Inner Product and Orthogonality

P. Sam Johnson

October 3, 2014

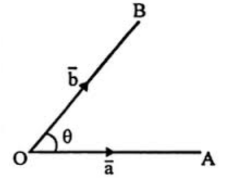
Overview

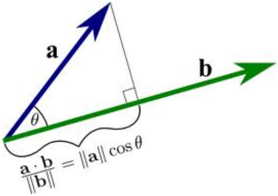
In the Euclidean space $\mathbb{R}^2$ and $\mathbb{R}^3$ there are two concepts, viz., length (or distance) and angle which have no analogues over a general field.

Fortunately there is a single concept usually known as inner product or scalar product which covers both the concepts of length and angle.

We discuss the concept of orthogonality and some applications.

Scalar product of vectors in $\mathbb{R}^2$





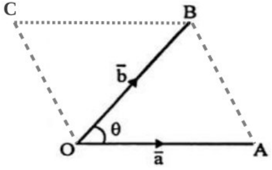
Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be vectors in $\mathbb{R}^2$ represented by the points A and B as in figure. Then the **scalar product** of a and b is defined to be

$$\langle a, b \rangle = \|a\| \; \|b\| \cos\theta$$

where $\|a\|$ is the length of OA, $\|b\|$ is the length of OB and θ is the angle between OA and OB.

Length, distance, angle : in terms of the inner product

Scalar product gives the following concepts of length, distance and angle.



1. **Length of a vector:** The length OA can be defined in terms of the scalar product since
$$OA^2 = \|a\|^2 = \langle a, a \rangle.$$
2. **Distance between vectors:** If $OABC$ is a parallelogram, the distance $AB = OC = \sqrt{\langle b - a, b - a \rangle}$ since $C = b - a$.
3. The **angle** θ can be obtained as
$$\theta = \cos^{-1}\left(\frac{\langle a, b \rangle}{\sqrt{\langle a, a \rangle . \langle b, b \rangle}}\right).$$

The above concepts and results have obvious analogues in $\mathbb{R}^3$. The concept of angle between vectors is generalized to "vector space with an inner product" (called **inner product space**).

Motivated by the scalar product (dot product) on $\mathbb{R}^2$ we now give the axiomatic definition of inner product on a vector space over $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$.

Definition

Let X be a vector space over $\mathbb{K}$. A function $\langle .,. \rangle : X \times X \to \mathbb{K}$ is an **inner product** *on X if for any $x, y, z \in X$ and $\alpha, \beta \in \mathbb{K}$ the following conditions are satisfied:*

1. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ *(linear with respect to first variable)*
2. $\langle x, x \rangle \geq 0$ *(positivity) and* $\langle x, x \rangle = 0 \iff x = 0$
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$. *(conjugate symmetry, often known as Hermitian symmetry) (the bar denotes the complex conjugate).*

One does not extend inner product to vector spaces over a general field mainly because $\langle x, x \rangle \geq 0$ has no meaning in a general field.

a vector space with an inner product	an **inner product space**
a real inner product space	a **Euclidean space**
a complex inner product space	a **unitary space**

Properties of an inner product

1. The restriction of an inner product to a subspace is an inner product.
2. In any inner product space, we have
   - $\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle$.
   - $\langle 0, y\rangle = \langle x, 0\rangle = 0$.
3. When the second argument is held fixed, **inner product is linear in the first argument**. Similarly, when the first argument is held fixed, **inner product is conjugate-linear in the second argument**.

Concept of length : Norm

Inner product combines the concepts of length and angle. We shall discuss the first concept, length.

Definition

A **norm** *on a (real or complex) vector space* V *is a map* $x \mapsto \|x\|$ *from* V *to* $\mathbb{R}$ *satisfying the following three conditions:*

1. $\|x\| \geq 0$; $x = 0$ *if* $\|x\| = 0$
2. $\|\alpha x\| = |\alpha|.\|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$.

A vector space together with a norm on it is called a **normed vector space** or **normed linear space** or simply **normed space**.

Each inner product induces a norm, defined by $\|x\| = \sqrt{\langle x, x\rangle}$.

Theorem

Every inner product space is a normed space.

In any inner-product space, we have the following

1. **length** of the vector x,

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

2. **distance** of two vectors x and y,

$$\|x - y\| = \sqrt{\langle x - y, x - y\rangle}.$$

Another concept : Angle

The **angle** between two nonzero vectors x and y is defined by the formula

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

It is understood that the angle θ should be chosen in the closed interval $[0, \pi]$.

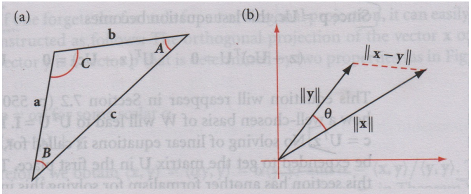
Inner product combines the concepts of length and angle. We shall discuss an important special case of the second concept, viz., the angle between two vectors being 90°.

Law of Cosines : Verification in $\mathbb{R}^2$

We call upon the **Law of Cosines** from trigonometry, which asserts that in a triangle having sides a, b, c, and opposing angles, A, B, C, the formula

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

holds.



Create a triangle having sides x, y, and $x - y$. Then in the law of cosines let $C = \theta, a = \|x\|, b = \|y\|$, and $c = \|x - y\|$. This produces the equation

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \cdot \|y\| \cos\theta.$$

Hence, we obtain

$$\|x\|^2 - 2\langle x, y\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \cdot \|y\| \cos\theta.$$

When this equation is simplified, we arrive at the equation

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} = \frac{\langle x, y\rangle}{\sqrt{\langle x, x\rangle} \cdot \sqrt{\langle y, y\rangle}}.$$

Theorem (Cauchy-Schwarz Inequality)

Let X be an inner product space. Then

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle \quad \text{for all } x, y \in X.$$

The equality occurs iff x and y are linearly dependent.

Examples of Inner Product Spaces

1. The space $\mathbb{K}^n$ of ordered n-tuples $(x_1, x_2, \ldots, x_n)$ of (real or complex) scalars is an inner product space with respect to the inner product (**canonical inner product**)

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

2. The space ℓ_2 of all sequences $(x_n)_{n=1}^{\infty}$ of (real or complex) scalars such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$, is an inner product space with the inner product defined by

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

3. Fix any finite subset A of $\mathbb{R}$ with size $\geq n$. Let $V = \mathcal{P}_n$ over $\mathbb{R}$.

$$\langle p, q\rangle := \sum_{a\in A} p(a)q(a)$$

is an inner product on V.

Examples of Inner Product Spaces

1. The space $C[a, b]$ of all continuous scalar-valued functions on the interval $[a, b]$ is an inner product space with the inner product defined by
$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}dx.$$
2. If $h \in V$ is such that $h(t) > 0$ for all $t \in [a, b]$,
$$\langle f, g\rangle = \int_a^b h(t)f(t)\ g(t)\ dt$$
is also an inner product.
3. $\langle A, B\rangle = tr(B^*A)$ is an inner product on $\mathbb{C}^{m\times n}$.
4. Let V be the vector space of all real-valued random variables with mean 0 and finite variance, defined on a fixed probability space. Let $F = \mathbb{R}$ and define $\langle x, y\rangle$ to be the covariance between x and y.

Exercises

1. Prove that the following

$$\langle x, y\rangle = y^T x \quad \text{and} \quad \langle x, y\rangle = x^* y$$

   are **not** inner products on $\mathbb{C}^n$.

2. In $\mathbb{C}^{m\times n}$, verify that

$$\langle A, B\rangle = \sum_{i=1}^{n} a_{ii}\overline{b}_{ii}$$

   is **not** an inner product.
   What are all the axioms which are violated?

Inner product associated with a matrix

Let V be an inner product space over $\mathbb{K}$ and $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ a basis of V. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_n)^T$ be the coordinate vectors of x and y respectively with respect to $\mathcal{B}$ and let $A = (a_{ij})$, where $a_{ij} = \langle u_j, u_i \rangle$. Then

$$\langle x, y \rangle := \left\langle \sum \alpha_i u_i, \sum \beta_j u_j \right\rangle = \sum \sum \overline{\beta_j} a_{ji} \alpha_i = \beta^* A \alpha. \tag{1}$$

The matrix A will satisfy the following conditions:

1. $A = A^*$
2. $\alpha^* A \alpha \geq 0$ for all $\alpha \in \mathbb{K}^n$,
3. if $\alpha^* A \alpha = 0$ then $\alpha = 0$.

Matrix associated with an inner product

Conversely, if A is a matrix satisfying the above three conditions, then $\langle ., . \rangle$ defined by (1) is an inner product on V.

Suppose $A = B^*B$, where B is a matrix with n columns and rank n. Then

$$\langle x, y \rangle = y^* B^* B x$$

is an inner product because

- $\overline{\langle y, x \rangle} = \overline{x^* B^* B y} = (x^* B^* B y)^* = y^* B^* B x = \langle x, y \rangle$.
- $\langle x, x \rangle = (Bx)^*(Bx) \geq 0$.
- If $\langle x, x \rangle = 0$ then $Bx = 0$ and so $x = 0$.

We shall later show that any matrix A satisfying the above three conditions, can be written as B^*B for some non-singular B.

Orthogonality

Let V be an inner product space, $x, y \in V$. Let A, B be subsets of V.

$\langle x, y \rangle = 0$ (we write $x \perp$ y)	x and y are **orthogonal** to each other
$x \perp y$ for every pair of distinct vectors x, y in A	A is **orthogonal**
A is orthogonal and every vector in A has norm 1	A is **orthonormal**
every vector in A is orthogonal to every vector in B	A is **orthogonal** to B

Properties of Orthogonality

- $x \perp y \iff y \perp x$.
- $0 \perp x$ for all x.
- $x \perp x \iff x = 0$.
- if $x \perp y, y \perp z$, then $x \perp (\alpha y + \beta z)$ for any $\alpha, \beta \in \mathbb{K}$.
- The empty set is orthonormal (in a vacuous sense).

Exercises

A set of vectors is orthogonal iff its elements are pair-wise orthogonal. **Is the corresponding statement for linear independence true?**.

Linear independence is a property of the entire set whereas orthogonality is a property of pairs.

Exercises

- *Any orthogonal set A not containing the null vector is linearly independent.*
- *Any orthonormal set is linearly independent.*
- *If the subspaces $S_1, S_2, \ldots, S_k$ are orthogonal to one another then $S_1 + S_2 + \cdots + S_k$ is direct.*

Orthogonal Complement

Definition

The **orthogonal complement** *of a set S in an inner-product space is the set*

$$\{x : x \perp s \text{ for all } s \in S\}.$$

The orthogonal complement of S is denoted by $S^{\perp}$.

Theorem

In an inner product space, the orthogonal complement of any subset is a subspace.

Theorem

In an inner product space, the orthogonal complement of a set is the same as the orthogonal complement of its span.

Theorem

If M is a subspace in an n-dimensional inner product space, then

$$dim(M) + dim(M^\perp) = n.$$

Theorem

Let M and N be subsets of an inner product space. If $M \subseteq N$, then $N^\perp \subseteq M^\perp$.

Theorem

If M is a finite dimensional subspace in an inner product space V, then V is the direct sum of M and $M^\perp$:

$$V = M + M^\perp \quad \text{and} \quad M \cap M^\perp = \{0\}.$$

Moreover, $M = M^{\perp\perp}$.

Pythagoras theorem

In a **real inner product space**, if $x \perp y$, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

The converse is true for real inner product space but not for complex inner product space.

More generally,

$$\Big\| \sum_{i=1}^{k} x_i \Big\|^2 = \sum_{i=1}^{k} \|x_i\|^2$$

if $\{x_1, x_2, \ldots, x_k\}$ is orthogonal. The converse is not true for both real and complex inner product spaces.

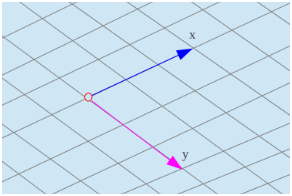
System corresponding to orthonormal basis

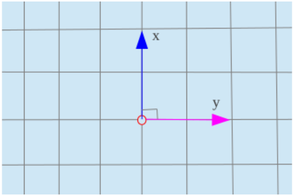
Definition

Let S be a subspace of an inner product space. We say that B is an **orthogonal basis** *(resp. an* **orthonormal basis***) of S if B is a basis of S and B is an orthonormal (resp. an orthonormal) set.*

We have seen that a basis corresponds to a **coordinate system**.

An orthonormal basis corresponds to a **system of rectangular coordinates** where the reference point on each axis is at unit distance from the origin.





For a given orthonormal basis, finding the coordinates with respect to such a coordinate system is easy as shown in the following.

Theorem

Let $B = \{x_1, x_2, \ldots, x_n\}$ be an orthonormal basis of an inner product space V. Then for any $x \in V$, we have

$$x = \sum_{j=1}^{n} \alpha_j x_j, \text{ where } \alpha_j = \langle x, x_j \rangle.$$

Exercises

Let $x_1, x_2, \ldots, x_k$ form an orthonormal set.

1. Show that $\| \sum_{i=1}^{k} \alpha_i x_i \|^2 = \sum_{i=1}^{k} \|\alpha_i\|^2$.
2. If z is the residual of x on $\{x_1, x_2, \ldots, x_k\}$, show that

$$\|z\|^2 = \|x\|^2 - \left\| \sum_{i=1}^{k} \langle x, x_i \rangle x_i \right\|^2 = \|x\|^2 - \sum_{i=1}^{k} |\langle x, x_i \rangle|^2.$$

3. **Bessel's inequality:**

$$\|x\|^2 \geq \sum_{i=1}^{k} |\langle x, x_i \rangle|^2$$

for any x. Show also that equality holds iff $x \in Sp(\{x_1, x_2, \ldots, x_k\})$.

Exercises

Let $B = \{x_1, x_2, \ldots, x_k\}$ be an orthonormal set in a finite-dimensional inner product space V. Show that the following statements are equivalent:

1. B is an orthonormal basis (maximal),
2. $\langle x, x_i \rangle = 0$ for $i = 1, 2, \ldots, k \Rightarrow x = 0$,
3. B generates V,
4. if $x \in V$ then $x = \sum_{i=1}^{k} \langle x, x_i \rangle x_i$,
5. if $x, y \in V$ then $\langle x, y \rangle = \sum_{i=1}^{k} \langle x, x_i \rangle . \langle x_i, y \rangle$,
6. if $x \in V$ then $\|x\|^2 = \sum_{i=1}^{k} |\langle x, x_i \rangle|^2$.

Orthogonal Projection

One vector x can be projected orthogonally onto another vector y, provided that y is not zero.

The idea is that the **projection** of x onto y should be a scalar multiple of y, say αy, such that $x - \alpha y$ is orthogonal to y.

What is the correct value of α?

Theorem

In any inner-product space, the orthogonal projection of a vector x onto a nonzero vector y is the point

$$p = \frac{\langle x, y \rangle}{\langle y, y \rangle} y.$$

It has the property that $x - p$ is orthogonal to y. Thus, x is split into an orthogonal pair of vectors in the equation $x = p + (x - p)$.

Construction of orthogonal projection matrix

Notice that our concept of projecting x onto y does not depend on the magnitude of the vector y. Actually the formula can be remembered more easily as

$$p = \langle x, y \rangle v$$

where v is the normalized y; that is, $y/\|y\|$.

The **calculation of an orthogonal projection** can be carried out in several different ways. For example, we can begin with the point z that is to be projected and the matrix U whose columns are the vectors u_i.

The point p that we seek is a linear combination of the columns of U and is therefore of the form

$$p = Uc$$

for some unknown vector c in $\mathbb{R}^n$.

The orthogonality condition is that $z - p$ should be orthogonal to all the columns of U, or in other terms,

$$(z - p)^T U = 0$$

Since $p = Uc$, this last equation becomes

$$\begin{aligned} (z - Uc)^T U &= 0 \\ U^T(z - Uc) &= 0 \\ U^T Uc &= U^T z. \end{aligned}$$

We shall see later how a well-chosen basis of W will lead to $U^T U = I$.

In this case we obtain $c = U^T z$.

Suppose we are working in the space $\mathbb{R}^n$, and we have an orthonormal set of n vectors, $u_1, u_2, \ldots, u_n$. Put them into a matrix U as columns.

The resulting matrix is square, and this property is crucial. The orthonormality now gives us the equation $U^T U = I$. Such a matrix U is said to be **orthogonal**.

It is obviously invertible as U^T is its inverse. Since U is square, $UU^T = I$, that the rows of U also form an orthonormal set of vectors! This is an impressive bit of magic.

Definition

A real matrix U is **orthogonal** *if $UU^T = U^T U = I$.*
A complex matrix U is **unitary** *if $UU^H = U^H U = I$.*

Theorem

Let $\{u_1, u_2, \ldots, u_n\}$ *be an orthonormal basis for a subspace* U *in an inner-product space. The orthogonal projection of any* x *onto* U *is the point*

$$p = \sum_{i=1}^{n} \langle x, u_i \rangle u_i.$$

Theorem

In order that a vector be orthogonal to a subspace (in an inner-product space), it is sufficient that the vector be orthogonal to each member of a set that spans the subspace.

Exercises

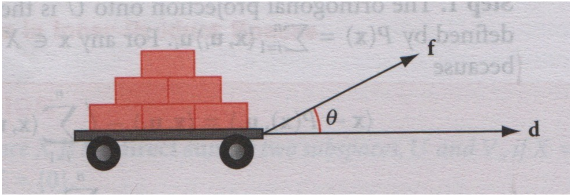
1. Let z be a fixed nonnull vector in the plane. What is the locus of the point x such that $\langle x, z \rangle = 0$? What happens if 0 is replaced by a non-zero scalars?
2. If x_1, x_2, y_1, y_2 are real numbers, show that
$$(x_1x_2 + y_1y_2)^2 \leq (x_1^2 + y_1^2)(x_2^2 + y_2^2).$$
Hence deduce that $PQ + QR \geq PR$ for any three points P, Q and R in the plane.
3. Suppose $A = \{x_1, x_2, \ldots, x_k\}$ be an orthogonal set (not a basis) of non-null vectors in V. Then for any $x \in V$, we call
$$z := x - \sum_{j=1}^{k} \frac{\langle x, x_j \rangle}{\langle x_j, x_j \rangle} x_j$$
the **residual of** x **with respect to** A. Prove that the residual z is orthogonal to each x_i.

Application : Work and Forces

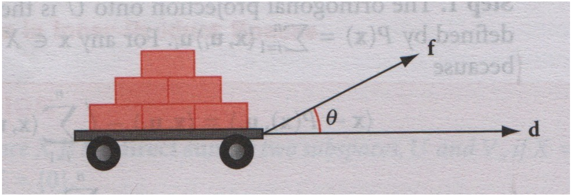
Let the vector f be the force exerted on an object, let the vector d be the displacement caused by the force, and let θ be the angle between f and d.

For example, suppose we are pulling a heavy load on a dolly with a constant force so that it moves horizontally along the ground.

The work done in moving the dolly through a distance d is given by the distance moved multiplied by the magnitude of the component of the force in the direction of motion.



Application : Work and Forces



The component of f in the direction d is

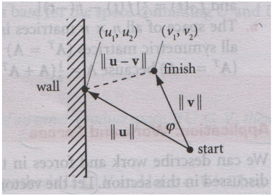
$$\|f\| \cos\theta.$$

By the definition, the work accomplished is

$$W = \|f\| \cdot \|d\| \cos\theta = \langle f, d \rangle.$$

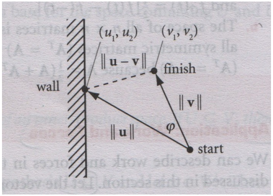
Application : Collision

The law of cosines can be applied to determine the final location of a ball after a glancing collision with a wall, as shown in the following figure.



Let $u = (u_1, u_2)$ be the initial position, $v = (v_1, v_2)$ be the final position, and $u - v$ be the change in position as a result of the collision. From the Law of Cosines, it follows that the magnitude of the change in position is:

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\| \cdot \|v\| \cos \varphi.$$



From the expression

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\| \cdot \|v\| \cos \varphi,$$

we can obtain

$$\langle u, v \rangle = \|u\| \cdot \|v\| \cos \varphi$$

which gives a connection between the inner product of the vectors u and v and the angle φ between them.

References

- Gilbert Strang, "*Linear Algebra and its Applications*", Cengage Learning, New Delhi, 2006.
- A. Ramachandra Rao and P. Bhimasankaram, "*Linear Algebra*", Hindustan Book Agency, 2000.
- Ward Cheney and David Kincaid, "*Linear Algebra - Theory and Applications*", Jones & Bartlett, 2010.