1. Inner product

1.1. Inner product.

1.1.1. *Inner product on real spaces.* Vectors in $\mathbb{R}^n$ have more properties than the ones listed in the definition of vector spaces: we can define their length, and the angle between two vectors.

Recall that two vectors are orthogonal if and only if their dot product is zero, and, more generally, the cosine of the angle between two *unit* vectors in $\mathbb{R}^3$ is their dot product. The notion of *inner product* extracts the essential properties of the dot product, while allowing it to be defined on quite general vector spaces. We will first define it for real vector spaces, and then we will formulate it for complex ones.

Definition 1. *An* **inner product** *on vector space V over $F = \mathbb{R}$ is an operation which associate to two vectors $\mathbf{x}, \mathbf{y} \in V$ a scalar $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{R}$ that satisfies the following properties:*
(i) its is positive definite: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = 0$,
(ii) it is symmetric: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
(iii) it is linear in the second argument: $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$ and $\langle \mathbf{x}, c\mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$

Note that by symmetry it follows that an inner product is linear in the first argument as well: $\langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle$ and $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$.

A function of two variables which is linear in one variable and linear in the other variable is called *bilinear*; hence, the inner product in a real vector space is bilinear.

Example 1. The dot product in $\mathbb{R}^2$ or $\mathbb{R}^3$ is clearly an inner product: if $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ then define

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

Example 2. More generally, an inner product on $\mathbb{R}^n$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \ldots + x_n y_n \tag{1}$$

Example 3. Here is another inner product on $\mathbb{R}^3$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = 5x_1 y_1 + 10x_2 y_2 + 2x_3 y_3$$

(some directions are weighted more than others).

Example 4. On spaces of functions the most useful inner products use integration. For example, consider $C[a, b]$ be the linear space of functions continuous on $[a, b]$. Then

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

is an inner product on $C[a, b]$ (check!).

Example 5. Sometimes a weight is useful: let $w(t)$ be a positive function. Then

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

is also an inner product on $C[a,b]$ (check!).

1.1.2. *Inner product on complex spaces.* For complex vector spaces extra care is needed. The blueprint of the construction here can be seen on the simplest case, $\mathbb{C}$ as a one dimensional vector space over $\mathbb{C}$: the inner product of $\langle z,z\rangle$ needs to be a positive number! It makes sense to define $\langle z,z\rangle = \overline{z}z$.

Definition 2. *An* **inner product** *on vector space V over $F=\mathbb{C}$ is an operation which associate to two vectors $\mathbf{x},\mathbf{y}\in V$ a scalar $\langle \mathbf{x},\mathbf{y}\rangle \in \mathbb{C}$ that satisfies the following properties:*
(i) it is positive definite: $\langle \mathbf{x},\mathbf{x}\rangle \geq 0$ and $\langle \mathbf{x},\mathbf{x}\rangle = 0$ if and only if $\mathbf{x}=0$,
(ii) it is linear in the second argument: $\langle \mathbf{x},\mathbf{y}+\mathbf{z}\rangle = \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{x},\mathbf{z}\rangle$ and $\langle \mathbf{x},c\mathbf{y}\rangle = c\langle \mathbf{x},\mathbf{y}\rangle$
(iii) it is conjugate symmetric*: $\langle \mathbf{x},\mathbf{y}\rangle = \overline{\langle \mathbf{y},\mathbf{x}\rangle}$.*

Note that conjugate symmetry combined with linearity implies that $\langle .,.\rangle$ is *conjugate-linear in the first variable*: $\langle \mathbf{x}+\mathbf{z},\mathbf{y}\rangle = \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{z},\mathbf{y}\rangle$ and $\langle c\mathbf{x},\mathbf{y}\rangle = \overline{c}\langle \mathbf{x},\mathbf{y}\rangle$.

A function of two variables which is linear in one variable and conjugate-linear in the other variable is called *sesquilinear*; the inner product in a complex vector space is sesquilinear.

Please keep in mind that most mathematical books use inner product linear in the first variable, and conjugate linear in the second one. You should make sure you know the convention used by each author.

Example 1'. On the one dimensional complex vector space $\mathbb{C}$ an inner product is $\langle z,w\rangle = \overline{z}w$.

Example 2'. More generally, an inner product on $\mathbb{C}^n$ is

$$\langle \mathbf{x},\mathbf{y}\rangle = \overline{x_1}y_1 + \ldots + \overline{x_n}y_n \tag{2}$$

Example 3'. Here is another inner product on $\mathbb{C}^3$:

$$\langle \mathbf{x},\mathbf{y}\rangle = 5\overline{x_1}y_1 + 10\overline{x_2}y_2 + 2\overline{x_3}y_3$$

(some directions are weighted more than others).

Example 4'. Let $C([a,b],\mathbb{C})$ be the linear space of complex-valued functions which are continuous on $[a,b]$.[1] Then

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

is an inner product on $C([a,b],\mathbb{C})$ (check!).

[1] A complex valued function $f(t)=u(t)+iv(t)$ is continuous if the $\mathbb{R}^2$-valued function $(u(t),v(t))$ is continuous.

Example 5'. Weights need to be positive: for $w(t)$ a given positive function. Then

$$\langle f,g\rangle = \int_a^b w(t)\overline{f(t)}\,g(t)\,dt$$

is also an inner product on $C([a,b],\mathbb{C})$ (check!).

1.2. **Inner product spaces.**

Definition 3. *A vector space V equipped with an inner product $(V,\langle .,.\rangle)$ is called an* **inner product space**.

Examples 1.-5. before are examples of inner product spaces over $\mathbb{R}$, while Examples 1'.-5'. are inner product spaces over $\mathbb{C}$.

In an inner product space we can do geometry. First of all, we can define length of vectors:

Definition 4. *Let $(V,\langle .,.\rangle)$ be an inner product space. The quantity*

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

is called the **norm** *of the vector x. For $V=\mathbb{R}^3$ with the usual inner product (which is the dot product) the norm of a vector is its length.*

Vectors of norm one, $\mathbf{x}$ with $\|\mathbf{x}\|=1$, are called **unit vectors**.

In an inner product space, the parallelogram law holds: "sum of the squares of the lengths of the four sides of a parallelogram equals the sum of the squares of the lengths of the two diagonals":

Proposition 5. *In an inner product space* the parallelogram law *holds: for any $\mathbf{x},\mathbf{y}\in V$*

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \tag{3}$$

The proof of (3) is a simple exercise, left to the reader.

In an inner product space we can define the angle between two vectors. Recall that in the usual Euclidian geometry in $\mathbb{R}^2$ or $\mathbb{R}^3$, the angle θ between two vectors $\mathbf{x},\mathbf{y}$ is calculated from $\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\,\cos\theta$. The existence of an angle θ with this property in any inner product space is guaranteed by the Cauchy-Schwartz inequality, one of the most useful, and deep, inequalities in mathematics, which holds in finite or infinite dimensional inner product spaces is:

Theorem 6. The Cauchy-Schwartz inequality

In an inner product vector space any two vectors $\mathbf{x},\mathbf{y}$ satisfy

$$|\langle\mathbf{x},\mathbf{y}\rangle| \le \|\mathbf{x}\|\,\|\mathbf{y}\| \tag{4}$$

with equality if and only if $\mathbf{x},\mathbf{y}$ are linearly dependent.

Remarks: 1. Recall that $\mathbf{x}, \mathbf{y}$ are linearly dependent means that one of the two vectors is $\mathbf{0}$ or the two vectors are scalar multiples of each other.

2. It suffices to prove inequality (4) in $Sp(\mathbf{x}, \mathbf{y})$, a space which is at most two-dimensional.

Proof of (4).

If $\mathbf{y} = \mathbf{0}$ then the inequality is trivially true. Otherwise, calculate, for any scalar c

$$0 \leq \langle \mathbf{x} - c\mathbf{y}, \mathbf{x} - c\mathbf{y} \rangle = \|\mathbf{x}\|^2 - 2\Re\left[c\langle \mathbf{x}, \mathbf{y} \rangle\right] + |c|^2 \|\mathbf{y}\|^2$$

and by choosing $c = \overline{\langle \mathbf{x}, \mathbf{y} \rangle} / \langle \mathbf{y}, \mathbf{y} \rangle$ we obtain

$$0 \leq \langle \mathbf{x}, \mathbf{x} \rangle - \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\langle \mathbf{y}, \mathbf{y} \rangle|}$$

which gives (4). Equality holds only for $\mathbf{x} - c\mathbf{y} = \mathbf{0}$, or $\mathbf{y} = \mathbf{0}$, therefore only when $\mathbf{x}, \mathbf{y}$ are linearly dependent. $\square$

In usual Euclidian geometry, the sides of triangle determines its angles. Similarly, in an inner product space, if we know the norm of vectors, then we know inner products. In other words, the inner product is completely recovered if we know the norm of every vector:

Theorem 7. The polarization identity*:*
In a real real inner space

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 \right)$$

In a complex inner space

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|^2 + i\|i\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 - i\| - i\mathbf{x} + \mathbf{y}\|^2 \right) = \frac{1}{4} \sum_{k=0}^{3} i^k \|i^k \mathbf{x} + \mathbf{y}\|^2$$

The proof is by a straightforward calculation.

2. Orthogonal bases

Let $(V, \langle ., . \rangle)$ be an inner product space.

Definition 8. *Two vectors* $\mathbf{x}, \mathbf{y} \in V$ *are called* **orthogonal** *if* $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

In this case we denote for short $\mathbf{x} \perp \mathbf{y}$ (and, of course, also $\mathbf{y} \perp \mathbf{x}$).

Note that the zero vector is orthogonal to any vector:

$$\langle \mathbf{x}, \mathbf{0} \rangle = 0 \quad \text{for all } \mathbf{x} \in V$$

since $\langle \mathbf{x}, \mathbf{0} \rangle = \langle \mathbf{x}, 0\mathbf{y} \rangle = 0 \langle \mathbf{x}, \mathbf{y} \rangle = 0$.

As in geometry:

Definition 9. *We say that a vector* $\mathbf{x} \in V$ *is orthogonal to a subset* S *go* V *if* $\mathbf{x}$ *is orthogonal to every vector in* S*:*

$$\mathbf{x} \perp S \quad \textit{if and only if} \quad \mathbf{x} \perp \mathbf{z} \ \textit{for all } \mathbf{z} \in S$$

Exercise. Show that $\mathbf{x} \perp Sp(\mathbf{y}_1, \ldots, \mathbf{y}_r)$ is and only if $\mathbf{x} \perp \mathbf{y}_j$ for all $j = 1, \ldots, r$.

Definition 10. *A set $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset V$ is called an* **orthogonal set** *if $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ for all $i \neq j$.*

The set is called **orthonormal** *if it is orthogonal and all $\mathbf{v}_j$ are unit vectors.*

Remark. *An orthogonal set which does not contain the zero vector is a linearly independent set*, since if $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ are orthogonal and $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$ then for any $j = 1, \ldots, k$

$$0 = \langle \mathbf{v}_j, c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k \rangle = c_1\langle \mathbf{v}_j, \mathbf{v}_1 \rangle + \ldots + c_j\langle \mathbf{v}_j, \mathbf{v}_j \rangle + \ldots + c_k\langle \mathbf{v}_j, \mathbf{v}_k \rangle$$

$$= c_j\langle \mathbf{v}_j, \mathbf{v}_j \rangle$$

and since $\mathbf{v}_j \neq \mathbf{0}$ then $\langle \mathbf{v}_j, \mathbf{v}_j \rangle \neq 0$ which implies $c_j = 0$. $\Box$.

Definition 11. *A basis for V with is an orthogonal set is called an* **orthogonal basis**.

An orthogonal basis made of unit vectors is called an **orthonormal basis**.

Of course, if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthogonal basis, then $\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}$, is an orthonormal basis.

For example, the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is an orthonormal basis in $\mathbb{R}^n$ (when equipped this the inner product given by the dot product).

Orthogonal bases make formulas simpler and calculations easier. As a first example, here is how coordinates of vectors are found:

Theorem 12. *a) Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthogonal basis for V.*

Then the coefficients of the expansion of any $\mathbf{x} \in V$ in the basis B are found as

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n \quad \text{where } c_j = \frac{\langle \mathbf{v}_j, \mathbf{x} \rangle}{\|\mathbf{v}_j\|^2} \tag{5}$$

b) In particular, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for V then

$$\mathbf{x} = x_1\mathbf{u}_1 + \ldots + x_n\mathbf{u}_n \quad \text{where } x_j = \langle \mathbf{u}_j, \mathbf{x} \rangle \tag{6}$$

Proof.

Consider the expansion of x in the basis B: $\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$. For each $j = 1, \ldots, n$

$$\langle \mathbf{v}_j, \mathbf{x} \rangle = \langle \mathbf{v}_j, c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n \rangle = c_1\langle \mathbf{v}_j, \mathbf{v}_1 \rangle + \ldots + c_j\langle \mathbf{v}_j, \mathbf{v}_j \rangle + \ldots + c_n\langle \mathbf{v}_j, \mathbf{v}_n \rangle$$

$$= c_j\langle \mathbf{v}_j, \mathbf{v}_j \rangle = c_j\|\mathbf{v}_j\|^2$$

which gives the formula (5) for c_j. $\Box$

When coordinates are given in an orthonormal basis, the inner product is the dot product of the coordinate vectors: