5. Inner Products, Norms, Distance and Angle

5.1 Note: In this section we shall be primarily interested in vector spaces over $\mathbb{R}$ or $\mathbb{C}$.

5.2 Definition: Recall that the set of complex numbers $\mathbb{C}$ is defined to be $\mathbb{C} = \mathbb{R}^2$. In $\mathbb{C}$ we write $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and for $x, y \in \mathbb{R}$ we write $x + iy = \begin{pmatrix} x \\ y \end{pmatrix}$. For $z = x + iy$ with $x, y \in \mathbb{R}$, we say that x and y are the **real** and **imaginary** parts of z, and write $x = \mathrm{Re}(z)$ and $y = \mathrm{Im}(z)$. For $z = x + iy$ and $w = u + iv$ with $x, y, u, v \in \mathbb{R}$, we define **addition** and **multiplication** by

$$z + w = (x + u) + i(y + v) \ , \ zw = (xu - yv) + i(xv + yu)$$

and we define the **conjugate** of z and the **length** (or **norm**) of z to be

$$\overline{z} = x - iy \ , \ |z| = \sqrt{z\,\overline{z}} = \sqrt{x^2 + y^2}.$$

These operations make $\mathbb{C}$ into a field. For $0 \neq z \in \mathbb{C}$, the **inverse** of z is given by

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

For a vector $z \in \mathbb{C}^n$, we can write $z = x + iy$ with $x, y \in \mathbb{R}^n$. We then define $\overline{z} = x - iy$ and we define $z^* = \overline{z}^T$. More generally, for $A \in M_{n\times l}(\mathbb{C})$ define the **adjoint** of A to be the matrix $A^* = \overline{A}^T \in M_{l\times n}(\mathbb{C})$, that is

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,l} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,l} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,l} \end{pmatrix}^* = \begin{pmatrix} \overline{a_{1,1}} & \overline{a_{2,1}} & \cdots & \overline{a_{n,1}} \\ \overline{a_{1,2}} & \overline{a_{2,2}} & \cdots & \overline{a_{n,2}} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1,l}} & \overline{a_{2,l}} & \cdots & \overline{a_{n,l}} \end{pmatrix}$$

5.3 Definition: There are several products in $\mathbb{C}^n$ analogous to the dot product on $\mathbb{R}^n$. For $z, w \in \mathbb{C}^n$ we define the (complex) **dot product** of z and w to be

$$z \cdot w = w^T z = \sum_{i=1}^n z_i w_i \in \mathbb{C}.$$

As second product can be defined as follows. Given $z = x + iy$ and $w = u + iv$ with $x, y, u, v \in \mathbb{R}^n$, we identify $\mathbb{C}^n$ with $\mathbb{R}^n$ by writing z and w as

$$z_{\mathbb{R}} = \big(x_1, y_1, x_2, y_2, \cdots, x_n, y_n\big)^T \in \mathbb{R}^{2n}$$
$$w_{\mathbb{R}} = \big(u_1, v_1, u_2, v_2, \cdots, u_n, v_n\big)^T \in \mathbb{R}^{2n},$$

and then we define the **real dot product** of z and w to be

$$z_{\mathbb{R}} \cdot w_{\mathbb{R}} = x \cdot u + y \cdot v = \sum_{i=1}^n (x_i u_i + y_i v_i) \in \mathbb{R}.$$

Finally, for $z, w \in \mathbb{C}^n$ we define a third product, called the **inner product** of z with w by

$$\langle z, w \rangle = w^* z = \sum_{i=1}^n z_i \overline{w_i} \in \mathbb{C}.$$

5.4 Remark: The latter two products can be used to define distance and angles in $\mathbb{C}^n$. Both give rise to the same definition of distance, but they give rise to different notions of orthogonality. They are related by $z_{\mathbb{R}} \bullet w_{\mathbb{R}} = \mathrm{Re}\big(\langle z, w\rangle\big)$. For the moment, we shall concentrate primarily on the third of these three products, namely the inner product.

5.5 Definition: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let W be a vector space over $\mathbb{F}$. An **inner product** over $\mathbb{F}$ is a function $\langle\ ,\ \rangle : W \times W \to \mathbb{F}$ (meaning that if $u, v \in W$ then $\langle u, v\rangle \in \mathbb{F}$) such that for all $u, v, w \in W$ and all $t \in \mathbb{F}$ we have

(1) (Sesquilinearity) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle\ ,\ \langle tu, v\rangle = t\,\langle u, v\rangle,$
$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle\ ,\ \langle u, tv\rangle = \overline{t}\,\langle u, v\rangle,$
(2) (Conjugate Symmetry) $\langle u, v\rangle = \overline{\langle v, u\rangle}$, and
(3) (Positive Definiteness) $\langle u, u\rangle \in \mathbb{R}$ and $\langle u, u\rangle \geq 0$ with $\langle u, u\rangle = 0 \iff u = 0$.

For $u, v \in W$, $\langle u, v\rangle$ is called the inner product of u with v. An inner product over $\mathbb{R}$ is called a **real inner product** and an inner product over $\mathbb{C}$ is called a **Hermitian inner product**. An **inner product space** over $\mathbb{F}$ is a vector space over $\mathbb{F}$ equipped with an inner product. Given two inner product spaces U and V over $\mathbb{F}$, a linear map $L : U \to V$ is called a **homomorphism** of inner product spaces (or we say that L **preserves inner product**) when $\big\langle L(x), L(y)\big\rangle = \langle x, y\rangle$ for all $x, y \in U$.

5.6 Definition: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let W be a vector space over $\mathbb{F}$. A **norm** on W is a map $\|\ \ \| : W \to \mathbb{R}$ such that for all $u, v \in W$ and all $t \in \mathbb{F}$ we have

(1) (Scaling) $\|tu\| = |t|\,\|u\|$,
(2) (Positive Definiteness) $\|u\| \geq 0$ with $\|u\| = 0 \iff u = 0$, and
(3) (Triangle Inequality) $\|u+v\| \leq \|u\| + \|v\|$.

For $u \in W$ the real number $\|u\|$ is called the **norm** (or **length**) of u, and we say that u is a **unit vector** when $\|u\| = 1$. A **normed linear space** over $\mathbb{F}$ is a vector space over $\mathbb{F}$ equipped with a norm. Given two normed linear spaces U and V over $\mathbb{F}$, a linear map $L : U \to V$ is called a **homomorphism** of normed linear spaces (or we say that L **preserves norm**) when $\big\|L(x)\big\| = \|x\|$ for all $x \in U$.

5.7 Definition: Let X be a set. A **metric** on X is a map $d : X \times X \to \mathbb{R}$ such that for all $a, b, c \in X$ we have

(1) (Symmetry) $d(a, b) = d(b, a)$,
(2) (Positive Definiteness) $d(a, b) \geq 0$ with $d(a, b) = 0 \iff a = b$, and
(3) (Triangle Inequality) $d(a, c) \leq d(a, b) + d(b, c)$.

For $a, b \in X$, $d(a, b)$ is called the **distance** between a and b. A **metric space** is a set which is equipped with a metric. Given two metric spaces X and Y, a map $F : X \to Y$ is called an **isometry** (or we say that F is **distance preserving**) when $d\big(F(a), F(b)\big) = d(a, a)$ for all $a, b \in X$.

5.8 Theorem: L*et W be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $u, v \in W$. Then if $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in W$, or if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in W$ then $u = v$.*

Proof: Suppose that $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in W$. Then $\langle x, u-v\rangle = \langle x, u\rangle - \langle x, v\rangle = 0$ for all $x \in W$. In particular, taking $x = u - v$ we have $\|u-v\|^2 = \langle u-v, u-v\rangle = 0$, and so $u = v$. Similarly, if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in W$ then $u = v$.

5.9 Theorem: *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let W be an inner product space over $\mathbb{F}$. For $u \in W$ define $\|u\| = \sqrt{\langle u, u\rangle}$. Then for all $u, v \in W$ and all $t \in \mathbb{F}$ we have*

(1) (Scaling) $\|tu\| = |t|\,\|u\|$,
(2) (Positive Definiteness) $\|u\| \geq 0$ with $\|u\| = 0 \iff u = 0$,
(3) $\|u+v\|^2 = \|u\|^2 + 2\,\mathrm{Re}\langle u, v\rangle + \|v\|^2$,
(4) (Polarization Identity) if $\mathbb{F} = \mathbb{R}$ then $\langle u, v\rangle = \frac{1}{4}\big(\|u+v\|^2 - \|u-v\|^2\big)$ and if $\mathbb{F} = \mathbb{C}$ then $\langle u, v\rangle = \frac{1}{4}\big(\|u+v\|^2 + i\|u+iv\|^2 - \|u-v\|^2 - i\|u-iv\|^2\big)$,
(5) (Cauchy-Schwarz Inequality) $|\langle u, v\rangle| \leq \|u\|\,\|v\|$ with $|\langle u, v\rangle| = \|u\|\,\|v\|$ if and only if $\{u, v\}$ is linearly dependent, and
(6) (Triangle Inequality) $\big|\|u\| - \|v\|\big| \leq \|u+v\| \leq \|u\| + \|v\|$.

In particular, $\|\ \ \|$ is a norm on W, which we call the norm **induced** *by the inner product.*

Proof: We only prove Part (5) and part of Part (6). To prove Cauchy's Inequality, suppose first that $\{u, v\}$ is linearly dependent. Then one of x and y is a multiple of the other, say $v = tu$ with $t \in \mathbb{F}$. Then $|\langle u, v\rangle| = |\langle u, tu\rangle| = \big|\overline{t}\,\langle u, u\rangle\big| = |t|\,\|u\|^2 = \|u\|\,\|tu\| = \|u\|\,\|v\|$.

Next we suppose that $\{u, v\}$ is linearly independent. Then $1 \cdot v + t \cdot u \neq 0$ for all $t \in \mathbb{F}$, so in particular $v - \frac{\langle v,u\rangle}{\|u\|^2}\,u \neq 0$. Thus we have

$$\begin{aligned} 0 < \left\|v - \tfrac{\langle v,u\rangle}{\|u\|^2}\,u\right\|^2 &= \left\langle v - \tfrac{\langle v,u\rangle}{\|u\|^2}\,u\,,\ v - \tfrac{\langle v,u\rangle}{\|u\|^2}\,u\right\rangle \\ &= \langle v, v\rangle - \overline{\tfrac{\langle v,u\rangle}{\|u\|^2}}\,\langle v, u\rangle - \tfrac{\langle v,u\rangle}{\|u\|^2}\,\langle u, v\rangle + \tfrac{\langle v,u\rangle}{\|u\|^2}\,\overline{\tfrac{\langle v,u\rangle}{\|u\|^2}}\,\langle u, u\rangle \\ &= \|v\|^2 - \tfrac{|\langle u,v\rangle|^2}{\|u\|^2} \end{aligned}$$

so that $\frac{|\langle u,v\rangle|^2}{\|u\|^2} < \|v\|^2$ and hence $|\langle u, v\rangle| \leq \|u\|\,\|v\|$. This proves Part (5).

Using Parts (3) and (5), and the inequality $|\mathrm{Re}(z)| \leq |z|$ for $z \in \mathbb{C}$ (which follows from Pythagoras' Theorem in $\mathbb{R}^2$), we have

$$\begin{aligned} \|u+v\|^2 = \|u\|^2 + 2\,\mathrm{Re}\langle u, v\rangle + \|v\|^2 &\leq \|u\|^2 + 2\,|\langle u, v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 = \big(\|u\|^2 + \|v\|^2\big). \end{aligned}$$

Taking the square root on both sides gives $\|u+v\| \leq \|u\| + \|v\|$.

5.10 Theorem: *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let W be a normed linear space over $\mathbb{F}$. For $a, b \in W$, define $d(a, b) = \|b - a\|$. Then d is a metric on W.*

Proof: The proof is left as an exercise.

5.11 Definition: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let W be an inner product space over $\mathbb{F}$. For $0 \neq u, v \in W$, we define the **real angle** between u and v to be

$$\theta_{\mathbb{R}}(u, v) = \cos^{-1}\frac{\mathrm{Re}\langle u, v\rangle}{\|u\|\,\|v\|} \in [0, \pi]$$

and we define the **complex angle** from u to v to be

$$\theta_{\mathbb{C}}(u, v) = \cos^{-1}\frac{\langle u, v\rangle}{\|u\|\,\|v\|} \in \mathbb{C}.$$

Here we use the complex cosine given by $\cos(z) = \dfrac{e^{iz} + e^{-iz}}{2}$ for $z \in \mathbb{C}$ with $0 \leq \mathrm{Re}(z) \leq \pi$. For $u, v \in W$, we say that u and v are **orthogonal** when $\langle u, v\rangle = 0$.

5.12 Example: The standard inner product on $\mathbb{R}^n$ is the dot product $x \bullet y = y^T x$.

5.13 Example: The standard inner product on $\mathbb{C}^n$ is given by $\langle z, w\rangle = w^* z = \sum\limits_{k=1}^{n} z_k \overline{w_k}$.

5.14 Example: Let $\langle\ ,\ \rangle$ be an inner product on a vector space W over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $L : W \to W$ be any bijective linear map. For $u, v \in W$ define $\langle u, v\rangle_L = \big\langle L(u), L(v)\big\rangle$. Then $\langle\ ,\ \rangle_L$ is another inner product on W.

5.15 Example: The standard inner product on the vector space $M_{n\times l}(\mathbb{F})$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, is given by

$$\langle A, B\rangle = \left\langle \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots, & a_{1,l} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,l} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,l} \end{pmatrix}, \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots, & b_{1,l} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,l} \\ \vdots & \vdots & & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,l} \end{pmatrix} \right\rangle = \sum_{\substack{1\le i\le n \\ 1\le j\le l}} a_{i,j}\overline{b_{i,j}}.$$

As you can check, this inner product can be expressed more elegantly as

$$\langle A, B\rangle = \mathrm{trace}(B^* A).$$

5.16 Example: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $\mathbb{F}^\omega$ be the vector space of all functions $f : \mathbb{N} \to \mathbb{F}$ or equivalently, the set of all sequences $a = (a_0, a_1, a_2, \cdots)$ with each $a_i \in \mathbb{F}$ (indeed the sequence $a = (a_0, a_1, \cdots)$ is equal, by definition, to the function $f : \mathbb{N} \to \mathbb{F}$ given by $f(k) = a_k$). Let $\mathbb{F}^\infty$ be the subspace

$$\begin{aligned} \mathbb{F}^\infty &= \Big\{ f : \mathbb{N} \to \mathbb{F} \Big| f(k) = 0 \text{ for all but finitely many } k \in \mathbb{N} \Big\} \\ &= \Big\{ a = (a_0, a_1, a_2, \cdots) \Big| \text{ each } a_i \in \mathbb{F} \text{ with } a_i = 0 \text{ for all but finitely many } k \in \mathbb{N} \Big\}. \end{aligned}$$

The vector space $\mathbb{F}^\infty$ has the standard basis $\big\{e_0, e_1, e_2, \cdots\big\}$ where $e_k = (e_{k0}, e_{k1}, e_{k2}, \cdots)$ has entries $e_{ki} = \delta_{ki}$ (the vector space $\mathbb{F}^\omega$, by contrast, has an uncountable basis). The standard inner product on $\mathbb{F}^\infty$ is given by

$$\langle a, b\rangle = \big\langle (a_0, a_1, \cdots), (b_0, b_1, \cdots)\big\rangle = \sum_{i=0}^{\infty} a_i \overline{b_i}.$$

Note that the sum on the right is a finite sum since only finitely many of the terms a_i and b_i are nonzero (the same sum would not be well-defined for $a, b \in \mathbb{F}^\omega$).

5.17 Example: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For $a, b \in \mathbb{R}$, let $\mathcal{C}^0\big([a, b], \mathbb{F}\big)$ denote the vector space of all continuous functions $f : [a, b] \to \mathbb{F}$ The standard inner product on $\mathcal{C}^0\big([a, b], \mathbb{F}\big)$ is given by

$$\langle f, g\rangle = \int_a^b f\,\overline{g}.$$

We recall here that for $h : [a, b] \to \mathbb{C}$ given by $h(z) = u(z) + iv(z)$ where $u, v : [a, b] \to \mathbb{R}$, the map h is continuous if and only if both u and v are continuous, and in this case we have $\int_a^b h = \int_a^b u + i \int_a^b v$. Note that this product is positive definite because for a continuos function f we have $\int_a^b |f|^2 \geq 0$ with $\int_a^b |f|^2 = 0 \iff f = 0$.

5.18 Example: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For $n \in \mathbb{N}$, let $P_n(\mathbb{F})$ denote the vector space of all polynomials of degree at most n with coefficients in $\mathbb{F}$, and let $P(\mathbb{F}) = \mathbb{F}[x]$ denote the vector space of all polynomials (of any degree) with coefficients in $\mathbb{F}$, that is

$$P_n(\mathbb{F}) = \left\{ \sum_{k=0}^{n} c_k x^k \,\middle|\, \text{each } c_k \in \mathbb{F} \right\},$$

$$\mathbb{F}[x] = P(\mathbb{F}) = \bigcup_{n=0}^{\infty} P_n(\mathbb{F}) = \left\{ \sum_{k=0}^{n} c_k\, x^k \,\middle|\, n \in \mathbb{N},\ \text{each } c_k \in \mathbb{F} \right\}.$$

The vector space $P_n(\mathbb{F})$ has standard basis $\{1, x, x^2, \cdots, x^n\}$ and the vector space $P(\mathbb{F})$ has standard basis $\{1, x, x^2, \cdots\}$. We define several inner products on $P_n(\mathbb{F})$. For the first product, we identify $P_n(\mathbb{F})$ with $\mathbb{F}^{n+1}$ by identifying the polynomial $\sum\limits_{k=0}^{n} c_k x^k$ with its vector of coefficients $(c_0, c_1, \cdots, c_n)^T$, and this gives rise to the inner product

$$\langle f, g \rangle = \left\langle \sum_{k=0}^{n} a_k x^k \,,\, \sum_{k=0}^{n} b_k x^k \right\rangle = \sum_{k=0}^{n} a_k \overline{b_k}.$$

For the second product, we choose $a, b \in \mathbb{R}$ with $a < b$, and then we identify $P_n(\mathbb{F})$ with a subspace of $\mathcal{C}^0\big([a,b], \mathbb{F}\big)$ by considering each polynomial $f \in P_n(\mathbb{F})$ as function $f : [a,b] \to \mathbb{F}$, and this gives rise to the inner product

$$\langle f, g \rangle = \int_a^b f\, \overline{g}\,.$$

We can define a third inner product on $P_n(\mathbb{F})$ as follows. We choose $n+1$ distinct points $a_0, a_1, \cdots, a_n \in \mathbb{F}$ and then we define

$$\langle f, g \rangle = \sum_{k=0}^{n} f(a_k)\overline{g(a_k)}.$$

Note that this product is positive definite since

$$\sum_{k=0}^{n} |f(a_k)|^2 \geq 0 \;\text{ with }\; \sum_{k=0}^{n} |f(a_k)|^2 = 0 \iff f(a_k) = 0 \text{ for all } k \iff f = 0$$

since the $n+1$ points a_k are distinct and since f is a polynomial of degree at most n.

Of these three inner products on $P_n(\mathbb{F})$, the formulas for the first two inner products can be used to define inner products on the infinite-dimensional space $P(\mathbb{F}) = \mathbb{F}[x]$.