\mathbf{u} \cdot \mathbf{a} = \|\mathbf{u}\| \|\mathbf{a}\| \cos \theta$, so (12) can also be written as

$$\|\text{proj}_{\mathbf{a}} \mathbf{u}\| = \|\mathbf{u}\| |\cos \theta| \quad (13)$$

(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 \mathbf{u} and \mathbf{v} are orthogonal vectors in R^n with the Euclidean inner product, then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad (14)$$

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

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

► EXAMPLE 6 Theorem of Pythagoras in R^4

We showed in Example 1 that the vectors

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

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

Solution We leave it for you to confirm that

$$\mathbf{u} + \mathbf{v} = (-1, 5, 1, 3)$$

$$\|\mathbf{u} + \mathbf{v}\|^2 = 36$$

$$\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = 30 + 6$$

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

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

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

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 $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

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

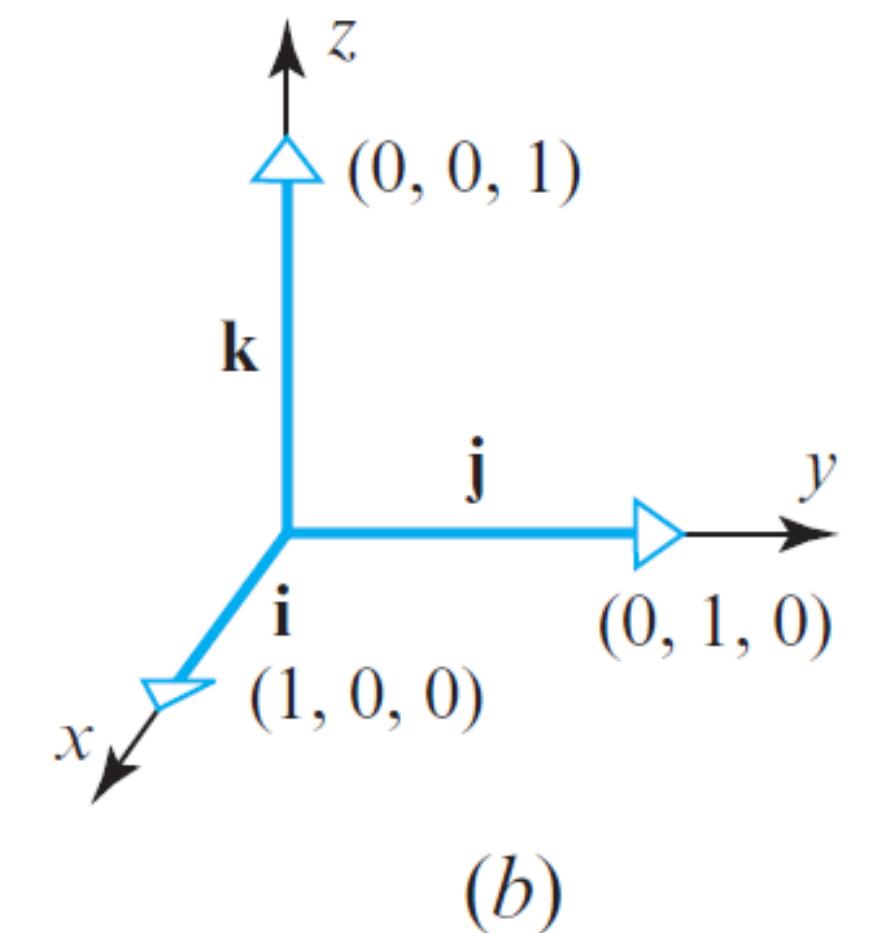
$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$$

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. ■

Standard basis for \mathbb{R}^n is an orthogonal set

Standard Basis

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



▲ Figure 3.2.2

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

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

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

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. ■

Examples of orthogonal set of vectors in \mathbb{R}^3

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

$$\begin{aligned} \mathbf{y} &= -8/6; \\ \mathbf{s} &= [\boxed{1} \boxed{0}; \\ &\quad \boxed{2} \boxed{4}; \\ &\quad \boxed{6} \boxed{y}]; \end{aligned}$$

$$\begin{aligned} \mathbf{s} & \text{checkOrthognality } \mathbf{u}_1 \mathbf{u}_2 = \mathbf{s}' * \mathbf{s} \\ & \text{checkOrthognality } = \begin{bmatrix} 41.0000 & 0.0000 \\ 0.0000 & 17.7778 \end{bmatrix} \end{aligned}$$

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

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

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

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. ■

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

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

are given by

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

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

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

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 . ■

Example

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

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

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

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 $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3(-\frac{1}{2}) + 1(-2) + 1(\frac{7}{2}) = 0$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1(-\frac{1}{2}) + 2(-2) + 1(\frac{7}{2}) = 0$$

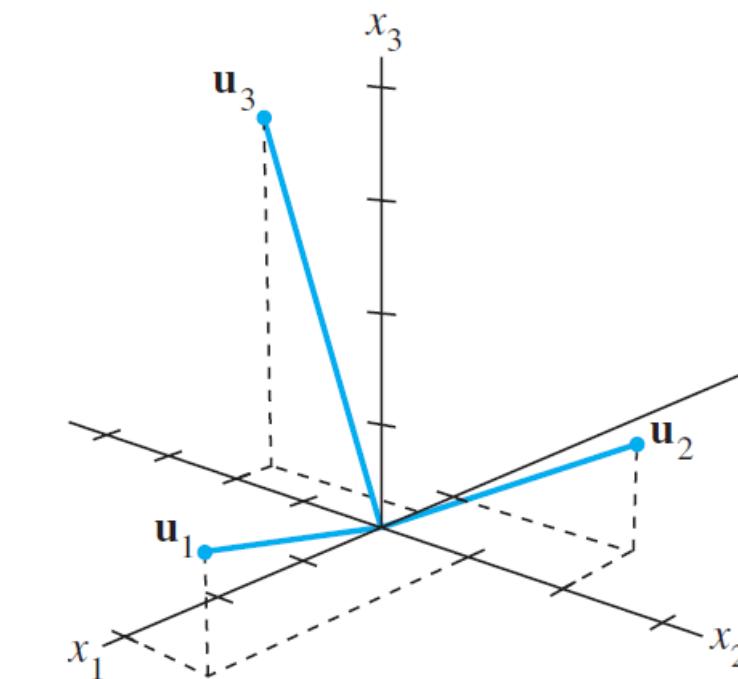


FIGURE 1

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

$$\mathbf{y} \cdot \mathbf{u}_1 = 11, \quad \mathbf{y} \cdot \mathbf{u}_2 = -12, \quad \mathbf{y} \cdot \mathbf{u}_3 = -33$$

$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 11, \quad \mathbf{u}_2 \cdot \mathbf{u}_2 = 6, \quad \mathbf{u}_3 \cdot \mathbf{u}_3 = 33/2$$

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.

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

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

Orthonormal Sets-vectors within have length 1!

Orthonormal Sets

A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ 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 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for \mathbb{R}^n . Any nonempty subset of $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is orthonormal, too. Here is a more complicated example.

THEOREM 4

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 .

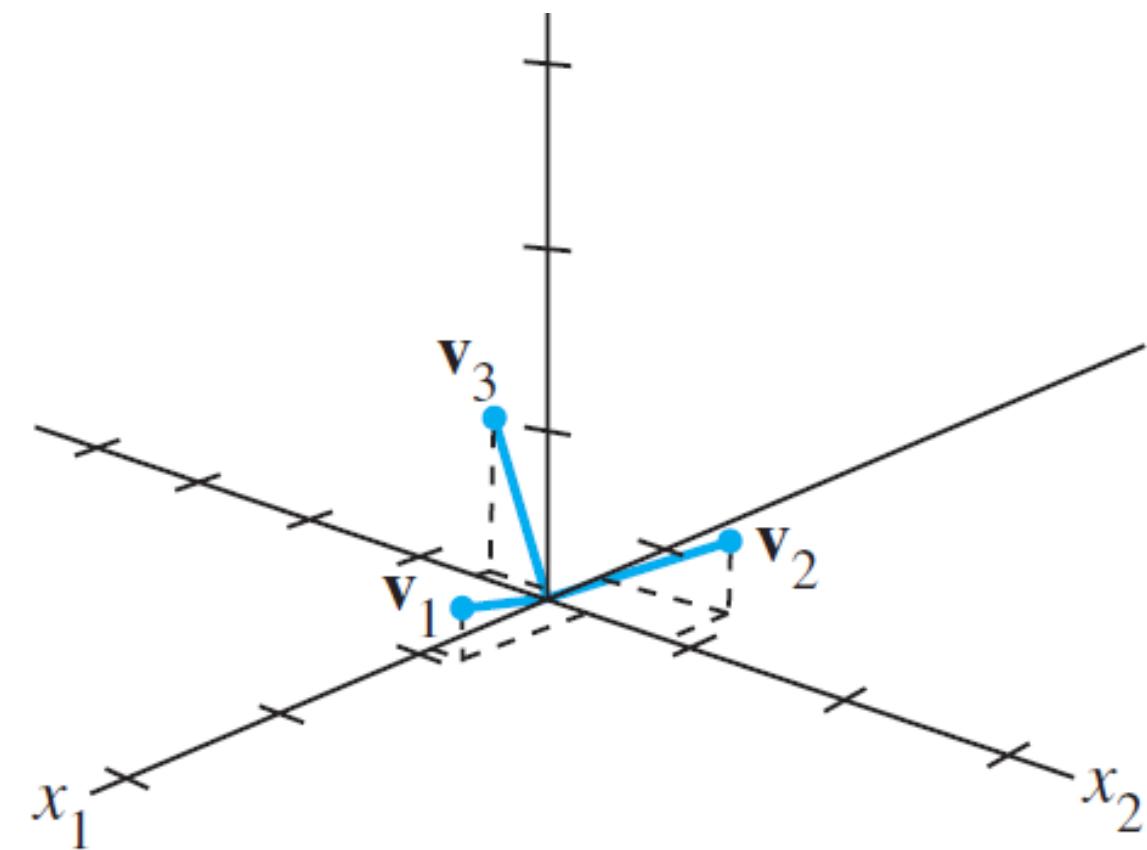


FIGURE 6

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

Lay's Linear Algebra and Applications

342 CHAPTER 6 Orthogonality and Least Squares

Linear Algebra: Orthonormal Basis
61,234 views • Jun 28, 2014

Worldwide Center of Mathematics
26.5K subscribers

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

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

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

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$

$$\mathbf{v}_2 \cdot \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$

$$\mathbf{v}_3 \cdot \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

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  
checkOrthogtnlty = s3'*s3
```

```
checkOrthogtnlty =  
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 $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3]$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T \mathbf{u}_1 & \mathbf{u}_1^T \mathbf{u}_2 & \mathbf{u}_1^T \mathbf{u}_3 \\ \mathbf{u}_2^T \mathbf{u}_1 & \mathbf{u}_2^T \mathbf{u}_2 & \mathbf{u}_2^T \mathbf{u}_3 \\ \mathbf{u}_3^T \mathbf{u}_1 & \mathbf{u}_3^T \mathbf{u}_2 & \mathbf{u}_3^T \mathbf{u}_3 \end{bmatrix} \quad (4)$$

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

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

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

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

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

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

- $\|U\mathbf{x}\| = \|\mathbf{x}\|$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$
- $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ 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.

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

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

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.

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

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

Orthogonal Decomposition

THEOREM 8

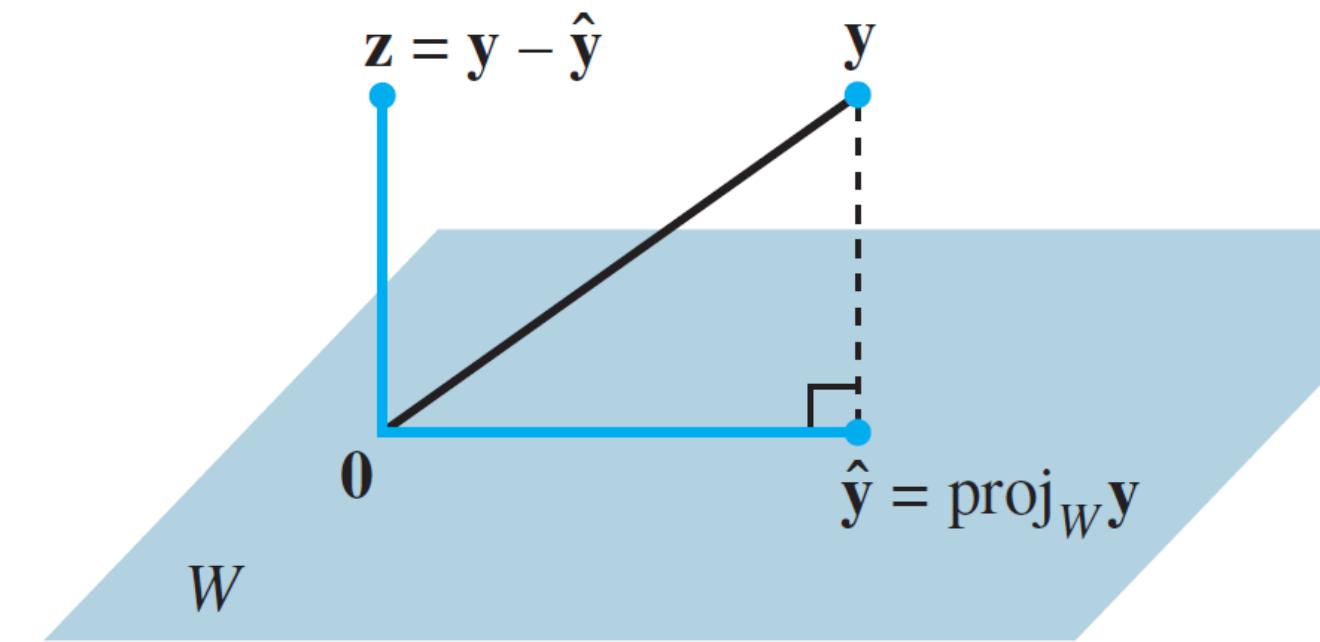


FIGURE 2 The orthogonal projection of \mathbf{y} onto W .

The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of \mathbf{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

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \quad (1)$$

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

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \mathbf{u}_p \quad (2)$$

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

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

¹We 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

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

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

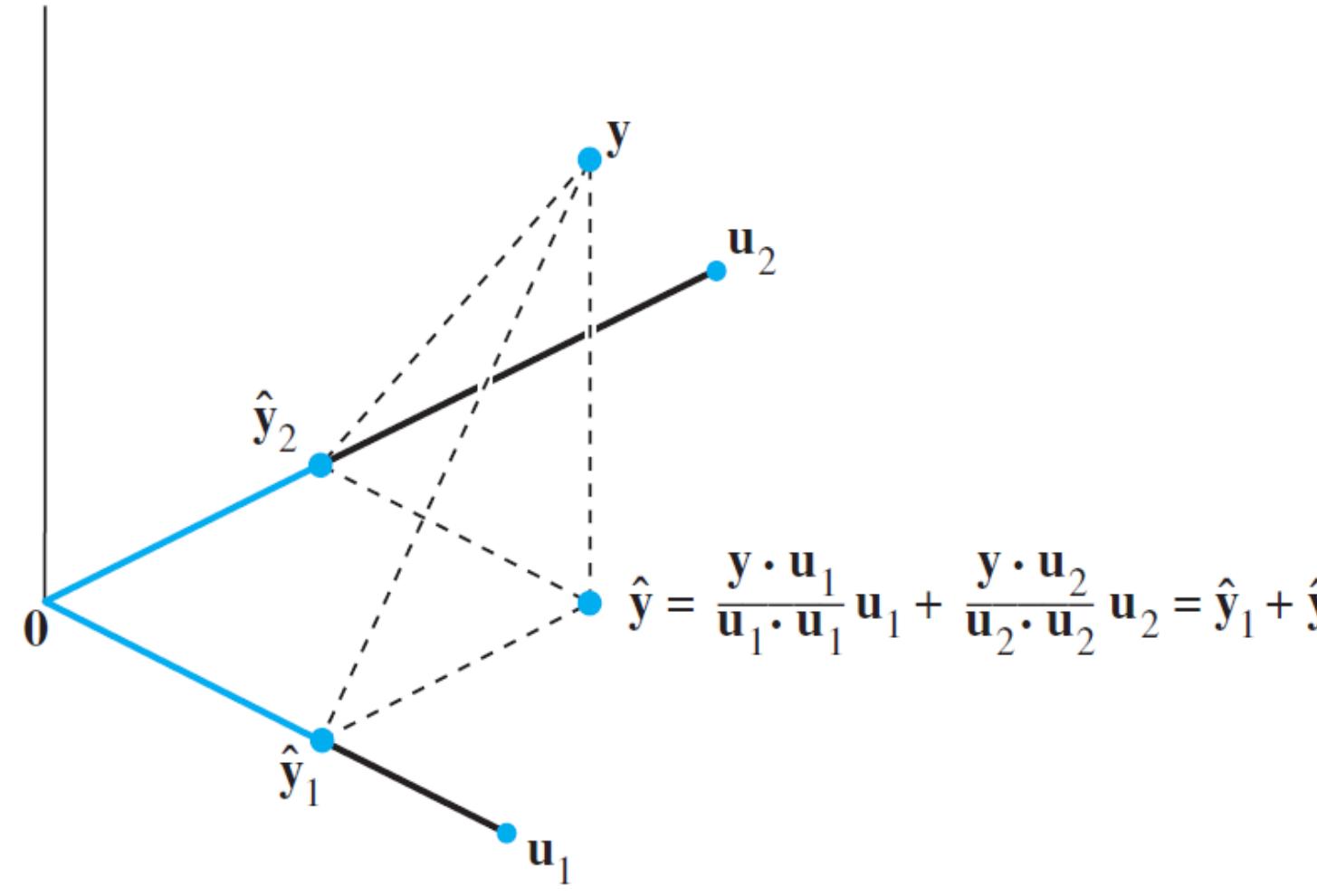


FIGURE 3 The orthogonal projection of \mathbf{y} is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

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.

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

$$\begin{aligned}\hat{\mathbf{y}} &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \\ &= \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}\end{aligned}$$

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

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \quad (3)$$

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 $\text{proj}_W \mathbf{y} = \mathbf{y}$.

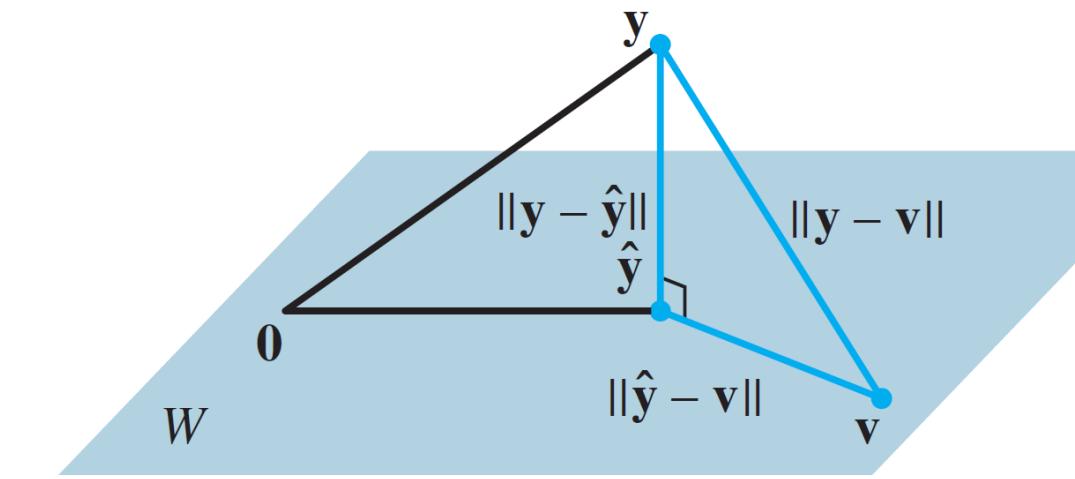


FIGURE 4 The orthogonal projection of \mathbf{y} onto W is the closest point in W to \mathbf{y} .

The vector $\hat{\mathbf{y}}$ in Theorem 9 is called **the best approximation to \mathbf{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

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

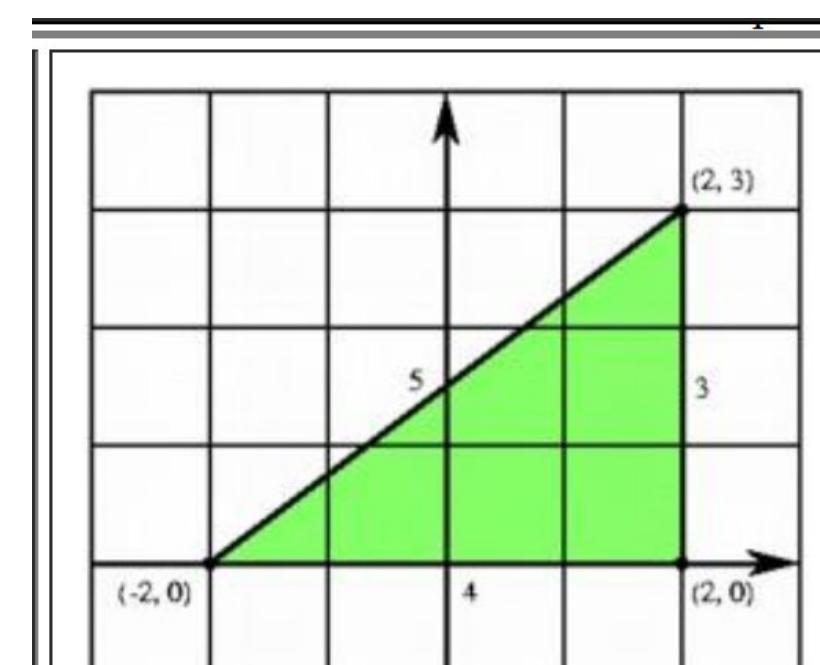
(See the colored right triangle in Fig. 4. The length of each side is labeled.) Now $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ because $\hat{\mathbf{y}} - \mathbf{v} \neq \mathbf{0}$, and so inequality (3) follows immediately. ■

Conclusion: if \mathbf{y} is in the space of W , then we can show
^
 \mathbf{y} to be equal to $\hat{\mathbf{y}}$ 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>



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

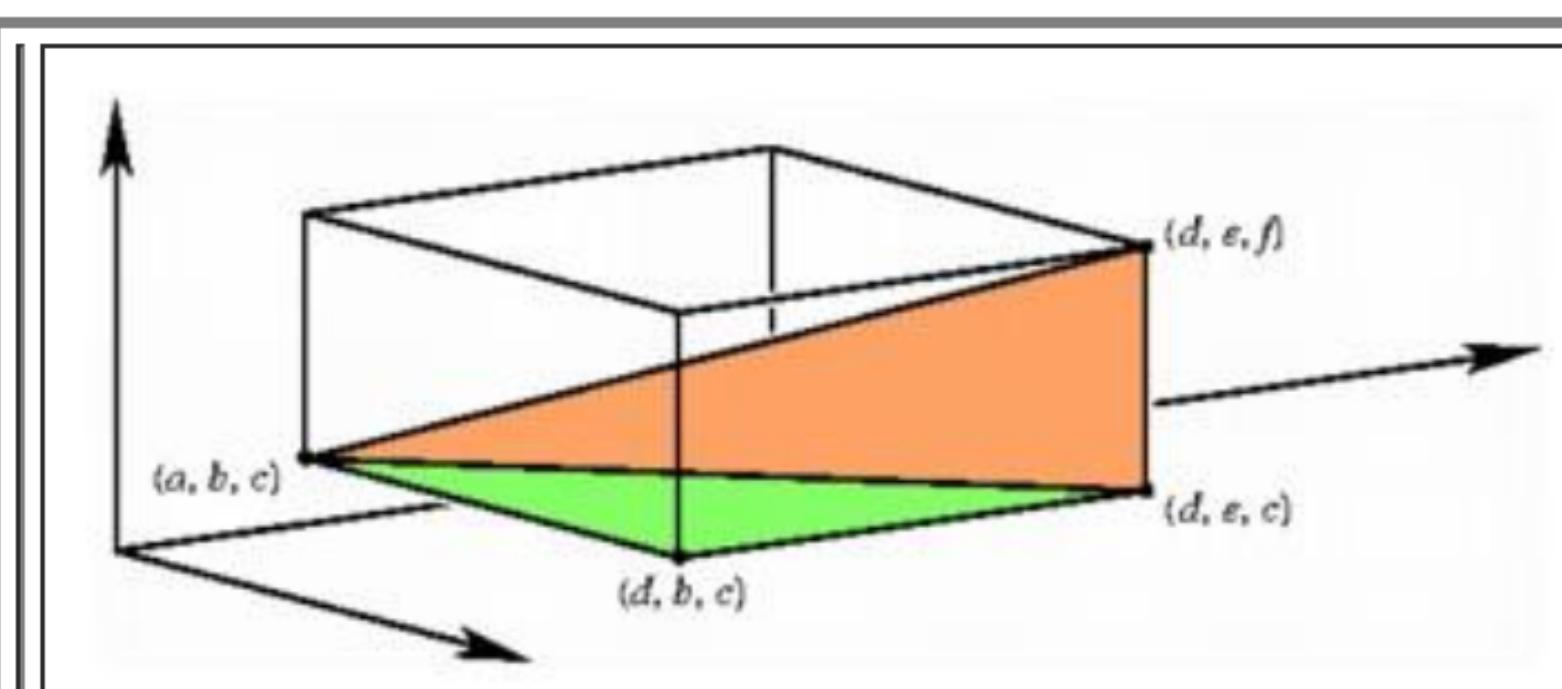
1

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



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.

The Overflow Blog

Coming together as a community connect

Defending yourself against corona scams

Featured on Meta

The Q1 2020 Community Roadmap 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 $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of \mathbb{R}^n , then

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

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

$$\text{Note : } p < n \quad \text{proj}_W \mathbf{y} = UU^T \mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \quad (5)$$

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). ■

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

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

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

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 \mathbf{u}_1 , and \mathbf{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 x example5_orthogonalSet.m x example6_Lay_pg352.m +  
1 % Lay Pg 352, orthogonality, projecting y onto U  
2 U = [- 7 -1; 1 1; 4 -2];  
3 y = [-9 1 6]';  
4  
5 U % sanity check  
6 y % to see the values  
7  
8 orthogonalityCheck = U'*U % quick way to find transpose column of U  
9 % and dot producing its other column  
10  
11 u1 = U(:,1) % extracting u1 and u2 from the column of U  
12 u2 = U(:,2)  
13  
14 u1'*u2 % checking u1 is orthogonal to u2 using dot product  
15  
16 Proj_y_onto_u1 = y'*u1/(u1'*u1) % remember to normalize using denominator  
17 Proj_y_onto_u2 = y'*u2/(u2'*u2) % to make it unit vector  
18  
19 est_y = Proj_y_onto_u1*u1 + Proj_y_onto_u2*u2  
20 est_y = y  
21
```

```
21 % Using projection matrix to find est_y2  
22 U_norm(:,1) = u1/sqrt((u1'*u1))  
23 U_norm(:,2) = u2/sqrt((u2'*u2))  
24  
25 est_y = U_norm*U_norm'*y  
26  
27  
28  
29 y2 = [-8.5, 1, 6]'  
30 est_y2 = U_norm*U_norm'*y2  
31 error = y2 - est_y2  
32
```

Screen shot showing the answer:

```
U =  
-7 -1  
1 1  
4 -2  
Proj_y_onto_u1 = 1.3333  
U_norm =  
-0.8616 -0.4082  
0.1231 0.4082  
0.4924 -0.8165  
y2 =  
-8.5000  
1.0000  
6.0000  
est_y2 =  
-8.5455  
0.8636  
5.9545  
error =  
0.0455  
0.1364  
0.0455  
est_y =  
-9.0000  
1.0000  
6.0000  
orthogonalityCheck =  
66 0  
0 6  
ans =  
1.0e-14 *  
0.1776  
0|  
0
```

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.3.1 (mod Tay Kian Boon)**

Lecture : **Orthogonality**

Topic : **Gram–Schmidt Process**

Concept : **Motivation and Review of Concepts**

Instructor: A/P Chng Eng Siong

TAs: Zhang Su, Vishal Choudhari

Motivation

Consider the problem of solving:

$$Ax = b$$

$A = \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix}$

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$b = \begin{bmatrix} 0.4 \\ 2.5 \\ 1.7 \end{bmatrix}$

M = 3 Rows
N = 2 Columns

Interpretation:

A 's Rows $\Rightarrow M$ Examples or Equations

A 's Columns $\Rightarrow N$ Features or Unknowns

$x \Rightarrow$ Model or Weights

Machine Learning Perspective

$b \Rightarrow$ Target Values

Goal: Solve x or find appropriate model

Based on M & N, there exist three cases:

$$\left[\begin{array}{c|c|c|c} \text{Blue} & \text{Red} & \dots & \text{Green} \\ \hline & & & \text{Yellow} \end{array} \right]$$

$$M \gg N$$

More equations,
less unknowns.

Hence, over-determined!

$$\left[\begin{array}{c|c|c|c} \text{Blue} & \text{Red} & \dots & \text{Green} \\ \hline & & & \text{Yellow} \end{array} \right]$$

$$M = N$$

$$\left[\begin{array}{c|c|c|c} \text{Blue} & \text{Red} & \dots & \text{Green} \\ \hline & & & \text{Yellow} \end{array} \right]$$

$$M \ll N$$

Less equations,
more unknowns.

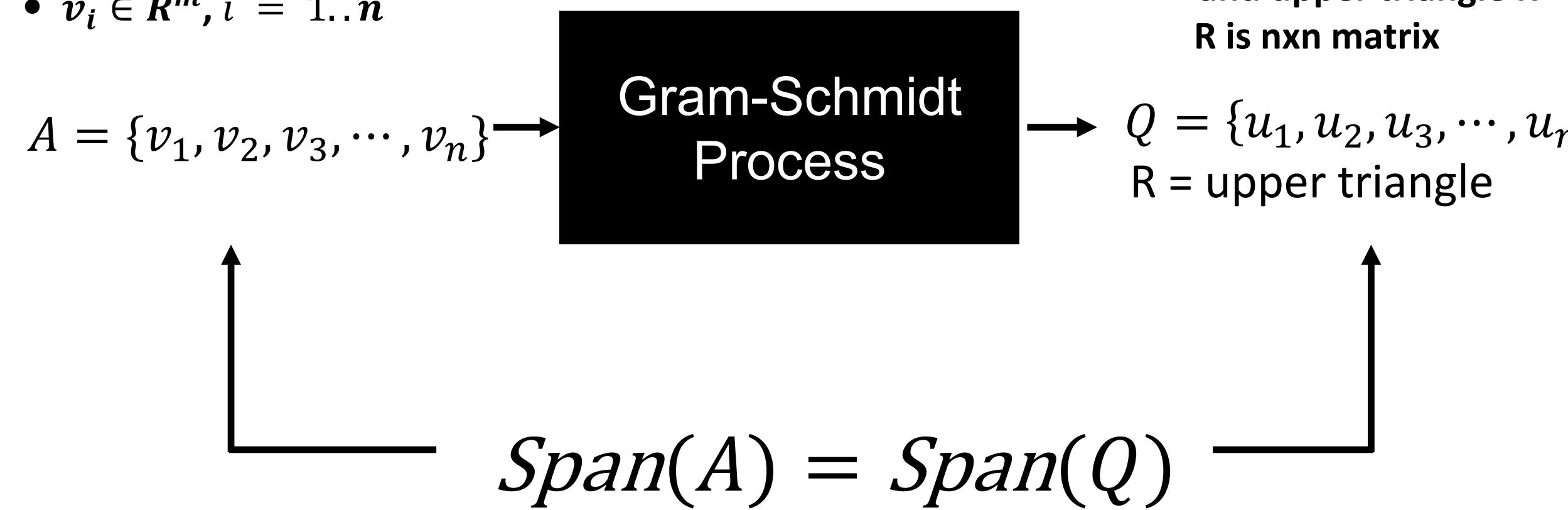
Hence, under-determined!

Motivation

What does Gram-Schmidt Process Do?

It orthogonalises a set of vectors!

- A is a mxn matrix,
assume columns are independent
- but are Not orthogonal
- $v_i \in R^m, i = 1..n$



Note: Q spans the same m-dimensional subspace of R^m as that of A

Ref: <http://www.seas.ucla.edu/~vandenbe/133A/lectures/qr.pdf> Slide 6.7

See Matlab: <https://www.mathworks.com/help/matlab/ref/qr.html>

(economy QR factorization vs full QR decomposition)

More explanations of full vs economy qr : http://www.ece.northwestern.edu/local-apps/matlabhelp/techdoc/math_anal/mat_li23.html

Applications for Gram-Schmidt Process

1. QR Factorisation:

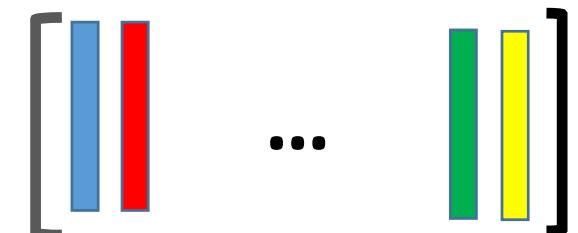
Ref: <https://www.quora.com/Why-is-QR-factorization-useful-and-important>

For large M and N , the system of equations of the form $Ax = b$ can be solved efficiently if A can be rewritten as:

$$A = QR$$

The columns of Q are orthonormal, and R is an upper triangular matrix.

2. Feature Selection:



When $M \ll N$, i.e, when there are more features than examples, features/columns most representative of b can be identified. The Gram Schmidt must be modified to take care of this (not in scope of work here)

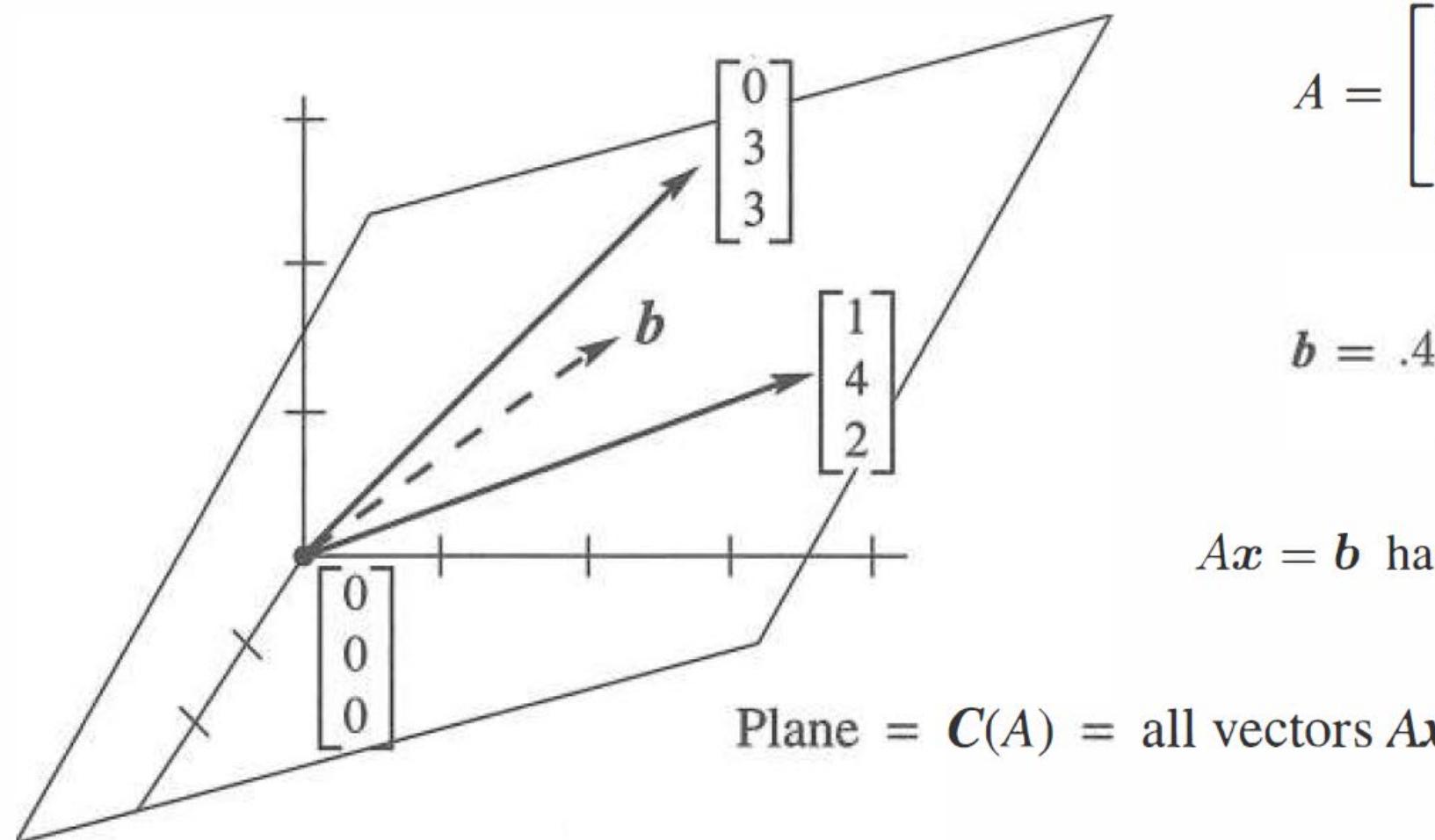
Ref: <https://qr.ae/pNKiO5>

$M \ll N$

Reviewing finding the solution of x for $Ax=b$

Matrix A is of dimension $M \times N$.

M Rows, N Columns



Chapter 3. Vector Spaces and Subspaces

$$A = \begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix}$$

$$b = .4 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + .3 \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}$$

$$Ax = b \text{ has } x = \begin{bmatrix} .4 \\ .3 \end{bmatrix}$$

Figure 3.2: The column space $C(A)$ is a plane containing the two columns. $Ax = b$ is solvable when b is on that plane. Then b is a combination of the columns.

In the example above, Matrix A is of dimension 3×2

3 Rows, 2 Columns

If $Ax = b$, then $x \in R^2$, and $b \in R^3$

By b being in a linear combination of columns
of A , b lies in the span of column vectors of A .

If A can be decomposed into

$$A = QR$$

where,

Q is an $m \times n$ matrix with orthonormal columns (having same col. space of A)

R is an upper triangular matrix.

The problem to find x can then be easily solved by:

$$Ax = b$$

$$QRx = b$$

$$Q^T QRx = Q^T b$$

$$Rx = Q^T b$$

Since R is upper triangle, x can be quickly found by back-substitution.

Note: if b is not in $C(A)$, then the found x will only result in the orthogonal projection of b onto $C(A)$.

Ref: https://www.mathwords.com/b/back_substitution.htm

Ref: Strang, Introduction to Linear Algebra

Note: solving for x is not trivial when
If A does not have orthogonal columns.

Reviewing Span

Linear Combinations

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ in \mathbb{R}^n and given scalars c_1, c_2, \dots, c_p , the vector \mathbf{y} defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p$$

is called a **linear combination** of $\mathbf{v}_1, \dots, \mathbf{v}_p$ with **weights** c_1, \dots, c_p .

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of vectors.

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in \mathbb{R}^n , then the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted by $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the **subset of \mathbb{R}^n spanned (or generated) by $\mathbf{v}_1, \dots, \mathbf{v}_p$** . That is, $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_p\mathbf{v}_p$$

with c_1, \dots, c_p scalars.

Note that $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ contains every scalar multiple of \mathbf{v}_1 (for example), since $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_p$. In particular, the zero vector must be in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

A Geometric Description of $\text{Span}\{\mathbf{v}\}$ and $\text{Span}\{\mathbf{u}, \mathbf{v}\}$

Let \mathbf{v} be a nonzero vector in \mathbb{R}^3 . Then $\text{Span}\{\mathbf{v}\}$ is the set of all scalar multiples of \mathbf{v} , which is the set of points on the line in \mathbb{R}^3 through \mathbf{v} and $\mathbf{0}$. See Fig. 10.

If \mathbf{u} and \mathbf{v} are nonzero vectors in \mathbb{R}^3 , with \mathbf{v} not a multiple of \mathbf{u} , then $\text{Span}\{\mathbf{u}, \mathbf{v}\}$ is the plane in \mathbb{R}^3 that contains \mathbf{u} , \mathbf{v} , and $\mathbf{0}$. In particular, $\text{Span}\{\mathbf{u}, \mathbf{v}\}$ contains the line in \mathbb{R}^3 through \mathbf{u} and $\mathbf{0}$ and the line through \mathbf{v} and $\mathbf{0}$. See Fig. 11.

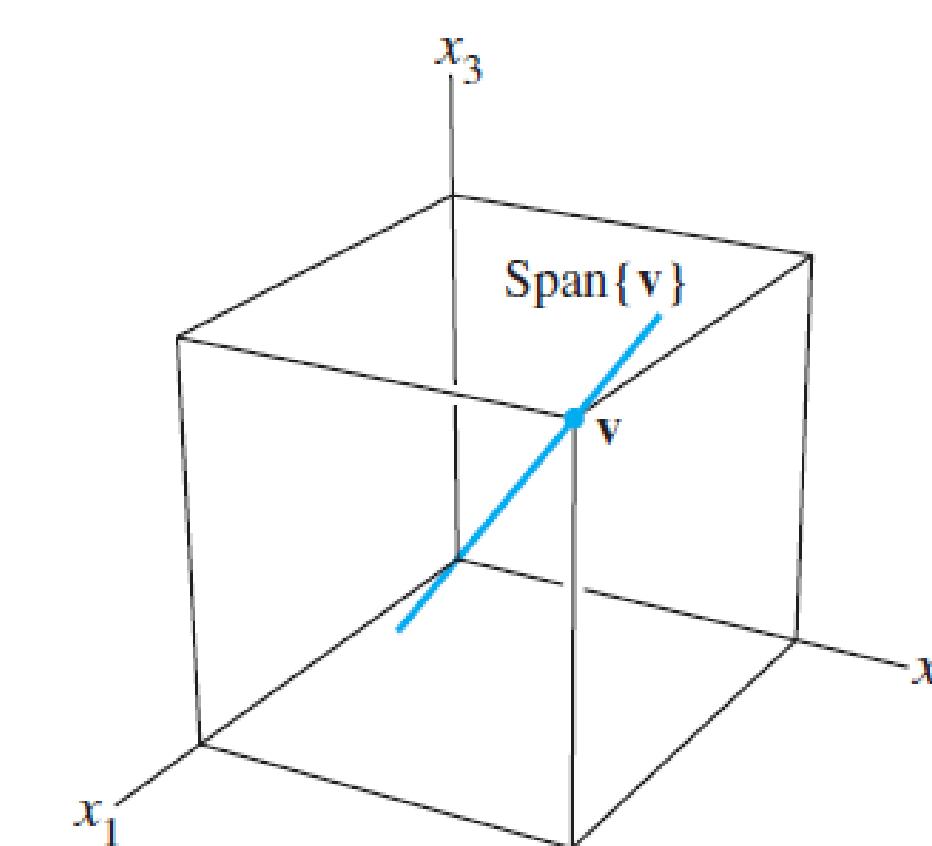


FIGURE 10 $\text{Span}\{\mathbf{v}\}$ as a line through the origin.

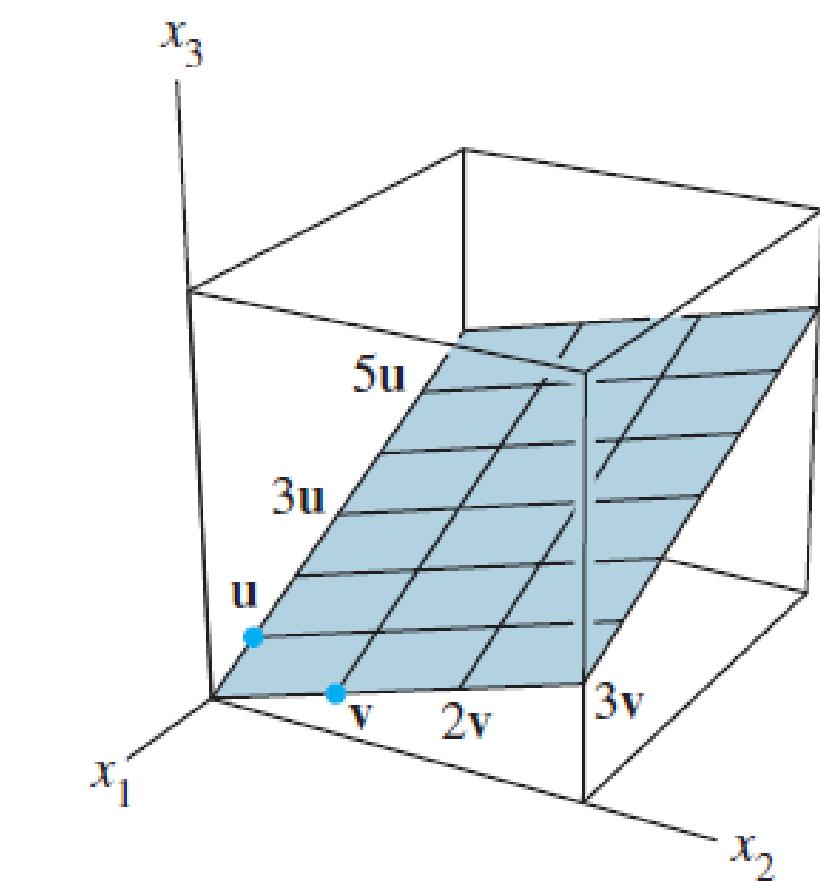


FIGURE 11 $\text{Span}\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

Reviewing Span

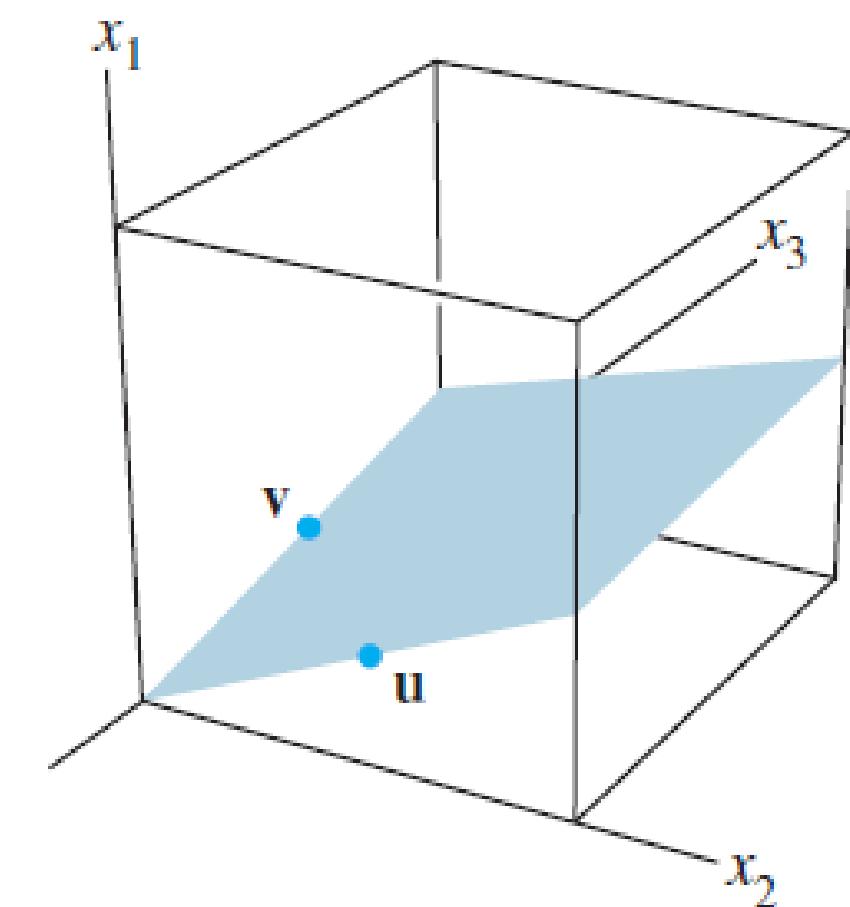
EXAMPLE 2 A single linear equation can be treated as a very simple system of equations. Describe all solutions of the homogeneous “system”

$$10x_1 - 3x_2 - 2x_3 = 0 \quad (1)$$

SOLUTION There is no need for matrix notation. Solve for the basic variable x_1 in terms of the free variables. The general solution is $x_1 = .3x_2 + .2x_3$, with x_2 and x_3 free. As a vector, the general solution is

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} .2x_3 \\ 0 \\ x_3 \end{bmatrix} \\ &= x_2 \begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix} \quad (\text{with } x_2, x_3 \text{ free}) \end{aligned} \quad (2)$$

This calculation shows that every solution of (1) is a linear combination of the vectors \mathbf{u} and \mathbf{v} , shown in (2). That is, the solution set is $\text{Span}\{\mathbf{u}, \mathbf{v}\}$. Since neither \mathbf{u} nor \mathbf{v} is a scalar multiple of the other, the solution set is a plane through the origin. See Fig. 2. ■



CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : **6.3.2 (mod Tay KB)**

Lecture : **Orthogonality**

Topic : **Gram–Schmidt Process**

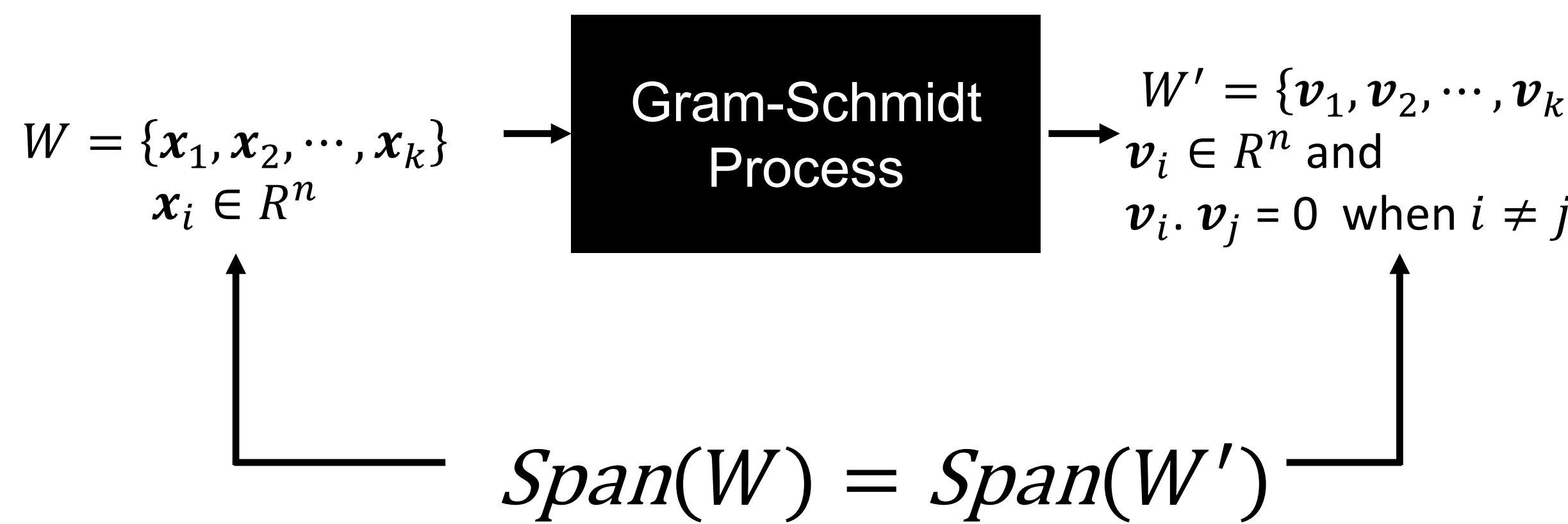
Concept : **The Gram-Schmidt Process**

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

The Gram Schmidt Process

What does Gram-Schmidt Process Do?

It orthogonalises a set of vectors!



The Gram–Schmidt process is a simple algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of R^n . The first two examples of the process are aimed at hand calculation.

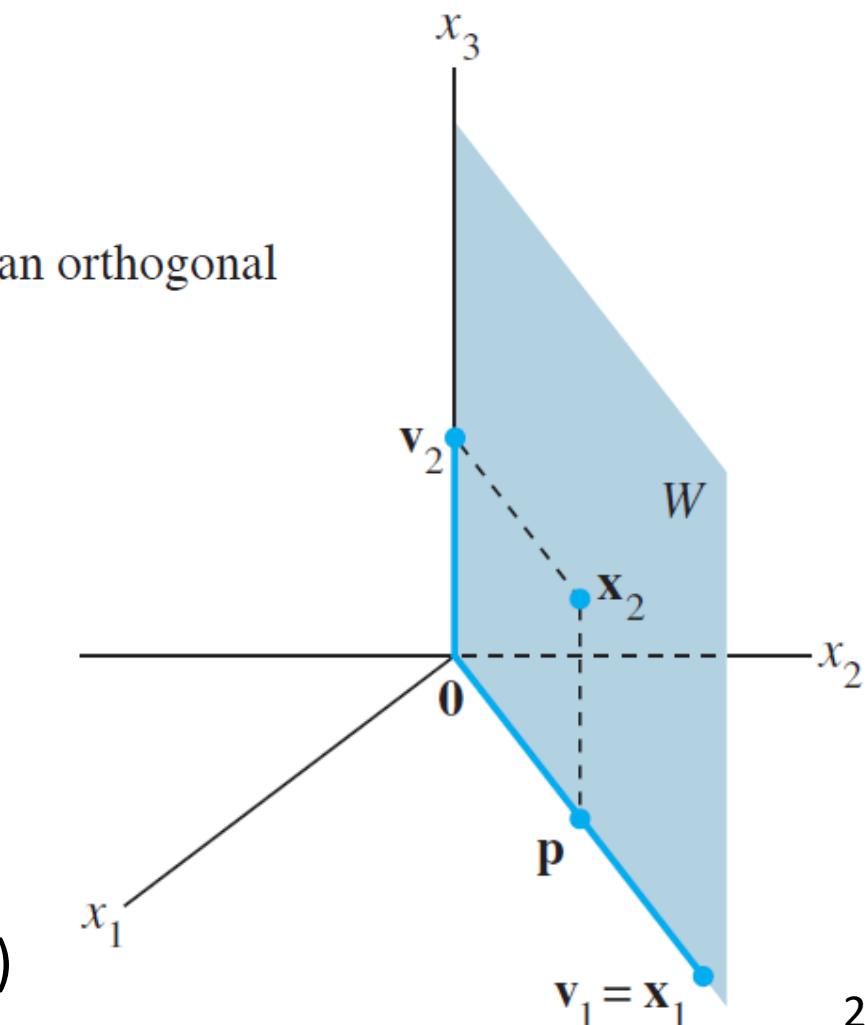
EXAMPLE 1 Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for W .

SOLUTION The subspace W is shown in Fig. 1, along with \mathbf{x}_1 , \mathbf{x}_2 , and the projection \mathbf{p} of \mathbf{x}_2 onto \mathbf{x}_1 . The component of \mathbf{x}_2 orthogonal to \mathbf{x}_1 is $\mathbf{x}_2 - \mathbf{p}$, which is in W because it is formed from \mathbf{x}_2 and a multiple of \mathbf{x}_1 . Let $\mathbf{v}_1 = \mathbf{x}_1$ and

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1} \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal set of nonzero vectors in W . Since $\dim W = 2$, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W . ■

FIGURE 1
Construction of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.



The Gram Schmidt Process

The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ for a nonzero subspace W of \mathbb{R}^n , define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2$$

⋮

$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}} \mathbf{v}_{p-1}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthogonal basis for W . In addition

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \quad \text{for } 1 \leq k \leq p \quad (1)$$

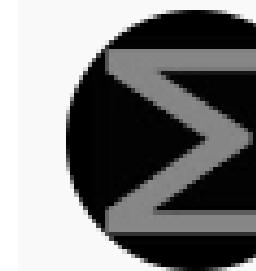


MIT OpenCourseWare
2.46M subscribers

Gram-Schmidt Orthogonalization

Instructor: Ana Rita Pires

View the complete course: <http://ocw.mit.edu/18-06SCF11>



Worldwide Center of Mathematics
26.5K subscribers

Watch these worked out examples:

1. GramSchmidt: <https://www.youtube.com/watch?v=Aslf3KGq2UE>
2. QR: <https://www.youtube.com/watch?v=6DybLNNKWyE>
3. MIT Gram Schmidt: <https://www.youtube.com/watch?v=TRktLuAktBQ&t=17s>

The Gram Schmidt Process

EXAMPLE 2 Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of \mathbb{R}^4 . Construct an orthogonal basis for W .

SOLUTION

Step 1. Let $\mathbf{v}_1 = \mathbf{x}_1$ and $W_1 = \text{Span}\{\mathbf{x}_1\} = \text{Span}\{\mathbf{v}_1\}$.

Step 2. Let \mathbf{v}_2 be the vector produced by subtracting from \mathbf{x}_2 its projection onto the subspace W_1 . That is, let

$$\begin{aligned}\mathbf{v}_2 &= \mathbf{x}_2 - \text{proj}_{W_1} \mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \quad \text{Since } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}\end{aligned}$$

As in Example 1, \mathbf{v}_2 is the component of \mathbf{x}_2 orthogonal to \mathbf{x}_1 , and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for the subspace W_2 spanned by \mathbf{x}_1 and \mathbf{x}_2 .

Note: $\mathbf{v}'_2 = \mathbf{v}_2 * 4$ to get rid of denominator in \mathbf{v}_2

Step 3. Let \mathbf{v}_3 be the vector produced by subtracting from \mathbf{x}_3 its projection onto the subspace W_2 . Use the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}'_2\}$ to compute this projection onto W_2 :

$$\begin{array}{c} \text{Projection of} \\ \mathbf{x}_3 \text{ onto } \mathbf{v}_1 \\ \downarrow \\ \text{proj}_{W_2} \mathbf{x}_3 = \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}'_2}{\mathbf{v}'_2 \cdot \mathbf{v}'_2} \mathbf{v}'_2 = \frac{2}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} \end{array}$$

Then \mathbf{v}_3 is the component of \mathbf{x}_3 orthogonal to W_2 , namely,

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

See Slide 3 of Chapter 6.2.5
for explanation.

The Gram Schmidt Process

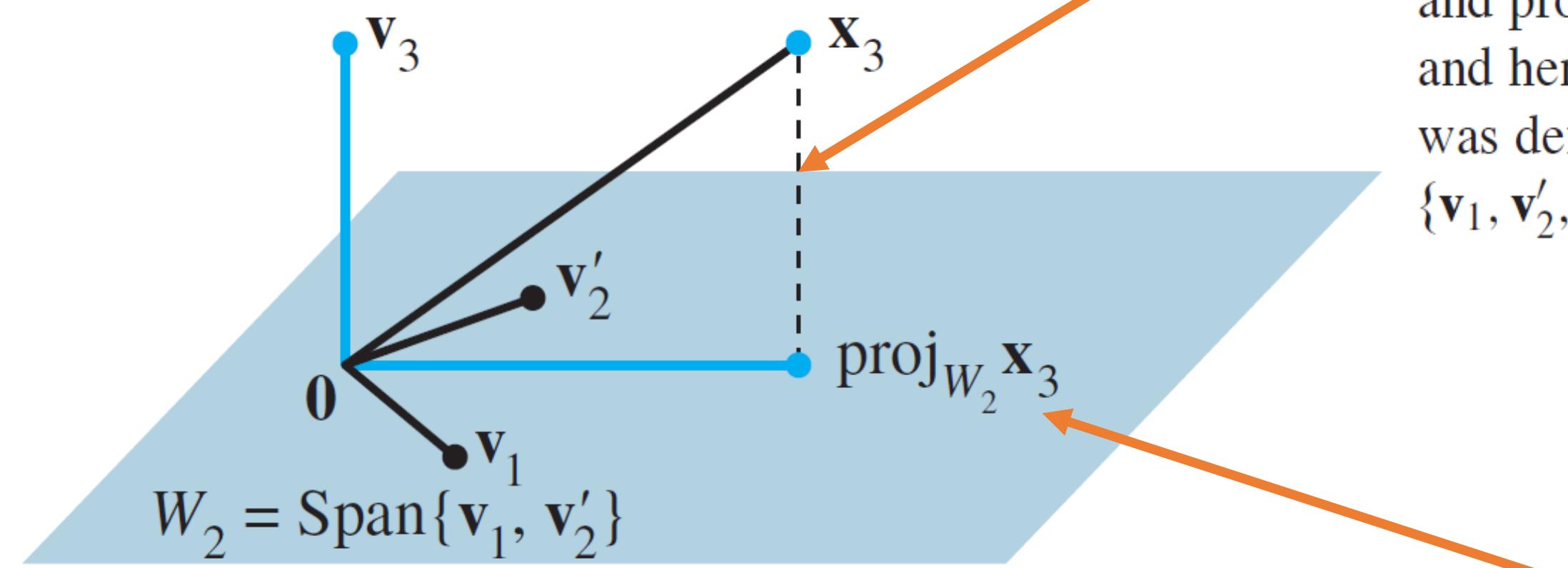


FIGURE 2 The construction of \mathbf{v}_3 from \mathbf{x}_3 and W_2 .

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

See Fig. 2 for a diagram of this construction. Observe that \mathbf{v}_3 is in W , because \mathbf{x}_3 and $\text{proj}_{W_2} \mathbf{x}_3$ are both in W . Thus $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ is an orthogonal set of nonzero vectors and hence a linearly independent set in W . Note that W is three-dimensional since it was defined by a basis of three vectors. Hence, by the Basis Theorem in Section 4.5, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ is an orthogonal basis for W . ■

$$\text{Projection of } \mathbf{x}_3 \text{ onto } \mathbf{v}_1 \quad \downarrow \quad \text{Projection of } \mathbf{x}_3 \text{ onto } \mathbf{v}_2' \quad \downarrow$$
$$\text{proj}_{W_2} \mathbf{x}_3 = \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}_2'}{\mathbf{v}_2' \cdot \mathbf{v}_2'} \mathbf{v}_2' = \frac{2}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

The Basis Theorem

Let V be a p -dimensional vector space, $p \geq 1$. Any linearly independent set of exactly p elements in V is automatically a basis for V . Any set of exactly p elements that spans V is automatically a basis for V .

CX1104: Linear Algebra for Computing

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_b$$

Chap. No : 6.3.3 (mod Tay KB)

Lecture : Orthogonality

Topic : Gram–Schmidt Process

Concept : Using GS Process for QR Factorisation

Instructor: A/P Chng Eng Siong

TAs: Zhang Su, Vishal Choudhari

Gram–Schmidt Process for QR Factorisation

The application of the Gram–Schmidt process to the column vectors of a **full column rank** matrix yields the QR decomposition.

QR decomposition helps in solving linear least squares problem (next chapter).

If $A \in R^{m \times n}$ has full column rank, i.e, A has linearly independent columns, A can be factored as follows:

$$A = QR$$

$$A = [q_1 \ q_2 \ \dots \ q_n] \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ 0 & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{nn} \end{bmatrix}$$

The Q Factor (economy qr):

- Q is $m \times n$ with orthonormal columns ($Q^T Q = I$)
- If A is square ($m = n$), then Q is orthogonal, i.e, $Q^T Q = QQ^T = I$

The R Factor:

- R is $n \times n$ upper triangular, with nonzero diagonal elements
- R is nonsingular (diagonal elements are nonzero)

- Vectors q_1, q_2, \dots, q_n are orthonormal m -dimensional vectors: $\|q_i\| = 1$ and $q_i^T q_j = 0$ if $i \neq j$
- Diagonal elements R_{ii} are nonzero.

NOTE:

Q is obtained by performing GS Process on A^2

Gram–Schmidt Process for QR Factorisation

THEOREM 12

The QR Factorization

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

$$A = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_n] = [Q \mathbf{r}_1 \ \cdots \ Q \mathbf{r}_n] = QR$$

$$\begin{bmatrix} \parallel & \parallel & \parallel \\ & \dots & \\ \parallel & \parallel & \parallel \end{bmatrix} = \begin{bmatrix} \parallel & \parallel & \parallel \\ & \parallel & \parallel \\ & & \parallel \end{bmatrix} \begin{bmatrix} & & \\ & & \\ & \text{blue triangle} & \\ & & 0 \end{bmatrix}$$

$\mathbf{R} = \text{Upper Triangle}$

$\mathbf{A} \in R^{m \times n}$ $\mathbf{Q} = \text{columns are orthonormal}$

```
% Gram Schmit - Chng Eng Siong, for QR decomposition
% 20 May 2020
function [Q,R]= my_GS(A)

[m,n] = size(A);
Q=zeros(m,n);
R=zeros(n,n);

x1 = A(:,1);
Q(:,1) = x1/norm(x1);
R(1,1) = norm(x1);

for i=2:n
    x_i = A(:,i);
    v_i = x_i;
    for j=1:i-1
        v_j = Q(:,j);
        RVal = (x_i'*v_j)/(v_j'*v_j);
        v_i = v_i - RVal*v_j;
        R(j,i) = RVal;
    end % of for j
    Q(:,i) = v_i/norm(v_i);
    R(i,i) = norm(v_i);
end % of for i
end % of funcyion
```

Gram–Schmidt Process for QR Factorisation

Gram–Schmidt Process

Consider a full column rank matrix A :

$$A = [x_1, x_2, x_3, \dots, x_n]$$

$$\text{Proj}_v x = \frac{\langle v, x \rangle}{\langle v, v \rangle} v$$

$$v_1 = x_1$$

$$v_2 = x_2 - \text{Proj}_{v_1} x_2$$

$$v_3 = x_3 - \text{Proj}_{v_1} x_3 - \text{Proj}_{v_2} x_3$$

$$v_i = x_i - \sum_{j=1}^{i-1} \text{Proj}_{v_j} x_i$$

$$e_1 = \frac{v_1}{\|v_1\|}$$

$$e_i = \frac{v_i}{\|v_i\|}$$

We can now express the a_i s over our newly computed orthonormal basis:

$$x_1 = \langle e_1, x_1 \rangle e_1$$

$$x_2 = \langle e_1, x_2 \rangle e_1 + \langle e_2, x_2 \rangle e_2$$

$$x_3 = \langle e_1, x_3 \rangle e_1 + \langle e_2, x_3 \rangle e_2 + \langle e_3, x_3 \rangle e_3$$

⋮

$$x_n = \sum_{j=1}^n \langle e_j, x_n \rangle e_j$$

This can be written in matrix form:

$$A = QR$$

where:

$$Q = [e_1, e_2, e_3, \dots, e_n]$$

and

$$R = \begin{pmatrix} \langle e_1, x_1 \rangle & \langle e_1, x_2 \rangle & \langle e_1, x_3 \rangle & \dots \\ 0 & \langle e_2, x_2 \rangle & \langle e_2, x_3 \rangle & \dots \\ 0 & 0 & \langle e_3, x_3 \rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Gram–Schmidt Process for QR Factorisation

Example [edit]

Consider the decomposition of

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}.$$

Recall that an orthonormal matrix Q has the property

$$Q^T Q = I.$$

Then, we can calculate Q by means of Gram–Schmidt as follows:

$$U = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3) = \begin{pmatrix} 12 & -69 & -58/5 \\ 6 & 158 & 6/5 \\ -4 & 30 & -33 \end{pmatrix};$$

$$Q = \left(\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \quad \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \quad \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \right) = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix}$$

Thus, we have

$$Q^T A = Q^T Q R = R;$$

$$R = Q^T A = \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

Hence,

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix} = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix} \times \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

Q : Orthogonal Matrix

R : Upper Triangular Matrix