Lecture no 6: INNER PRODUCT SPACES. EUCLIDEAN SPACE. ORTHOGONALITY [1]

Up to this point, we have made qualitative deductions about vector spaces, but we have not derived quantitative properties. An example of a quantitative property that would almost certainly be of interest is the *length of a vector*.

In this chapter, we introduce one way of forming the product of vectors, called the *inner product*, that associates a scalar with each pair of vectors. Among other things, an inner product enables us to generalize the notion of length that occurs in *Euclidean spaces*.

The terms "*inner product*" and "*dot product*" are often used synonymously, but they are not exactly the same. The dot product is used only when the scalar field is $\mathbb{R}$. The inner product generalizes the dot product and applies when the scalar field is $\mathbb{R}$ or $\mathbb{C}$.

1. Inner Product Spaces

Definition. Let V be a real vector space. Suppose to each pair of vectors $u, v \in V$ there is assigned a real number, denoted by $\langle u, v\rangle$. This function is called a *(real) inner product* on V if it satisfies the following axioms:

[I_1] (*Linear Property*): $\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$.

[I_2] (*Symmetric Property*): $\langle u, v\rangle = \langle v; u\rangle$.

[I_3] (*Positive Definite Property*): $\langle u, u\rangle \geq 0$.; and $\langle u, u\rangle = 0$ if and only if $u = 0$.

The vector space V with an inner product is called a *(real) inner product space*.

Axiom [I_1] states that an inner product function is linear in the first position. Using [I_1] and the symmetry axiom [I_2], we obtain

$$\langle u, \ cv_1 + dv_2\rangle = \langle cv_1 + dv_2, \ u\rangle = c\langle v_1, u\rangle + d\langle v_2, u\rangle = c\langle u, v_1\rangle + d\langle u, v_2\rangle$$

That is, the inner product function is also linear in its second position. Combining these

1

9	Л	2	Евклідов простір. Аксіоми скалярного добутку. Існування ортонормованого базису евклідового простору. Ортонормовані системи векторів. Процес ортогоналізації Грама-Шмідта.
10	ПР	2	Побудова ортонормованого базису евклідового простору.

two properties and using induction yields the following general formula:

$$\left\langle \sum_i a_i u_i,\ \sum_j b_j v_j \right\rangle = \sum_i \sum_j a_i b_j \langle u_i, v_j \rangle$$

That is, an inner product of linear combinations of vectors is equal to a linear combination of the inner products of the vectors.

Example Let V be a real inner product space. Then, by linearity

$$\begin{aligned}\langle 3u_1 - 4u_2,\ 2v_1 - 5v_2 + 6v_3 \rangle = 6\langle u_1, v_1\rangle - 15\langle u_1, v_2\rangle + 18\langle u_1, v_3\rangle \\ - 8\langle u_2, v_1\rangle + 20\langle u_2, v_2\rangle - 24\langle u_2, v_3\rangle\end{aligned}$$

$$\begin{aligned}\langle 2u - 5v,\ 4u + 6v\rangle &= 8\langle u, u\rangle + 12\langle u, v\rangle - 20\langle v, u\rangle - 30\langle v, v\rangle \\ &= 8\langle u, u\rangle - 8\langle v, u\rangle - 30\langle v, v\rangle\end{aligned}$$

Remark Axiom $[I_1]$ by itself implies $\langle 0, 0\rangle = \langle 0v, 0\rangle = 0\langle v, 0\rangle = 0$. Thus, $[I_1]$, $[I_2]$, $[I_3]$ are equivalent to $[I_1]$, $[I_2]$, and the following axiom:

$[I'_3]$ If $u \neq 0$; then $\langle u, u\rangle$ is positive.

That is, a function satisfying $[I_1]$, $[I_2]$, $[I'_3]$ is an inner product.

Norm of a Vector

By the third axiom $[I_3]$ of an inner product, $\langle u, u\rangle$ is nonnegative for any vector u. Thus, its positive square root exists. We use the notation

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

This nonnegative number is called the *norm or length* of u. The relation $||u||^2 = \langle u, u\rangle$ will be used frequently.

Remark If $||u|| = 1$ or, equivalently, if $\langle u, u\rangle = 1$ then u is called a unit vector and it is said to be *normalized*. Every nonzero vector v in V can be multiplied by the reciprocal of its length to obtain the *unit vector*

$$\hat{v} = {}^{v}/_{||v||}$$

which is a positive multiple of v. This process is called *normalizing v*.

Examples of Inner Product Spaces

1. Euclidean n-Space $\mathbf{R}^n$

Consider the vector space $\mathbf{R}^n$. The dot product or scalar product in $\mathbf{R}^n$ is defined by

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

where $u = (a_i)$ and $v = (b_i)$. This function defines an inner product on $\mathbf{R}^n$. The norm $\|u\|$ of the vector $u = (a_i)$ in this space is as follows

$$\|u\| = \sqrt{u \cdot u} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

On the other hand, by the Pythagorean theorem, the distance from the origin O in $\mathbf{R}^3$ to a point $P(a; b; c)$ is given by $\sqrt{a^2 + b^2 + c^2}$. This is precisely the same as the above-defined norm of the vector $v = (a; b; c)$ in $\mathbf{R}^3$.

Because the *Pythagorean theorem* is a consequence of the axioms of Euclidean geometry, the vector space $\mathbf{R}^n$ with the above inner product and norm is called *Euclidean n-space*.

Although there are many ways to define an inner product on $\mathbf{R}^n$, we shall assume this inner product unless otherwise stated or implied. It is called the *usual (or standard) inner product* on $\mathbf{R}^n$.

Remark Frequently the vectors in $\mathbf{R}^n$ will be represented by *column vectors*—that is, by $n \times 1$ column matrices. In such a case, the formula

$$\langle u, v \rangle = u^T v$$

defines the usual inner product on $\mathbf{R}^n$.

Example Let $u = (1; 3;-4; 2)$, $v = (4;-2; 2; 1)$, $w = (5;-1;-2; 6)$ in $\mathbf{R}^4$.

(a) Show $\langle 3u - 2v, w \rangle = 3\langle u, w \rangle - 2\langle v, w \rangle$.

By definition,

$$\langle u, w \rangle = 5 - 3 + 8 + 12 = 22 \qquad \text{and} \qquad \langle v, w \rangle = 20 + 2 - 4 + 6 = 24$$

Note that $3u - 2v = (-5, 13, -16, 4)$. Thus,

$$\langle 3u - 2v, \ w \rangle = -25 - 13 + 32 + 24 = 18$$

As expected, $3\langle u, w \rangle - 2\langle v, w \rangle = 3(22) - 2(24) = 18 = \langle 3u - 2v, \ w \rangle$.

(b) Normalize u and v.

By definition,

$$\|u\| = \sqrt{1+9+16+4} = \sqrt{30} \qquad \text{and} \qquad \|v\| = \sqrt{16+4+4+1} = 5$$

We normalize u and v to obtain the following unit vectors in the directions of u and v, respectively:

$$\hat{u} = \frac{1}{\|u\|}u = \left(\frac{1}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{-4}{\sqrt{30}}, \frac{2}{\sqrt{30}}\right) \qquad \text{and} \qquad \hat{v} = \frac{1}{\|v\|}v = \left(\frac{4}{5}, \frac{-2}{5}, \frac{2}{5}, \frac{1}{5}\right)$$

Function Space C[*a; b*] **and Polynomial Space P**(*t*)

The notation **C**[*a; b*] is used to denote the vector space of all continuous functions on the closed interval [*a; b*]—that is, where $a \le t \le b$. The following defines an inner product on **C**[*a; b*], where $f(t)$ and $g(t)$ are functions in **C**[*a; b*]:

$$\langle f, g\rangle = \int_a^b f(t)g(t)\, dt$$

It is called the *usual inner product* on **C**[*a; b*].

The vector space **P**(*t*) of all polynomials is a *subspace* of **C**[*a; b*] for any interval [*a; b*], and hence, the above is also an *inner product* on **P**(*t*).

Example

Consider $f(t) = 3t - 5$ and $g(t) = t^2$ in the polynomial space $\mathbf{P}(t)$ with inner product

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt.$$

(a) Find $\langle f, g\rangle$.

We have $f(t)g(t) = 3t^3 - 5t^2$. Hence,

$$\langle f, g\rangle = \int_0^1 (3t^3 - 5t^2)\, dt = \tfrac{3}{4}t^4 - \tfrac{5}{3}t^3\Big|_0^1 = \tfrac{3}{4} - \tfrac{5}{3} = -\tfrac{11}{12}$$

(b) Find $\|f\|$ and $\|g\|$.

We have $[f(t)]^2 = f(t)f(t) = 9t^2 - 30t + 25$ and $[g(t)]^2 = t^4$. Then

$$\|f\|^2 = \langle f, f\rangle = \int_0^1 (9t^2 - 30t + 25)\, dt = 3t^3 - 15t^2 + 25t\Big|_0^1 = 13$$

$$\|g\|^2 = \langle g, g\rangle = \int_0^1 t^4\, dt = \tfrac{1}{5}t^5\Big|_0^1 = \tfrac{1}{5}$$

Therefore, $\|f\| = \sqrt{13}$ and $\|g\| = \sqrt{\tfrac{1}{5}} = \tfrac{1}{5}\sqrt{5}$.

Matrix Space M = $\mathbf{M}_{m,n}$

Let **M** = $\mathbf{M}_{m;n}$, the vector space of all real $m \times n$ matrices. An *inner product* is defined on **M** by

$$\langle A, B\rangle = \mathrm{tr}\,(B^T A),$$

where, as usual, tr () is the *trace*—the sum of the diagonal elements. If $A = [a_{ij}]$ and $B = [b_{ij}]$, then

$$\langle A, B\rangle = \mathrm{Tr}\,(B^T A) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}, \text{ and}$$

$$\|A\|^2 = \langle A, A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2$$

That is, $\langle A, B\rangle$ is the sum of the products of the corresponding entries in A and B and, in particular, $\langle A, A\rangle$ is the sum of the squares of the entries of A.

Hilbert Space

Let V be the vector space of all infinite sequences of real numbers $(a_1; a_2; a_3; \ldots)$ satisfying

$$\sum_{i=1}^{\infty} a_i^2 = a_1^2 + a_2^2 + \cdots < \infty$$

that is, the sum *converges*. Addition and scalar multiplication are defined in V *componentwise*, that is, if

$$u = (a_1, a_2, \ldots) \quad \text{and} \quad v = (b_1, b_2, \ldots)$$

then $$u + v = (a_1 + b_1,\ a_2 + b_2,\ \ldots) \quad \text{and} \quad ku = (ka_1, ka_2, \ldots)$$

An *inner product* is defined in v by

$$\langle u, v\rangle = a_1b_1 + a_2b_2 + a_3b_3 + \ldots$$

The above sum converges absolutely for any pair of points in V. Hence, the *inner product* is well defined. This inner product space is called l_2-*space* or *Hilbert space*.

2. Cauchy–Schwarz Inequality, Applications

The following formula is called the *Cauchy–Schwarz inequality* or *Schwarz inequality*. It is used in many branches of mathematics.

Theorem : (Cauchy–Schwarz) For any vectors u and v in an inner product space V,

$$\langle u, v\rangle^2 \le \langle u, u\rangle\langle v, v\rangle \qquad \text{or} \qquad |\langle u, v\rangle| \le \|u\|\|v\|$$

Next we examine this inequality in specific cases.

Example

(a) Consider any real numbers $a_1, \ldots, a_n$, $b_1, \ldots, b_n$. Then, by the Cauchy–Schwarz inequality,

$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \le (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

That is, $(u \cdot v)^2 \le \|u\|^2\|v\|^2$, where $u = (a_i)$ and $v = (b_i)$.

(b) Let f and g be continuous functions on the unit interval $[0, 1]$. Then, by the Cauchy–Schwarz inequality,

$$\left[\int_0^1 f(t)g(t)\,dt\right]^2 \le \int_0^1 f^2(t)\,dt \int_0^1 g^2(t)\,dt$$

That is, $(\langle f, g\rangle)^2 \le \|f\|^2\|v\|^2$. Here V is the inner product space $C[0, 1]$.

The next theorem gives the basic properties of a norm. The proof of the third property requires the Cauchy–Schwarz inequality.

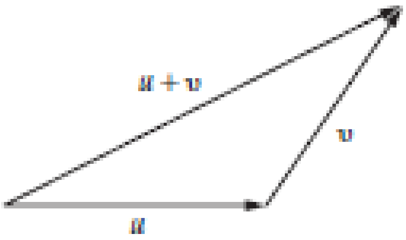
Theorem : Let *V* be an inner product space. Then the norm in *V* satisfies the following properties:

$[N_1]$ $\|v\| \ge 0$; and $\|v\| = 0$ if and only if $v = 0$.

$[N_2]$ $\|kv\| = |k|\|v\|$.

$[N_3]$ $\|u + v\| \le \|u\| + \|v\|$.

The property $[N_3]$ is called the *triangle inequality*, because if we view $u + v$ as the side of the triangle formed with sides u and v (as shown in Figure), then $[N_3]$ states that the length of one side of a triangle cannot be greater than the sum of the lengths of the other two sides.



Triangle Inequality

Angle Between Vectors

For any nonzero vectors *u* and *v* in an inner product space *V*, the *angle* between *u* and *v* is defined to be the angle θ such that $0 \le \theta \le \pi$ and

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}$$

By the Cauchy–Schwartz inequality, $-1 \le \cos\theta \le 1$, and so the angle exists and is unique.

Example

(a) Consider vectors $u = (2, 3, 5)$ and $v = (1, -4, 3)$ in $\mathbf{R}^3$. Then

$$\langle u, v \rangle = 2 - 12 + 15 = 5, \qquad \|u\| = \sqrt{4 + 9 + 25} = \sqrt{38}, \qquad \|v\| = \sqrt{1 + 16 + 9} = \sqrt{26}$$

Then the angle θ between u and v is given by

$$\cos\theta = \frac{5}{\sqrt{38}\sqrt{26}}$$

Note that θ is an acute angle, because $\cos\theta$ is positive.

(b) Let $f(t) = 3t - 5$ and $g(t) = t^2$ in the polynomial space $\mathbf{P}(t)$ with inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$. By Example 7.3,

$$\langle f, g \rangle = -\tfrac{11}{12}, \qquad \|f\| = \sqrt{13}, \qquad \|g\| = \tfrac{1}{5}\sqrt{5}$$

Then the "angle" θ between f and g is given by

$$\cos\theta = \frac{-\frac{11}{12}}{(\sqrt{13})(\frac{1}{5}\sqrt{5})} = -\frac{55}{12\sqrt{13}\sqrt{5}}$$

Note that θ is an obtuse angle, because $\cos\theta$ is negative.

3. Orthogonality

Let V be an inner product space. The vectors $u, v \in V$ are said to be *orthogonal* and u is said to be *orthogonal to v* if

$$\langle u, v \rangle = 0$$

The relation is clearly *symmetric*—if u is *orthogonal to* v, then $\langle v, u \rangle = 0$, and so v is *orthogonal to* u. We note that $0 \in V$ is orthogonal to every $v \in V$, because

$$\langle 0, v \rangle = \langle 0v, v \rangle = 0\langle v, v \rangle = 0.$$

Conversely, if u is orthogonal to every $v \in V$, then $\langle u, u \rangle = 0$ and hence $u = 0$ by $[I_3]$. Observe that u and v are orthogonal if and only if $\cos\theta = 0$, where θ is the angle between u and v. Also, this is true if and only if u and v are "*perpendicular*"—that is, $\theta = \pi/2$ (or $\theta = 90$).

Example

(a) Consider the vectors $u = (1, 1, 1)$, $v = (1, 2, -3)$, $w = (1, -4, 3)$ in $\mathbf{R}^3$. Then

$$\langle u, v \rangle = 1 + 2 - 3 = 0, \qquad \langle u, w \rangle = 1 - 4 + 3 = 0, \qquad \langle v, w \rangle = 1 - 8 - 9 = -16$$

Thus, u is orthogonal to v and w, but v and w are not orthogonal.

(b) Consider the functions $\sin t$ and $\cos t$ in the vector space $C[-\pi, \pi]$ of continuous functions on the closed interval $[-\pi, \pi]$. Then

$$\langle \sin t, \cos t \rangle = \int_{-\pi}^{\pi} \sin t \cos t\, dt = \tfrac{1}{2} \sin^2 t\big|_{-\pi}^{\pi} = 0 - 0 = 0$$

Thus, $\sin t$ and $\cos t$ are orthogonal functions in the vector space $C[-\pi, \pi]$.

Remark A vector $w = (x_1; x_2; \ldots; x_n)$ is *orthogonal to* $u = (a_1; a_2; \ldots; a_n)$ in $\mathbf{R}^n$ if

$$\langle u, w\rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

That is, w is *orthogonal to* u if w satisfies a homogeneous equation whose coefficients are the elements of u.

Example Find a nonzero vector w that is orthogonal to $u_1 = (1; 2; 1)$ and $u_2 = (2; 5; 4)$ in $\mathbf{R}^3$.

Let $w = (x, y, z)$. Then we want $\langle u_1, w\rangle = 0$ and $\langle u_2, w\rangle = 0$. This yields the homogeneous system

$$\begin{aligned} x + 2y + z &= 0 \\ 2x + 5y + 4z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y + z &= 0 \\ y + 2z &= 0 \end{aligned}$$

Here z is the only free variable in the echelon system. Set $z = 1$ to obtain $y = -2$ and $x = 3$. Thus, $w = (3, -2, 1)$ is a desired nonzero vector orthogonal to u_1 and u_2.

Any multiple of w will also be orthogonal to u_1 and u_2. Normalizing w, we obtain the following unit vector orthogonal to u_1 and u_2:

$$\hat{w} = \frac{w}{\|w\|} = \left(\frac{3}{\sqrt{14}}, -\frac{2}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right)$$

Orthogonal Complements

Let S be a subset of an inner product space V. The *orthogonal complement* of S, denoted by $S^\perp$ (read "*S perpendicular*") consists of those vectors in V that are orthogonal to every vector $u \in S$, that is,

$$S^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ for every } u \in S\}$$

In particular, for a given vector u in V, we have

$$u^\perp = \{v \in V : \langle v, u\rangle = 0\}$$

that is, $u^\perp$ consists of all vectors in V that are orthogonal to the given vector u.

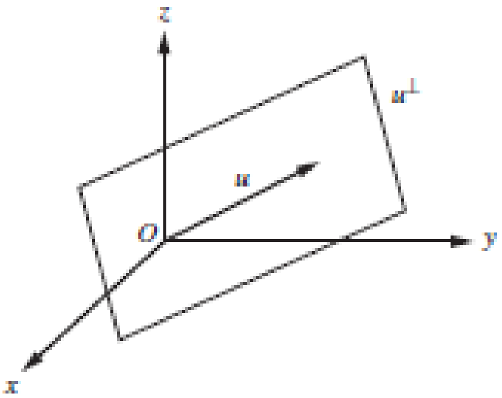
We show that $S^\perp$ is a subspace of V. Clearly $0 \in S^\perp$, because 0 is orthogonal to every vector in V. Now suppose $v, w \in S^\perp$. Then, for any scalars a and b and any vector $u \in S$, we have

$$\langle av + bw, \ u\rangle = a\langle v, u\rangle + b\langle w, u\rangle = a \cdot 0 + b \cdot 0 = 0$$

Thus, $av + bw \in S^\perp$, and therefore $S^\perp$ is a subspace of V.

<u>We state this result formally</u>: Let S be a *subset* of a vector space V. Then $S^\perp$ is a subspace of V.

Remark **1**: Suppose u is a nonzero vector in $\mathbf{R}^3$. Then there is a geometrical description of $u^\perp$. Specifically, $u^\perp$ is *the plane* in $\mathbf{R}^3$ through the origin O and perpendicular to the vector u. This is shown in Figure:



Orthogonal Complement $u^\perp$

***Remark* 2**: Let W be the solution space of an $m \times n$ homogeneous system $AX = 0$, where $A = [a_{ij}]$ and $X = [x_i]$. Recall that W may be viewed as the kernel of the linear mapping A: $\mathbf{R}^n \to \mathbf{R}^m$. Now we can give another interpretation of W using the notion of orthogonality.

Specifically, each solution vector $w = (x_1; x_2; \dots ; x_n)$ is orthogonal to each row of A, hence, W is the orthogonal complement of the row space of A.

Example Find a basis for the subspace $u^\perp$ of $\mathbf{R}^3$, where $u = (1; 3;-4)$

Note that $u^\perp$ consists of all vectors $w = (x, y, z)$ such that $\langle u, w \rangle = 0$, or $x + 3y - 4z = 0$. The free variables are y and z.

(1) Set $y = 1$, $z = 0$ to obtain the solution $w_1 = (-3, 1, 0)$.

(2) Set $y = 0$, $z = 1$ to obtain the solution $w_2 = (4, 0, 1)$.

The vectors w_1 and w_2 form a basis for the solution space of the equation, and hence a basis for $u^\perp$.

Suppose W is a subspace of V. Then both W and $W^\perp$ are subspaces of V. The next theorem, is a basic result in linear algebra.

Theorem : Let W be a subspace of V. Then V is the direct sum of W and $W^\perp$, that is, $V = W \oplus W^\perp$