In Chapter (3), we defined the dot product of vectors in R^n and we used that concept to define notions of length, angle, distance and orthogonality.

In this Chapter, we will generalize those ideas so they are applicable in any vector space, not just R^n.

SEC (6.1) INNER PRODUCTS

In this Section, we will use the most important properties of the dot product on R^n as axioms, which if satisfied by the vectors in a vector space V, will enable us to extend the notions of length, distance, angle and perpendicularity to general vector spaces.

Definition: An Inner Product on a real vector space V is a function that associates a real number $\langle u, v \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors u, v and w in V and all scalars k

(i) $\langle u, v \rangle = \langle v, u \rangle$ [Symmetry Axiom]

(ii) $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ [Additivity Axiom]

(iii) $\langle ku, v \rangle = k\langle u, v \rangle$ [Homogeneity Axiom]

(iv) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ iff $\mathbf{v} = \mathbf{0}$ [Positivity Axiom]

A real vector space with an inner product is called a Real Product Space.

NOTE: This defi. applies only to real vector spaces. Since we will have little need for complex vector spaces from this point on, you can assume that all vector spaces under discussion are real, even though some of theorems are also valid in complex vector spaces.

Because the axioms for a real inner product space are based on properties of the dot product, these inner product space axioms will be satisfied automatically if we define the inner product of two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n to be | $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ & $\mathbf{v} = (v_1, v_2, \cdots, v_n)$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

This inner product is commonly called the Euclidean Inner Product (or Standard Inner Product) on R^n to distinguish it from other possible inner products that might be defined **on** R^n. We call R^n with Euclidean inner product as Euclidean n-space.

Inner products can be used to define notions of norm and distance in a general inner product space just as we did with dot products in R^n. Recall from Sec (3.2) that if u and v are vectors in Euclidean n-space, then norm and distance can be expressed in terms of the dot product as

$$\|v\| = \sqrt{v \cdot v}$$

and $$d(u, v) = \|u - v\| = \sqrt{(u-v) \cdot (u-v)}$$

Motivated by these formulas, we make the following definition —

Definition: If V is a real inner product space, then the Norm (or Length) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{<\mathbf{v},\mathbf{v}>}$$

and the Distance between two vectors is denoted by $d(\mathbf{u},\mathbf{v})$ and is defined by

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \sqrt{<\mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}>}$$

A vector of norm 1 is called a Unit Vector.

The following Theorem shows that norms and distances in real inner product spaces have many of the properties that you might expect —

THEOREM 1 If $\mathbf{u}$ & $\mathbf{v}$ are vectors in a real inner product space V, and if k is a scalar, then

(i) $\|\mathbf{v}\| \geq 0$ with equality iff $\mathbf{v} = 0$.

(ii) $\|k\mathbf{v}\| = |k| \, \|\mathbf{v}\|$.

(iii) $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$.

(iv) $d(\mathbf{u},\mathbf{v}) \geq 0$ with equality iff $\mathbf{u} = \mathbf{v}$.

WEIGHTED EUCLIDEAN INNER PRODUCT:

Although the Euclidean inner product is the most important inner product on R^n, there are various applications in which it is desirable to modify it by weighting each term differently.

More precisely, if $w_1, w_2, \dots, w_n$ are positive real numbers, which we will call weights, and 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 it can be shown that the formula

$$<\mathbf{u},\mathbf{v}> = w_1 u_1 v_1 + w_2 u_2 v_2 + \dots + w_n u_n v_n \qquad \text{——— (1)}$$

defines an inner product on R^n that we call Weighted Euclidean Inner Product with the weights $w_1, w_2, \dots, w_n$.

Note that the standard Euclidean inner product is the special case of the weighted Euclidean inner product in which all the weights are 1.

Example 1 **Weighted Euclidean Inner Product**

Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be vectors in R^2. Verify that the weighted Euclidean inner product $<\mathbf{u},\mathbf{v}> = 3u_1v_1 + 2u_2v_2$ satisfies the four inner product axioms.

Soln. Axiom 1 We have
$$<\mathbf{u},\mathbf{v}> = 3u_1v_1 + 2u_2v_2$$
$$= 3v_1u_1 + 2v_2u_2$$
i.e., $$<\mathbf{u},\mathbf{v}> = <\mathbf{v},\mathbf{u}>$$

Axiom 2 If $\mathbf{w} = (w_1, w_2)$ then
$$<\mathbf{u}+\mathbf{v}, \mathbf{w}> = 3(u_1+v_1)w_1 + 2(u_2+v_2)w_2$$
$$= 3(u_1w_1 + v_1w_1) + 2(u_2w_2 + v_2w_2)$$
$$= (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2)$$
$$= <\mathbf{u},\mathbf{w}> + <\mathbf{v},\mathbf{w}>$$

<u>Axiom ③</u> We have $\langle k\mathbf{u},\mathbf{v}\rangle = 3(ku_1)v_1 + 2(ku_2)v_2$
$$= k(3u_1v_1 + 2u_2v_2)$$
i.e., $\langle k\mathbf{u},\mathbf{v}\rangle = k\langle \mathbf{u},\mathbf{v}\rangle$

<u>Axiom ④</u>
$$\langle \mathbf{v},\mathbf{v}\rangle = 3(v_1v_1) + 2(v_2v_2)$$
$$= 3v_1^2 + 2v_2^2 \geq 0 \text{ with equality iff } v_1 = v_2 = 0, \text{ i.e., iff } \mathbf{v} = \mathbf{0}.$$

<u>NOTE</u>: It is important to keep in mind that Norm and Distance depend on the inner product being used. If the inner product is changed, then the Norms and Distances between vectors also change.

For example, for the vectors $\mathbf{u} = (1,0)$ and $\mathbf{v} = (0,1)$ in R^2 with Euclidean inner product $\langle \mathbf{u},\mathbf{v}\rangle = u_1v_1 + u_2v_2$

we have $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u},\mathbf{u}\rangle} = \sqrt{(1)(1)+(0)(0)} = 1$

and $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \|(1,-1)\|$
$$= \sqrt{\langle (1,-1),(1,-1)\rangle}$$
$$= \sqrt{(1)(1) + (-1)(-1)} = \sqrt{2}$$

but if we change to the weighted Euclidean inner product
$$\langle \mathbf{u},\mathbf{v}\rangle = 3u_1v_1 + 2u_2v_2$$

then we have $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u},\mathbf{u}\rangle}$
$$= \sqrt{\langle (1,0),(1,0)\rangle}$$
$$= \sqrt{3(1)(1) + 2(0)(0)} = \sqrt{3}$$

and $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$
$$= \|(1,-1)\|$$
$$= \sqrt{\langle (1,-1),(1,-1)\rangle}$$
$$= \sqrt{3(1)(1) + 2(-1)(-1)}$$
$$= \sqrt{5}$$

UNIT CIRCLES AND SPHERES IN INNER PRODUCT SPACES.

If V is an inner product space, then the set of points in V that satisfy $\|\mathbf{u}\| = 1$ is called the Unit Sphere or sometimes the Unit Circle in V.

Example ③ Unusual Unit Circles in R^2

(i) Sketch the unit circle in an xy-coordinate system in R^2 using Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 .$$

(ii) Sketch the unit circle in an xy-coordinate system in R^2 using weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{9} u_1 v_1 + \frac{1}{4} u_2 v_2$.

Soln. (i) If $\mathbf{u} = (x, y)$ in R^2, then

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$
$$= \sqrt{\langle (x,y), (x,y) \rangle}$$
$$= \sqrt{(x)(x) + (y)(y)} = \sqrt{x^2+y^2}$$

so the equ. of unit circle is $\|\mathbf{u}\| = 1$

i.e., $\sqrt{x^2+y^2} = 1$

$\Rightarrow x^2+y^2 = 1$

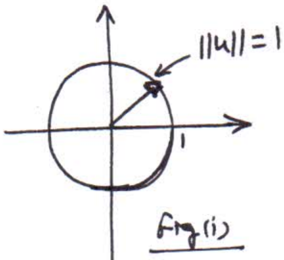

Fig (i)

The Unit Circle using Standard Euclidean Inner

As expected, the graph of this equ. is a circle of radius 1 centered at Origin (See Fig (i))

Soln (ii) If $\mathbf{u} = (x, y) \in R^2$, then

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$
$$= \sqrt{\langle (x,y), (x,y) \rangle}$$
$$= \sqrt{\tfrac{1}{9}(x)(x) + \tfrac{1}{4}(y)(y)}$$
$$= \sqrt{\tfrac{1}{9}x^2 + \tfrac{1}{4}y^2}$$

so the equ. of unit circle is $\|\mathbf{u}\| = 1$

i.e., $\sqrt{\frac{1}{9}x^2 + \frac{1}{4}y^2} = 1$

$\Rightarrow \dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$

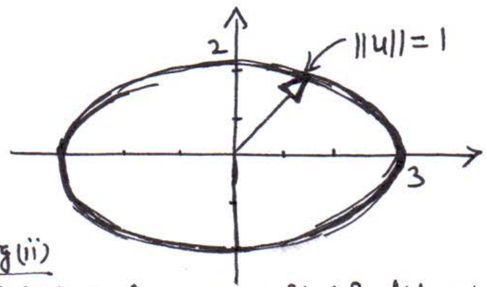

Fig (ii)

The Unit Circle using weighted Euclidean inner produ

The graph of this equ. is the ellipse shown in Fig (ii).

INNER PRODUCTS GENERATED BY MATRICES.

The Euclidean inner product and the weighted Euclidean inner products are special cases of a general class of inner products on R^n called Matrix Inner Products. To define this class of inner products, let $\mathbf{u}$ and $\mathbf{v}$ are vectors in R^n that are expressed in column form, and let A is an invertible $n \times n$ matrix. It can be shown that if $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v} \qquad \text{——— ①}$$

also defines an inner product; it is called the Inner Product on R^n generated by A.

Recall from Section (3.2) that if $\mathbf{u}$ and $\mathbf{v}$ are in column form, then $\mathbf{u} \cdot \mathbf{v}$ can be written as $\mathbf{v}^T\mathbf{u}$ from which it follows that ① can be expressed as

$$\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{v})^T A\mathbf{u}$$

or $$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^T A^T A \mathbf{u} \qquad \text{——— ②}$$

Example ④ Matrices Generating Weighted Euclidean Inner Products

The standard Euclidean and weighted Euclidean inner products are examples of matrix inner products. The standard Euclidean inner product on R^n is generated by the $n \times n$ identity matrix, since setting $A = I$ in formula ① yields

$$\langle \mathbf{u}, \mathbf{v} \rangle = I\mathbf{u} \cdot I\mathbf{v}$$
$$= \mathbf{u} \cdot \mathbf{v}$$

and the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n \qquad \text{——— ③}$$

is generated by the matrix

$$A = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{w_n} \end{bmatrix} \qquad \text{——— ④}$$

This can be seen by first observing that $A^T A$ is the $n \times n$ diagonal matrix whose diagonal entries are the weights $w_1, w_2, \cdots, w_n$ and then observing that ② simplifies to ③ when A is the matrix in formula ④.

Example ⑤ Example ① Revisited

The weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1 v_1 + 2u_2 v_2$ discussed in Example ① is the inner product on R^2 generated by

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

OTHER EXAMPLES OF INNER PRODUCTS.

So far, we have only considered examples of inner products on $\mathbb{R}^n$. We will now consider examples of inner products on some of the other kinds of vector spaces that we discussed earlier.

Example ⑥ An Inner Product on M_{nn}

① If U and V are $n \times n$ matrices, then the formula

$$\langle U, V \rangle = \operatorname{trace}(U^T V) \qquad \text{——— ①}$$

defines an inner product on the vector space M_{nn}

for the 2x2 matrices, $U = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}$ and $V = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$

we have
$$U^T V = \begin{bmatrix} u_1 & u_3 \\ u_2 & u_4 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$$

$$U^T V = \begin{bmatrix} u_1 v_1 + u_3 v_3 & u_1 v_2 + u_3 v_4 \\ u_2 v_1 + u_4 v_3 & u_2 v_2 + u_4 v_4 \end{bmatrix}$$

∴ from ①, $\langle U, V \rangle = \operatorname{trace}(U^T V)$

$$\langle U, V \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4 \qquad \text{——— ②}$$

which is just the dot product of the corresponding entries in the two matrices.

for example, if $U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$

then $\langle U, V \rangle = (1)(-1) + (2)(0) + (3)(3) + (4)(2)$

$= 16$

② The Norm of a matrix U relative to this inner product is

$$\|U\| = \sqrt{\langle U, U \rangle}$$

$$= \sqrt{(u_1)(u_1) + (u_2)(u_2) + (u_3)(u_3) + (u_4)(u_4)}$$

$$\|U\| = \sqrt{u_1^2 + u_2^2 + u_3^2 + u_4^2} \qquad \text{——— ③}$$

For example, if $U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

then $\|U\| = \sqrt{(1)^2 + (2)^2 + (3)^2 + (4)^2} = \sqrt{30}$

Question: Compute $\langle u, v \rangle$ using inner product on $M_{2\times2}$, where $u = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$ & $v = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}$.

Soln. We have
$$\langle u, v \rangle = \left\langle \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \right\rangle$$

$$= (1)(-1) + (1)(0) + (2)(1) + (3)(2)$$

$$= -1 + 0 + 2 + 6 = 7$$

Example ⑦ The Standard Inner Product on P_n

If $p = a_0 + a_1x + \cdots + a_nx^n$ and $q = b_0 + b_1x + \cdots + b_nx^n$ are polynomials in P_n, then the following formula defines an inner product on P_n that we will call the Standard Inner Product on this space:

$$\langle p, q\rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n \quad \text{——— ①}$$

The Norm of a polynomial p relative to this inner product is

$$\|p\| = \sqrt{\langle p, p\rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2} \quad \text{——— ②}$$

Example ⑧ The Evaluation Inner Product on P_n

If $p = p(x) = a_0 + a_1x + \cdots + a_nx^n$ and $q = q(x) = b_0 + b_1x + \cdots + b_nx^n$ are polynomials in P_n and if $x_0, x_1, \ldots, x_n$ are distinct real numbers (called sample points), then the formula

$$\langle p, q\rangle = p(x_0)\,q(x_0) + p(x_1)\,q(x_1) + \cdots + p(x_n)\,q(x_n) \quad \text{——— ①}$$

defines an inner product on P_n, called the Evaluation Inner Product at $x_0, x_1, \ldots, x_n$.

Algebraically, this can be viewed as the dot product in R^n of the n-tuples

$$\big(p(x_0), p(x_1), \ldots, p(x_n)\big) \quad \text{and} \quad \big(q(x_0), q(x_1), \ldots, q(x_n)\big)$$

The Norm of a polynomial p relative to the Evaluation Inner Product is

$$\|p\| = \sqrt{\langle p, p\rangle} = \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + \cdots + [p(x_n)]^2} \quad \text{——— ②}$$

Example ⑨ Working With the Evaluation Inner Product

Let P_2 have the evaluation inner product at the points $x_0 = -2$, $x_1 = 0$ and $x_2 = 2$. Compute $\langle p, q\rangle$ and $\|p\|$ for the polynomials $p = p(x) = x^2$ and $q = q(x) = 1 + x$.

Solution:

$$\begin{aligned}\langle p, q\rangle &= p(-2)\,q(-2) + p(0)\,q(0) + p(2)\,q(2)\\ &= (-2)^2(1-2) + (0)(1+0) + (2)^2(1+2)\\ &= -4 + 0 + 12 = 8\end{aligned}$$

$$\begin{aligned}\|p\| &= \sqrt{[p(x_0)]^2 + [p(x_1)]^2 + [p(x_2)]^2}\\ &= \sqrt{[p(-2)]^2 + [p(0)]^2 + [p(2)]^2}\\ &= \sqrt{(4)^2 + 0 + (4)^2}\\ &= 4\sqrt{2}\end{aligned}$$

$$\left\{\begin{aligned} \because\ p(-2) &= (-2)^2 = 4\\ p(0) &= 0\\ p(2) &= (2)^2 = 4\end{aligned}\right.$$

Example ⑩ Working With Standard Inner Product on P_n

If P_2 have usual inner product on polynomials and $p = 1 - 2x + 3x^2$, $q = 3 + x^2$ are polynomials. Then find — (i) $\|p\|$ (ii) $\|q\|$ (iii) $\langle p, q\rangle$.

Soln. (i) $\|p\| = \sqrt{\langle p, p\rangle} = \sqrt{(1)^2 + (-2)^2 + (3)^2} = \sqrt{14}$

(ii) $\|q\| = \sqrt{\langle q, q\rangle} = \sqrt{(3)^2 + (0)^2 + (1)^2} = \sqrt{10}$

$\boxed{\begin{array}{l}\text{for } p = 1-2x+3x^2,\ a_0=1, a_1=-2, a_2=3\\ \text{for } q = 3+x^2,\ b_0=3, b_1=0, b_2=1\end{array}}$

(iii) $\langle p, q\rangle = (1)(3) + (-2)(0) + (3)(1) = 6$.

SEC (6.2) ANGLE AND ORTHOGONALITY IN INNER PRODUCT SPACES.

In Section (3.2), we defined the notion of 'angle' between vectors in R^n. In this Section, we will extend this idea to general vector spaces. This will enable us to extend the notion of orthogonality as well.

CAUCHY-SCHWARZ INEQUALITY

Recall from Sec (3.2) that the angle θ between two vectors u and v in R^n is

$$\theta = \cos^{-1}\left(\frac{u.v}{\|u\|\ \|v\|}\right) \quad \text{——— (1)}$$

We were assured that this formula was valid because it followed from Cauchy-Schwarz inequality that

$$-1 \le \frac{u.v}{\|u\|\ \|v\|} \le 1 \quad \text{——— (2)}$$

as required for the inverse cosine to be defined.

THEOREM (1) Cauchy-Schwarz Inequality

If u and v are vectors in a real inner product space V, then $|\langle u,v\rangle| \le \|u\|\ \|v\|$ — (3)

The following two alternative forms of Cauchy-Schwarz inequality are useful to know —

$$\langle u,v\rangle^2 \le \langle u,u\rangle \langle v,v\rangle$$

$$\langle u,v\rangle^2 \le \|u\|^2\ \|v\|^2$$

ANGLE BETWEEN VECTORS

Let θ is the angle between u and v in a real inner product space, then

$$\cos\theta = \frac{\langle u,v\rangle}{\|u\|\ \|v\|} \ , \quad 0 \le \theta \le \pi \quad \text{——— (4)}$$

Example (1) Cosine of an Angle Between Two Vectors in R^4

Let R^4 have the Euclidean inner product. Find the cosine of the angle θ between the vectors $u = (4,3,1,-2)$ and $v = (-2,1,2,3)$.

Soln. Here $\|u\| = \sqrt{\langle u,u\rangle} = \sqrt{\langle (4,3,1,-2),(4,3,1,-2)\rangle}$

$$= \sqrt{(4)(4)+(3)(3)+(1)(1)+(-2)(-2)} = \sqrt{30}$$

$$\|v\| = \sqrt{\langle v,v\rangle} = \sqrt{\langle (-2,1,2,3),(-2,1,2,3)\rangle}$$

$$= \sqrt{(-2)(-2)+(1)(1)+(2)(2)+(3)(3)} = \sqrt{18}$$

& $\langle u,v\rangle = \langle (4,3,1,-2),(-2,1,2,3)\rangle$

$$= (4)(-2)+(3)(1)+(1)(2)+(-2)(3) = -9$$

Now $\cos\theta = \dfrac{\langle u,v\rangle}{\|u\|\ \|v\|} = \dfrac{-9}{\sqrt{30}\sqrt{18}}$

$$= \frac{-9}{3\sqrt{30}\sqrt{2}} = \frac{-3}{\sqrt{60}}$$

THEOREM ② If u, v & w are vectors in a real inner product space V and if k is any scalar, then

(i) $\|u+v\| \le \|u\| + \|v\|$ [Triangle Inequality for Vectors]

(ii) $d(u,v) \le d(u,w) + d(w,v)$ [Triangle Inequality for Distances]

ORTHOGONALITY: You should be able to see from formula ④ that if u and v are non-zero vectors, then the angle between them is $\theta = \frac{\pi}{2}$ iff $\langle u,v\rangle = 0$. Accordingly, we make the following defi. (which is applicable even if one or both of the vectors is zero).

Definition: Two vectors u and v in an inner product space are called Orthogonal if $\langle u,v\rangle = 0$.

Orthogonality depends on the inner product in the sense that for different inner products two vectors can be Orthogonal with respect to one but not the other.

Example ② Orthogonality Depends on the Inner Product

The vector $u = (1,1)$ and $v = (1,-1)$ are Orthogonal with respect to Euclidean inner product on R^2 ie., $\langle u,v\rangle = u_1v_1 + u_2v_2$, since

$$\langle u,v\rangle = (1)(1) + (1)(-1)$$
$$= 0$$

However, they are not Orthogonal with respect to weighted Euclidean inner product defined by $\langle u,v\rangle = 3u_1v_1 + 2u_2v_2$, since

$$\langle u,v\rangle = 3(1)(1) + 2(1)(-1)$$
$$= 1 \neq 0$$

Example ③ Orthogonal Vectors in M_{22}

Show that the matrices $U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ are Orthogonal if M_{22} has inner product defined in Sec 6.1 (Example ⑥).

Soln. Here $U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$

$$\therefore \langle U,V\rangle = (1)(0) + (0)(2) + (1)(0) + (1)(0)$$
$$= 0$$

Thus, the matrices U & V are Orthogonal.

Example (4) Orthogonal Vectors in P_2

Let P_2 have the inner product $\langle \mathbf{p},\mathbf{q}\rangle = \int_{-1}^{1} p(x)q(x)\,dx$ and let $\mathbf{p}=x$ and $\mathbf{q}=x^2$.
Find $\|\mathbf{p}\|$, $\|\mathbf{q}\|$ and show that $\mathbf{p}=x$ and $\mathbf{q}=x^2$ are Orthogonal relative to given inner product.

Soln.

$$\|\mathbf{p}\| = [\langle \mathbf{p},\mathbf{p}\rangle]^{1/2} = \left[\int_{-1}^{1} x\,x\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^2\,dx\right]^{1/2} = \left[\left(\frac{x^3}{3}\right)_{-1}^{1}\right]^{1/2} = \sqrt{\frac{2}{3}}$$

$$\|\mathbf{q}\| = [\langle \mathbf{q},\mathbf{q}\rangle]^{1/2} = \left[\int_{-1}^{1} x^2x^2\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^4\,dx\right]^{1/2} = \left[\left(\frac{x^5}{5}\right)_{-1}^{1}\right]^{1/2} = \sqrt{\frac{2}{5}}$$

Now $\langle \mathbf{p},\mathbf{q}\rangle = \int_{-1}^{1} x\,x^2\,dx = \int_{-1}^{1} x^3\,dx = \left[\frac{x^4}{4}\right]_{-1}^{1} = \frac{1}{4}\left[(1)^4-(-1)^4\right] = 0$

Since $\langle \mathbf{p},\mathbf{q}\rangle = 0$, the vectors $\mathbf{p}=x$ and $\mathbf{q}=x^2$ are Orthogonal relative to given inner product

THEOREM (3) Generalized Theorem of Pythagoras

If $\mathbf{u}$ and $\mathbf{v}$ are Orthogonal vectors in an inner product space, then

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

Example (5) Theorem of Pythagoras in P_2

In Example (4), we showed that $\mathbf{p}=x$ and $\mathbf{q}=x^2$ are Orthogonal with respect to inner product

$$\langle \mathbf{p},\mathbf{q}\rangle = \int_{-1}^{1} p(x)q(x)\,dx \quad \text{on } P_2.$$

It follows from Theorem (3) that

$$\|\mathbf{p}+\mathbf{q}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{q}\|^2$$

Thus from the computations in Example (4), we have

$$\|\mathbf{p}+\mathbf{q}\|^2 = \left(\sqrt{\frac{2}{3}}\right)^2 + \left(\sqrt{\frac{2}{5}}\right)^2$$

$$= \frac{2}{3} + \frac{2}{5} = \frac{16}{15}$$

We can check this result by direct integration as —

$$\|\mathbf{p}+\mathbf{q}\|^2 = \langle \mathbf{p}+\mathbf{q},\, \mathbf{p}+\mathbf{q}\rangle$$

$$= \int_{-1}^{1} (x+x^2)(x+x^2)\,dx$$

$$= \int_{-1}^{1} (x^2+2x^3+x^4)\,dx$$

$$= \left[\frac{x^3}{3} + 2\frac{x^4}{4} + \frac{x^5}{5}\right]_{-1}^{1}$$

$$= \frac{1}{3}\left[(1)^3-(-1)^3\right] + \frac{1}{2}\left[(1)^4-(-1)^4\right] + \frac{1}{5}\left[(1)^5-(-1)^5\right]$$

$$= \frac{2}{3} + 0 + \frac{2}{5} = \frac{16}{15}.$$

SEC (6.3) GRAM-SCHMIDT PROCESS ; QR-DECOMPOSITION

In many problems involving vector spaces, the problem solver is free to choose any basis for the vector space that seems appropriate. In inner product spaces, the solution of a problem is often greatly simplified by choosing a basis in which the vectors are orthogonal to one another.

ORTHOGONAL AND ORTHONORMAL SETS : Recall from Sec (6.2) that two vectors in an inner product space are said to be Orthogonal if their inner product is **zero**. The following definition extends the notion of Orthogonality to set of vectors in an inner product space.

Definition : A set of two or more vectors in a real inner product space is said to be Orthogonal if all pairs of distinct vectors in the set are orthogonal.

An orthogonal set in which each vector has norm 1 is said to be Orthonormal.

Example (1) An Orthogonal Set in R^3

Show that the set $S=\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$ is Orthogonal, where $\mathbf{u}_1=(0,1,0)$, $\mathbf{u}_2=(1,0,1)$, $\mathbf{u}_3=(1,0,-1)$ and assume that R^3 has the Euclidean inner product.

Soln. We have $\mathbf{u}_1=(0,1,0)$, $\mathbf{u}_2=(1,0,1)$, $\mathbf{u}_3=(1,0,-1)$

Now $\langle \mathbf{u}_1,\mathbf{u}_2\rangle = (0)(1)+(1)(0)+(0)(1) = 0$

$\langle \mathbf{u}_1,\mathbf{u}_3\rangle = (0)(1)+(1)(0)+(0)(-1) = 0$

& $\langle \mathbf{u}_2,\mathbf{u}_3\rangle = (1)(1)+(0)(0)+(1)(-1) = 0$

Hence the set of vectors $S=\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3\}$ is Orthogonal.

NORMALIZATION : If V is a non-zero vector in an inner product space, then it follows that

$$\left\| \frac{1}{\|V\|} V \right\| = \left| \frac{1}{\|V\|} \right| \|V\| = \frac{1}{\|V\|}\|V\| = 1$$

from which we see that multiplying a non-zero vector by the reciprocal of its norm produces a vector of norm 1. This process is called Normalizing vector V.

It follows that any Orthogonal set of non-zero vectors can be converted to an Orthonormal set by normalizing each of its vectors.

Example ② **Constructing an Orthonormal Set**

The Euclidean norms of the vectors in Example ① are

$$\|\mathbf{u}_1\| = \sqrt{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} = \sqrt{(0)(0) + (1)(1) + (0)(0)} = 1$$

$$\|\mathbf{u}_2\| = \sqrt{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} = \sqrt{(1)(1) + (0)(0) + (1)(1)} = \sqrt{2}$$

$$\|\mathbf{u}_3\| = \sqrt{\langle \mathbf{u}_3, \mathbf{u}_3 \rangle} = \sqrt{(1)(1) + (0)(0) + (-1)(-1)} = \sqrt{2}$$

Consequently, normalizing $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$ yields

$$\mathbf{v}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{1}(0,1,0) = (0,1,0)$$

$$\mathbf{v}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$$\mathbf{v}_3 = \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = \frac{1}{\sqrt{2}}(1,0,-1) = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

We can verify that the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is Orthonormal by showing that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$$

and $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$.

In R^2, any two non-zero perpendicular vectors are linearly independent because neither is a scalar multiple of the other and in R^3, any three non-zero mutually perpendicular vectors are linearly independent because no one lies in the plane of the other two (and hence is not expressible as a linear combination of the other two). The following theorem generalizes these observations —

THEOREM ① If $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an Orthogonal set of non-zero vectors in an inner product space, then the set S is linearly independent.

> Since an Orthonormal set is Orthogonal, it follows that every Orthonormal set is linearly independent.

ORTHONORMAL BASIS: In an inner product space, a basis consisting of orthonormal vectors is called an Orthonormal Basis. A familiar example of an orthonormal basis is the standard basis for R^n with the Euclidean inner product:

$$\mathbf{e}_1 = (1,0,0,\cdots,0),\ \mathbf{e}_2 = (0,1,0,\cdots,0),\ \cdots\cdots,\ \mathbf{e}_n = (0,0,0,\cdots,1).$$

Example ③ **An Orthonormal Basis**

~~And an Orthonormal basis~~ In Example ②, we showed that the vectors

$$\mathbf{v}_1 = (0,1,0),\quad \mathbf{v}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \text{ and } \mathbf{v}_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

form an Orthonormal set with respect to the Euclidean inner product on R^3. By Theorem ①, these vectors form a linearly independent set and since R^3 is three-dimensional, it follows that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an Orthonormal basis for R^3.

COORDINATES RELATIVE TO ORTHONORMAL BASIS.

One way to express a vector $\mathbf{u}$ as a linear combination of basis vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is to convert the vector equ. $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$

to a linear system and solve this system for the coefficients $c_1, c_2, \cdots, c_n$.

However, if the basis happen to be Orthogonal or Orthonormal, then the following theorem shows that the coefficients can be obtained more simply by computing appropriate inner products.

THEOREM (2)

(i) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an Orthogonal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n\rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n \quad \text{--- (1)}$$

(ii) If $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an Orthonormal basis for an inner product space V, and if $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n\rangle \mathbf{v}_n \quad \text{--- (2)}$$

Example (4) A Coordinate Vector Relative to an Orthonormal Basis

Show that the vectors $\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$, $\mathbf{v}_3 = \left(\frac{3}{5}, 0\, \frac{4}{5}\right)$ form an Orthonormal basis for R^3 with the Euclidean inner product. Express the vector $\mathbf{u} = (1,1,1)$ as a linear combination of the vectors in $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and find the Co-ordinate vector $(\mathbf{u})_S$.

Soln. Here $\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$, $\mathbf{v}_3 = \left(\frac{3}{5}, 0, \frac{4}{5}\right)$

Now $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = (0)\left(-\frac{4}{5}\right) + (1)(0) + (0)\left(\frac{3}{5}\right) = 0$

$\langle \mathbf{v}_1, \mathbf{v}_3\rangle = (0)\left(\frac{3}{5}\right) + (1)(0) + (0)\left(\frac{4}{5}\right) = 0$

& $\langle \mathbf{v}_2, \mathbf{v}_3\rangle = \left(-\frac{4}{5}\right)\left(\frac{3}{5}\right) + (0)(0) + \left(\frac{3}{5}\right)\left(\frac{4}{5}\right) = 0$

Also $\|\mathbf{v}_1\| = \sqrt{(0)(0) + (1)(1) + (0)(0)} = 1$

$\|\mathbf{v}_2\| = \sqrt{\left(-\frac{4}{5}\right)\left(-\frac{4}{5}\right) + (0)(0) + \left(\frac{3}{5}\right)\left(\frac{3}{5}\right)} = \sqrt{\frac{16}{25} + \frac{9}{25}} = 1$

& $\|\mathbf{v}_3\| = \sqrt{\left(\frac{3}{5}\right)\left(\frac{3}{5}\right) + (0)(0) + \left(\frac{4}{5}\right)\left(\frac{4}{5}\right)} = \sqrt{\frac{9}{25} + \frac{16}{25}} = 1$

Hence, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ form an Orthonormal basis for R^3.

Now $\langle \mathbf{u}, \mathbf{v}_1\rangle = \langle (1,1,1), (0,1,0)\rangle = (1)(0) + (1)(1) + (1)(0) = 1$

$\langle \mathbf{u}, \mathbf{v}_2\rangle = \langle (1,1,1), \left(-\frac{4}{5}, 0, \frac{3}{5}\right)\rangle = (1)\left(-\frac{4}{5}\right) + (1)(0) + (1)\left(\frac{3}{5}\right) = -\frac{1}{5}$

$\langle \mathbf{u}, \mathbf{v}_3\rangle = \langle (1,1,1), \left(\frac{3}{5}, 0, \frac{4}{5}\right)\rangle = (1)\left(\frac{3}{5}\right) + (1)(0) + (1)\left(\frac{4}{5}\right) = \frac{7}{5}$

Therefore, by Theorem (2), we have

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 + \langle \mathbf{u}, \mathbf{v}_3\rangle \mathbf{v}_3$$

$$\mathbf{u} = \mathbf{v}_1 - \frac{1}{5}\mathbf{v}_2 + \frac{7}{5}\mathbf{v}_3$$

Thus, the Coordinate vector of $\mathbf{u}$ relative to orthonormal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is

$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1\rangle, \langle \mathbf{u}, \mathbf{v}_2\rangle, \langle \mathbf{u}, \mathbf{v}_3\rangle)$$

$$= \left(1, -\frac{1}{5}, \frac{7}{5}\right).$$

Example (5) An Orthonormal Basis from an Orthogonal Basis

(i) Show that the vectors $w_1 = (0,2,0)$, $w_2 = (3,0,3)$, $w_3 = (-4,0,4)$ form an Orthogonal basis for R^3 with the Euclidean inner product, and use that basis to find an Orthonormal basis by normalizing each vector.

(ii) Express vector $u = (1,2,4)$ as a linear combination of orthonormal basis vectors obtained in part (i).

Solution (i) Here $w_1 = (0,2,0)$, $w_2 = (3,0,3)$, $w_3 = (-4,0,4)$

Now $\langle w_1, w_2 \rangle = (0)(3) + (2)(0) + (0)(3) = 0$

$\langle w_1, w_3 \rangle = (0)(-4) + (2)(0) + (0)(4) = 0$

& $\langle w_2, w_3 \rangle = (3)(-4) + (0)(0) + (3)(4) = 0$

So the vectors w_1, w_2, w_3 form an Orthogonal set.
It follows from Theorem (1) that these vectors are linearly independent and hence form a basis for R^3.

To find Orthonormal basis:

$$||w_1|| = \sqrt{\langle w_1, w_1 \rangle} = \sqrt{(0)(0)+(2)(2)+(0)(0)} = 2$$

$$||w_2|| = \sqrt{\langle w_2, w_2 \rangle} = \sqrt{(3)(3)+(0)(0)+(3)(3)} = \sqrt{18} = 3\sqrt{2}$$

$$\& \ ||w_3|| = \sqrt{\langle w_3, w_3 \rangle} = \sqrt{(-4)(4)+(0)(0)+(4)(4)} = \sqrt{32} = 4\sqrt{2}$$

Consequently, normalizing w_1, w_2 & w_3 yields

$$v_1 = \frac{w_1}{||w_1||} = \frac{1}{2}(0,2,0) = (0,1,0)$$

$$v_2 = \frac{w_2}{||w_2||} = \frac{1}{3\sqrt{2}}(3,0,3) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$$v_3 = \frac{w_3}{||w_3||} = \frac{1}{4\sqrt{2}}(-4,0,4) = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

We can verify that the set $S = \{v_1, v_2, v_3\}$ is Orthonormal.

(ii) It follows from Theorem (2) that

$$u = \langle u, v_1 \rangle v_1 + \langle u, v_2 \rangle v_2 + \langle u, v_3 \rangle v_3 \quad \text{——— (1)}$$

Now $\langle u, v_1 \rangle = \langle (1,2,4), (0,1,0) \rangle = (1)(0) + (2)(1) + (4)(0) = 2$

$\langle u, v_2 \rangle = \langle (1,2,4), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \rangle = (1)\left(\frac{1}{\sqrt{2}}\right) + (2)(0) + (4)\left(\frac{1}{\sqrt{2}}\right) = \frac{5}{\sqrt{2}}$

$\langle u, v_3 \rangle = \langle (1,2,4), \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \rangle = (1)\left(-\frac{1}{\sqrt{2}}\right) + (2)(0) + (4)\left(\frac{1}{\sqrt{2}}\right) = \frac{3}{\sqrt{2}}$

Hence from (1),

$$(1,2,4) = 2(0,1,0) + \frac{5}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + \frac{3}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

THEOREM ③ Every non-zero finite dimensional inner product space has an Orthonormal Basis.

The step-by-step construction of an Orthogonal (or Orthonormal) basis is called the Gram-Schmidt Process. We provide the following summary of the steps —

The GRAM-SCHMIDT PROCESS : To convert a basis $\{u_1, u_2, \cdots, u_r\}$ into an Orthogonal basis $\{v_1, v_2, \cdots, v_r\}$, perform the following computations —

Step ① Take $v_1 = u_1$.

Step ② Take $v_2 = u_2 - \dfrac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$

Step ③ Take $v_3 = u_3 - \dfrac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \dfrac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$

⋮

(Continue for r steps)

Optional Step : To convert the orthogonal basis into an orthonormal basis $\{q_1, q_2, \cdots, q_r\}$, normalize the orthogonal basis vectors.

Example Using Gram-Schmidt Process

Assume that the vector space R^3 has the Euclidean inner product. Apply Gram-Schmidt process to transform the basis vectors $u_1 = (1,1,1)$, $u_2 = (0,1,1)$, $u_3 = (0,0,1)$ into an Orthogonal basis $\{v_1, v_2, v_3\}$ and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{q_1, q_2, q_3\}$.

Soln.

Step ① $\quad v_1 = u_1 = (1,1,1)$

Step ②

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$$

$$= (0,1,1) - \left[\frac{\langle (0,1,1),(1,1,1) \rangle}{\langle v_1, v_1 \rangle}\right] v_1$$

$$= (0,1,1) - \left[\frac{\langle (0,1,1),(1,1,1) \rangle}{\langle (1,1,1),(1,1,1) \rangle}\right] v_1$$

$$= (0,1,1) - \frac{2}{3}(1,1,1)$$

$$= (0,1,1) - \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = \left(0-\frac{2}{3}, 1-\frac{2}{3}, 1-\frac{2}{3}\right)$$

$$\Rightarrow v_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Step (3)

$$V_3 = u_3 - \frac{\langle u_3, V_1\rangle}{\|V_1\|^2} V_1 - \frac{\langle u_3, V_2\rangle}{\|V_2\|^2} V_2$$

$$= (0,0,1) - \left[\frac{\langle (0,0,1),(1,1,1)\rangle}{\langle (1,1,1),(1,1,1)\rangle}\right] V_1 - \left[\frac{\langle (0,0,1),(-\frac{2}{3},\frac{1}{3},\frac{1}{3})\rangle}{\langle (-\frac{2}{3},\frac{1}{3},\frac{1}{3}),(-\frac{2}{3},\frac{1}{3},\frac{1}{3})\rangle}\right] V_2$$

$$= (0,0,1) - \frac{1}{3}(1,1,1) - \frac{(\frac{1}{3})}{(\frac{2}{3})}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right)$$

$$= (0,0,1) - \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) - \frac{1}{2}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right)$$

$$= (0,0,1) - \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) - \left(-\frac{1}{3},\frac{1}{6},\frac{1}{6}\right)$$

$$= \left(0-\frac{1}{3}+\frac{1}{3},\ 0-\frac{1}{3}-\frac{1}{6},\ 1-\frac{1}{3}-\frac{1}{6}\right)$$

$$\Rightarrow \quad V_3 = \left(0,-\frac{1}{2},\frac{1}{2}\right)$$

Thus $V_1 = (1,1,1)$, $V_2 = \left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right)$ & $V_3 = \left(0,-\frac{1}{2},\frac{1}{2}\right)$ form an Orthogonal basis for R^3.

The norms of these vectors are —

$$\|V_1\| = \sqrt{\langle V_1, V_1\rangle} = \sqrt{(1)(1)+(1)(1)+(1)(1)} = \sqrt{3}$$

$$\|V_2\| = \sqrt{\langle V_2, V_2\rangle} = \sqrt{\left(-\frac{2}{3}\right)\left(-\frac{2}{3}\right)+\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)+\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)} = \frac{\sqrt{6}}{3}$$

$$\& \quad \|V_3\| = \sqrt{\langle V_3, V_3\rangle} = \sqrt{(0)(0)+\left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right)+\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = \frac{1}{\sqrt{2}}$$

So an orthonormal basis for R^3 is

$$q_1 = \frac{V_1}{\|V_1\|} = \frac{1}{\sqrt{3}}(1,1,1) = \left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)$$

$$q_2 = \frac{V_2}{\|V_2\|} = \frac{1}{\frac{\sqrt{6}}{3}}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right) = \frac{3}{\sqrt{6}}\left(-\frac{2}{3},\frac{1}{3},\frac{1}{3}\right) = \left(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)$$

$$\& \quad q_3 = \frac{V_3}{\|V_3\|} = \frac{1}{\frac{1}{\sqrt{2}}}\left(0,-\frac{1}{2},\frac{1}{2}\right) = \sqrt{2}\left(0,-\frac{1}{2},\frac{1}{2}\right) = \left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$$

SEC 6.4 BEST APPROXIMATION ; LEAST SQUARES

In this Section, we will be concerned with linear systems that cannot be solved exactly and for which an approximate solution is needed. Such systems commonly occur in applications where measurement errors 'perturb' the coefficients of a consistent system sufficiently to produce inconsistency.

LEAST SQUARES SOLUTIONS OF LINEAR SYSTEMS.

Suppose that $A\mathbf{x}=\mathbf{b}$ is an inconsistent linear system of m-equations in n-unknowns in which we suspect the inconsistency to be caused by measurement errors in the coefficients of 'A'. Since no exact solution is possible, we will look for a vector $\mathbf{x}$ that comes as 'close as possible' to being a solution in the sense that it minimizes $\|\mathbf{b}-A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. We can think of $A\mathbf{x}$ as an approximation to '$\mathbf{b}$' and $\|\mathbf{b}-A\mathbf{x}\|$ as the error in that approximation – the smaller the error, the better the approximation. This leads to the following problem –

LEAST SQUARES PROBLEM.

Given a linear system $A\mathbf{x}=\mathbf{b}$ of m-equations in n-unknowns, find a vector '$\mathbf{x}$' that minimizes $\|\mathbf{b}-A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. We call such an '$\mathbf{x}$' a least squares solution of the system, we call $\mathbf{b}-A\mathbf{x}$ the least squares error vector and we call $\|\mathbf{b}-A\mathbf{x}\|$ the least squares error.

NOTE: To clarify the above terminology, suppose that the matrix form of $\mathbf{b}-A\mathbf{x}$ is

$$\mathbf{b}-A\mathbf{x} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$

The term 'least squares solution' results from the fact that minimizing $\|\mathbf{b}-A\mathbf{x}\|$ also minimizes $\|\mathbf{b}-A\mathbf{x}\|^2 = e_1^2+e_2^2+ \cdots + e_m^2$.

THEOREM 1 Best Approximation Theorem.

If W is a finite-dimensional subspace of an inner product space V, and if '$\mathbf{b}$' is a vector in V, then $\text{proj}_W \mathbf{b}$ is the Best Approximation to '$\mathbf{b}$' from W in the sense that

$$\|\mathbf{b}-\text{proj}_W \mathbf{b}\| < \|\mathbf{b}-\mathbf{w}\|$$

for every vector $\mathbf{w}$ in W that is different from $\text{proj}_W \mathbf{b}$.

LEAST SQUARES SOLU. OF LINEAR SYSTEMS.

THEOREM ② For every linear system $A\mathbf{x}=\mathbf{b}$, the associated normal system

$$A^TA\mathbf{x} = A^T\mathbf{b} \quad \text{——— ①}$$

is consistent and all solutions of ① are least squares solutions of $A\mathbf{x}=\mathbf{b}$.

Moreover, if W is a column space of A, and '$\mathbf{x}$' is any least squares solu. of $A\mathbf{x}=\mathbf{b}$, then the orthogonal projection of '$\mathbf{b}$' on 'W' is

$$\text{proj}_W \mathbf{b} = A\mathbf{x} \quad \text{——— ②}$$

NOTE: If a linear system is consistent, then its exact solutions are the same as its least squares solutions, in which case the error is zero.

Example ① **Least Squares Solution**

Find all least squares solutions of the linear system

$$\begin{aligned} x_1 - x_2 &= 4 \\ 3x_1 + 2x_2 &= 1 \\ -2x_1 + 4x_2 &= 3 \end{aligned}$$

Also find the error vector and the error.

Solution: The given system of equ. in matrix form is $A\mathbf{x}=\mathbf{b}$, where

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

It follows that
$$A^TA = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix}$$

i.e., $$A^TA = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

and $$A^T\mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

i.e., $$A^T\mathbf{b} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

so the Normal System $$A^TA\mathbf{x} = A^T\mathbf{b}$$

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

$$14x_1 - 3x_2 = 1 \quad \text{——— (i)}$$
$$-3x_1 + 21x_2 = 10 \quad \text{——— (ii)}$$

Solving this system yields a unique least squares solu., $x_1 = \frac{17}{95}$, $x_2 = \frac{143}{285}$.

The Error Vector is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{95}{57} \end{bmatrix} = \begin{bmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{bmatrix}$$

and the Error is
$$\|\mathbf{b} - A\mathbf{x}\| = \sqrt{\left(\frac{1232}{285}\right)^2 + \left(-\frac{154}{285}\right)^2 + \left(\frac{4}{3}\right)^2} \approx 4.556.$$

Example ② Orthogonal Projection on a Subspace

Find the orthogonal projection of the vector $\mathbf{u} = (-3,-3,8,9)$ on the subspace of R^4 spanned by the vectors $\mathbf{u}_1 = (3,1,0,1)$, $\mathbf{u}_2 = (1,2,1,1)$, $\mathbf{u}_3 = (-1,0,2,-1)$.

Soln: The subspace W of R^4 spanned by $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$ is the column space of matrix

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

Thus, if '$\mathbf{u}$' is expressed as a column vector, we can find the orthogonal projection of $\mathbf{u}$ on W by finding a least squares soln. of the system $A\mathbf{x} = \mathbf{u}$ and then calculating $\text{proj}_W \mathbf{u} = A\mathbf{x}$ from the least squares soln.

The system $A\mathbf{x} = \mathbf{u}$ is
$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \\ 8 \\ 9 \end{bmatrix}$$

so
$$A^TA = \begin{bmatrix} 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ -1 & 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}$$

i.e.,
$$A^TA = \begin{bmatrix} 11 & 6 & -4 \\ 6 & 7 & 0 \\ -4 & 0 & 6 \end{bmatrix}$$

and
$$A^T\mathbf{u} = \begin{bmatrix} 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ -1 & 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} -3 \\ -3 \\ 8 \\ 9 \end{bmatrix}$$

i.e.,
$$A^T\mathbf{u} = \begin{bmatrix} -3 \\ 8 \\ 10 \end{bmatrix}$$

The Normal System is
$$A^TA\mathbf{x} = A^T\mathbf{u}$$
$$\begin{bmatrix} 11 & 6 & -4 \\ 6 & 7 & 0 \\ -4 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 8 \\ 10 \end{bmatrix}$$

$$11x_1 + 6x_2 - 4x_3 = -3 \quad \text{—— (i)}$$
$$6x_1 + 7x_2 = 8 \quad \text{—— (ii)}$$
$$-4x_1 + 6x_3 = 10 \quad \text{—— (iii)}$$

Solving this system of eqn. yields least squares soln., $x_1 = -1$, $x_2 = 2$ & $x_3 = 1$

so
$$\text{proj}_W \mathbf{u} = A\mathbf{x} = A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= \begin{bmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} -2 \\ 3 \\ 4 \\ 0 \end{bmatrix}$$

OR $\text{proj}_W \mathbf{u} = (-2, 3, 4, 0)$.