Real and complex inner products

We discuss inner products on finite dimensional real and complex vector spaces. Although we are mainly interested in complex vector spaces, we begin with the more familiar case of the usual inner product.

1 Real inner products

Let $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n) \in \mathbb{R}^n$. We define the *inner product* (or *dot product* or *scalar product*) of v and w by the following formula:

$$\langle v, w\rangle = v_1w_1 + \cdots + v_nw_n.$$

Define the *length* or *norm* of v by the formula

$$\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Note that we can define $\langle v, w\rangle$ for the vector space k^n, where k is any field, but $\|v\|$ only makes sense for $k = \mathbb{R}$.

We have the following properties for the inner product:

1. (Bilinearity) For all v, u, $w \in \mathbb{R}^n$, $\langle v + u, w\rangle = \langle v, w\rangle + \langle u, w\rangle$ and $\langle v, u + w\rangle = \langle v, u\rangle + \langle v, w\rangle$. For all v, $w \in \mathbb{R}^n$ and $t \in \mathbb{R}$, $\langle tv, w\rangle = \langle v, tw\rangle = t\langle v, w\rangle$.

2. (Symmetry) For all v, $w \in \mathbb{R}^n$, $\langle v, w\rangle = \langle w, v\rangle$.

3. (Positive definiteness) For all $v \in \mathbb{R}^n$, $\langle v, v\rangle = \|v\|^2 \geq 0$, and $\langle v, v\rangle = 0$ if and only if $v = 0$.

The inner product and norm satisfy the familiar inequalities:

1. (Cauchy-Schwarz) For all $v, w \in \mathbb{R}^n$, $|\langle v, w\rangle| \leq \|v\|\|w\|$, with equality $\iff$ v and w are linearly dependent.

2. (Triangle) For all $v, w \in \mathbb{R}^n$, $|v + w\| \leq \|v\| + \|w\|$, with equality $\iff$ v is a positive scalar multiple of w or vice versa.

3. For all $v, w \in \mathbb{R}^n$, $|\|v\| - \|w\|| \leq \|v - w\|$.

Recall that the standard basis $e_1, \ldots, e_n$ is *orthonormal*:

$$\langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j. \end{cases}$$

More generally, vectors $u_1, \ldots, u_n \in \mathbb{R}^n$ are *orthonormal* if, for all i, j, $\langle u_i, u_j \rangle = \delta_{ij}$, i.e. $\langle u_i, u_i \rangle = \|u_i\|^2 = 1$, and $\langle u_i, u_j \rangle = 0$ for $i \neq j$. In this case, $u_1, \ldots, u_n$ are linearly independent and hence automatically a basis of $\mathbb{R}^n$. One advantage of working with an orthonormal basis $u_1, \ldots, u_n$ is that, for an arbitrary vector v, it is easy to read off the coefficients of v with respect to the basis $u_1, \ldots, u_n$, i.e. if $v = \sum_{i=1}^n t_i u_i$ is written as a linear combination of the u_i, then clearly

$$\langle v, u_i \rangle = \sum_{j=1}^n t_j \langle u_j, u_i \rangle = t_i.$$

Equivalently, for all $v \in \mathbb{R}^n$,

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

We have the following:

Proposition 1.1 (Gram-Schmidt)**.** *Let $v_1, \ldots, v_n$ be a basis of $\mathbb{R}^n$. Then there exists an orthonormal basis $u_1, \ldots, u_n$ of $\mathbb{R}^n$ such that, for all i, $1 \leq i \leq n$,*

$$\operatorname{span}\{v_1, \ldots, v_i\} = \operatorname{span}\{u_1, \ldots, u_i\}.$$

In particular, for every subspace W of $\mathbb{R}^n$, there exists an orthonormal basis $u_1, \ldots, u_n$ of $\mathbb{R}^n$ such that $u_1, \ldots, u_a$ is a basis of W.

Proof. Given the basis $v_1, \ldots, v_n$, we define the u_i inductively as follows. Since $v_1 \neq 0$, $\|v_1\| \neq 0$. Set $u_1 = \frac{1}{\|v_1\|} v_1$, a unit vector (i.e. $\|u_1\| = 1$). Now suppose inductively that we have found $u_1, \ldots, u_i$ such that, for all $k, \ell \leq i$, $\langle u_k, u_\ell \rangle = \delta_{k\ell}$, and such that $\operatorname{span}\{v_1, \ldots, v_i\} = \operatorname{span}\{u_1, \ldots, u_i\}$. Define $v'_{i+1} = v_{i+1} - \sum_{j=1}^i \langle v_{i+1}, u_j \rangle u_j$. Clearly

$$\operatorname{span}\{v_1, \ldots, v_{i+1}\} = \operatorname{span}\{u_1, \ldots, u_i, v'_{i+1}\}.$$

Thus, $v'_{i+1} \neq 0$, since otherwise $\dim \operatorname{span}\{v_1, \dots, v_{i+1}\}$ would be less than i. Also, for $k \leq i$,

$$\langle v'_{i+1}, u_k \rangle = \langle v_{i+1}, u_k \rangle - \sum_{j=1}^{i} \langle v_{i+1}, u_j \rangle \langle u_j, u_k \rangle = \langle v_{i+1}, u_k \rangle - \langle v_{i+1}, u_k \rangle = 0.$$

Set $u_{i+1} = \frac{1}{\|v'_{i+1}\|} v'_{i+1}$. Then u_{i+1} is a unit vector and (since u_{i+1} is a scalar multiple of v'_{i+1}) $\langle u_{i+1}, u_k \rangle = 0$ for all $k \leq i$. This completes the inductive definition of the basis $u_1, \dots, u_n$, which has the desired properties. The final statement is then clear, by starting with a basis $v_1, \dots, v_n$ of $\mathbb{R}^n$ such that $v_1, \dots, v_a$ is a basis of W. $\square$

The construction of the proof above leads to the construction of *orthogonal projections*. If W is a subspace of $\mathbb{R}^n$, then there are many different complements to W, i.e. subspaces W' such that $\mathbb{R}^n$ is the direct sum $W \oplus W'$. Given the inner product, there is a natural choice:

Definition 1.2. Let $X \subseteq \mathbb{R}^n$. Then

$$X^\perp = \{v \in \mathbb{R}^n : \langle v, x \rangle = 0 \text{ for all } x \in X\}.$$

It is easy to see from the definitions that $X^\perp$ is a subspace of $\mathbb{R}^n$ and that $X^\perp = W^\perp$, where W is the smallest subspace of $\mathbb{R}^n$ containing X, which we can take to be the set of all linear combinations of elements of X. In particular, if $W = \operatorname{span}\{w_1, \dots, w_a\}$, then

$$W^\perp = \{v \in \mathbb{R}^n : \langle v, w_i \rangle = 0, 1 \leq i \leq a\}.$$

Proposition 1.3. *If W is a vector subspace of $\mathbb{R}^n$, then $\mathbb{R}^n$ is the direct sum of W and $W^\perp$. In this case, the projection $p\colon \mathbb{R}^n \to W$ with kernel $W^\perp$ is called the* orthogonal projection onto *W.*

Proof. We begin by giving a formula for the orthogonal projection. Let $u_1, \dots, u_n$ be an an orthonormal basis of $\mathbb{R}^n$ such that $u_1, \dots, u_a$ is a basis of W, and define

$$p_W(v) = \sum_{i=1}^{a} \langle v, u_i \rangle u_i.$$

Clearly $\operatorname{Im} p_W \subseteq W$. Moreover, if $w \in W$, then there exist $t_i \in \mathbb{R}$ with $w = \sum_{i=1}^{a} t_i u_i$, and in fact $t_i = \langle w, u_i \rangle$. Thus, for all $w \in W$,

$$w = \sum_{i=1}^{a} \langle w, u_i \rangle u_i = p_W(w).$$

Finally, $v \in \operatorname{Ker} p_W \iff \langle v, u_i \rangle = 0$ for $1 \leq i \leq a \iff v \in W^\perp$. It follows that $\mathbb{R}^n = W \oplus W^\perp$ and that p_W is the corresponding projection. $\square$

2 Symmetric and orthogonal matrices

Let A be an $m \times n$ matrix with real coefficients, corresponding to a linear map $\mathbb{R}^n \to \mathbb{R}^m$ which we will also denote by A. If $A = (a_{ij})$, we define the *transpose* ${}^tA$ to be the $n \times m$ matrix (a_{ji}); in case A is a square matrix, ${}^tA$ is the reflection of A about the diagonal going from upper left to lower right. Since ${}^tA$ is an $n \times m$ matrix, it corresponds to a linear map (also denoted by ${}^tA$) from $\mathbb{R}^m$ to $\mathbb{R}^n$. Since $a(e_i) = \sum_{j=1}^m a_{ji}e_j$, it is easy to see that one has the formula: for all $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$,

$$\langle Av, w\rangle = \langle v, {}^tAw\rangle,$$

where the first inner product is of two vectors in $\mathbb{R}^m$ and the second is of two vectors in $\mathbb{R}^n$. In fact, using bilinearity of the inner product, it is enough to check that $\langle Ae_i, e_j\rangle = \langle e_i, {}^tAe_j\rangle$ for $1 \leq i \leq n$ and $1 \leq j \leq m$, which follows immediately. From this formula, or directly, it is easy to check that

$${}^t(BA) = {}^tA{}^tB$$

whenever the product is defined. In other words, taking transpose **reverses** the order of multiplication. Finally, we leave it as an exercise to check that, if $m = n$ and A is invertible, then so is ${}^tA$, and in fact

$$({}^tA)^{-1} = {}^t(A^{-1}).$$

Note that all of the above formulas make sense when we replace $\mathbb{R}$ by an arbitrary field k.

If A is a square matrix, then ${}^tA$ is also a square matrix, and we can compare A and ${}^tA$.

Definition 2.1. Let $A \in \mathbb{M}_n(\mathbb{R})$, or more generally let $A \in \mathbb{M}_n(k)$. Then A is *symmetric* if $A = {}^tA$. Equivalently, for all $v, w \in \mathbb{R}^n$,

$$\langle Av, w\rangle = \langle v, Aw\rangle.$$

Definition 2.2. Let $A \in \mathbb{M}_n(\mathbb{R})$. Then A is *an orthogonal matrix* if, for all $v, w \in \mathbb{R}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$. In other words, A preserves the inner product.

Lemma 2.3. *Let $A \in \mathbb{M}_n(\mathbb{R})$. Then the following are equivalent:*

(i) *A is orthogonal, i.e. for all $v, w \in \mathbb{R}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$.*

(ii) *For all $v \in \mathbb{R}^n$, $\|Av\| = \|v\|$, i.e. A preserves length.*

(iii) *A is invertible, and $A^{-1} = {}^tA$.*

(iv) *The columns of A are an orthonormal basis of $\mathbb{R}^n$.*

(v) *The rows of A are an orthonormal basis of $\mathbb{R}^n$.*

Proof. (i) $\implies$ (ii): Clear, since we can take $w = v$. (ii) $\implies$ (i): Follows from *the polarization identity*: For all $v, w \in \mathbb{R}^n$,

$$2\langle v, w\rangle = \|v + w\|^2 - \|v\|^2 - \|w\|^2.$$

(i) $\implies$ (iii): Suppose that, for all $v, w \in \mathbb{R}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$. Now $\langle Av, Aw\rangle = \langle v, {}^tAAw\rangle$, and hence, for all $w \in \mathbb{R}^n$, $\langle v, w\rangle = \langle v, {}^tAAw\rangle$ for all $v \in \mathbb{R}^n$. It follows that

$$\langle v, w - {}^tAAw\rangle = 0.$$

In other words, for every $w \in \mathbb{R}^n$, $w - {}^tAAw$ is orthogonal to every $v \in \mathbb{R}^n$, hence $w - {}^tAAw = 0$, $w = {}^tAAw$, and so ${}^tAA = \mathrm{Id}$. Thus $A^{-1} = {}^tA$.
(iii) $\implies$ (i): If $A^{-1} = {}^tA$, then, for all $v, w \in \mathbb{R}^n$,

$$\langle Av, Aw\rangle = \langle v, {}^tAAw\rangle = \langle v, A^{-1}Aw\rangle = \langle v, w\rangle.$$

(iii) $\iff$ (iv): In general, the entries of ${}^tAA$ are the inner products $\langle c_i, c_j\rangle$, where $c_1, \dots, c_n$ are the columns of A. Thus, the columns of A are an orthonormal basis of $\mathbb{R}^n$ $\iff$ ${}^tAA = \mathrm{Id}$ $\iff$ $A^{-1} = {}^tA$. (iii) $\iff$ (v): Similar, using the fact that the entries of A^tA are the inner products $\langle r_i, r_j\rangle$, where $r_1, \dots, r_n$ are the rows of A. $\square$

Definition 2.4. The *orthogonal group* $O(n)$ is the subgroup of $GL(n, \mathbb{R})$ defined by

$$O(n) = \{A \in GL(n, \mathbb{R}) : A^{-1} = {}^tA\}.$$

Thus $O(n)$ is the set of all orthogonal $n \times n$ matrices.

Proposition 2.5. *$O(n)$ is a subgroup of $GL(n, \mathbb{R})$.*

Proof. Clearly $\mathrm{Id} \in O(n)$. Next, we show that $O(n)$ is closed under matrix multiplication: if $A, B \in O(n)$, then, for all $v, w \in \mathbb{R}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$ and $\langle Bv, Bw\rangle = \langle v, w\rangle$. Thus $\langle ABv, ABw\rangle = \langle Bv, Bw\rangle = \langle v, w\rangle$, and so $AB \in O(n)$. Finally, if $A \in O(n)$, then $\langle Av, Aw\rangle = \langle v, w\rangle$ for all $v, w \in \mathbb{R}^n$. Replacing v by $A^{-1}v$ and w by $A^{-1}w$ gives: for all $v, w \in \mathbb{R}^n$,

$$\langle A(A^{-1}v), A(A^{-1}w)\rangle = \langle A^{-1}v, A^{-1}w\rangle.$$

Since $\langle A(A^{-1}v), A(A^{-1}w)\rangle = \langle v, w\rangle$, we see that $\langle A^{-1}v, A^{-1}w\rangle = \langle v, w\rangle$ for all $v, w \in \mathbb{R}^n$, so that $A^{-1} \in O(n)$. $\square$

Remark 2.6. It is also easy to prove the above proposition by using: (i) if $A, B \in O(n)$, then

$$^t(AB) = {}^tB^tA = B^{-1}A^{-1} = (AB)^{-1},$$

(ii) $^tI = I$, and (iii) if $A \in O(n)$, then

$$^t(A^{-1}) = ({}^tA)^{-1} = (A^{-1})^{-1} \quad (= A).$$

It is easy to see from $^tAA = \mathrm{Id}$ that, if $A \in O(n)$, then $\det A = \pm 1$. We define the *special orthogonal group* $SO(n)$ to be the subgroup

$$SO(n) = \{A \in O(n) : \det A = 1\}.$$

Since $SO(n) = \operatorname{Ker} \det\colon O(n) \to \mathbb{R}$ (the restriction of the determinant homomorphism to the group $O(n)$), $SO(n)$ is in fact a normal subgroup of $O(n)$ of index two.

3 General inner products

Let V be a finite dimensional $\mathbb{R}$-vector space and let $B\colon V \times V \to \mathbb{R}$ be a general bilinear function. More generally, for any field k and finite dimensional k-vector space V, let $B\colon V \times V \to k$ be a bilinear function. Note that we require the range of B to be the field k, not some general k-vector space.

Definition 3.1. The bilinear function B is a *symmetric bilinear form* if, for all $v, w \in V$, $B(v, w) = B(w, v)$.

In general, a bilinear function $B\colon V \times W \to U$ defines two linear maps $(F_B)_1\colon V \to \mathrm{Hom}(W, U)$ and $(F_B)_2\colon W \to \mathrm{Hom}(V, U)$, by the formulas

$$(F_B)_1(v)(w) = B(v, w);$$
$$(F_B)_2(w)(v) = B(v, w).$$

In other words, by the definition of bilinear, for a fixed v, the function $w \mapsto B(v, w)$ is a linear map from W to U, thus an element of $\mathrm{Hom}(W, U)$, and this function depends linearly on v. This defined $(F_B)_1$, by the property that

$$(F_B)_1(v)(w) = B(v, w).$$

The function $(F_B)_2$ is defined similarly. Conversely, if $\Phi\colon V \to \mathrm{Hom}(W, U)$ is linear, then, by definition, if we define $B\colon V \times W \to U$ via

$$B(v, w) = \Phi(v)(w),$$

then B is bilinear. This construction sets up an isomorphism from the vector space of bilinear maps from $V \times W$ to U is identified with $\mathrm{Hom}(V, \mathrm{Hom}(W, U))$, and also with $\mathrm{Hom}(W, \mathrm{Hom}(V, U))$. In case $V = W$ and $U = k$, $(F_B)_1$ and $(F_B)_2$ are both elements of $\mathrm{Hom}(V, k) = V^*$, and the condition that B is symmetric is just the condition that $(F_B)_1 = (F_B)_2$.

Remark 3.2. A more abstract way to give this construction (but only in the finite dimensional case) is as follows. The vector space of bilinear functions from $V \times W$ to U is identified with $\mathrm{Hom}(V \otimes W, U)$. In case V, W, U are finite dimensional, there are "natural" isomorphisms

$$\begin{aligned}\mathrm{Hom}(V \otimes W, U) &\cong (V \otimes W)^* \otimes U \cong (V^* \otimes W^*) \otimes U \cong V^* \otimes (W^* \otimes U) \\ &\cong \mathrm{Hom}(V, W^* \otimes U) \cong \mathrm{Hom}(V, \mathrm{Hom}(W, U)).\end{aligned}$$

There is a similar isomorphism $\mathrm{Hom}(V \otimes W, U \cong \mathrm{Hom}(W, \mathrm{Hom}(V, U))$.

Definition 3.3. The symmetric bilinear form $B\colon V \times V \to k$ is *non-degenerate* if $(F_B)_1$ and $(F_B)_2$ are isomorphisms.

Lemma 3.4. *The symmetric bilinear form B is non-degenerate $\iff$ for all $v \in V$, $v \neq 0$, there exists a $w \in V$ such that $B(v, w) \neq 0$.*

Proof. Since V and hence V^* are finite dimensional, and $\dim V^* = \dim V$, $(F_B)_1$ is an isomorphism $\iff$ it is injective, $\iff$ $\mathrm{Ker}(F_B)_1 = \{0\}$. This is equivalent to the condition that, for all $v \in V$, if $v \neq 0$ then $(F_B)_1(v) \neq 0$, which in turn is equivalent to the statement that, for all $v \in V$, $v \neq 0$, there exists a $w \in V$ such that $B(v, w) \neq 0$. $\square$

Definition 3.5. For $k = \mathbb{R}$, a symmetric bilinear form B is *positive definite* if, for all $v \in V$, $B(v, v) \geq 0$ and $B(v, v) = 0 \iff v = 0$.

In the case $k = \mathbb{R}$, if B is positive definite, then it is non-degenerate, since we can just take $w = v$ in the definition of non-degenerate. However, there are many non-degenerate symmetric bilinear forms B which are not positive definite, and for other fields (such as $\mathbb{C}$), the notion of positivity makes no sense and it is often the case that, for example, for every symmetric bilinear form B, there exists a vector $v \in V$ such that $B(v, v) = 0$. For example, in case $k = \mathbb{C}$, this happens for every finite dimensional $\mathbb{C}$-vector space of dimension at least 2.

Let V be a finite dimensional $\mathbb{R}$-vector space and B a positive definite symmetric bilinear form on V. Then we can define the length with respect to B as follows:

$$\|v\|_B = \sqrt{B(v, v)}.$$

It is easy to see that the proofs of the Cauchy-Schwarz and triangle inequalities can be modified to cover this case.

If B is a positive definite symmetric bilinear form on a finite dimensional $\mathbb{R}$-vector space V, then we define a *B-orthonormal basis of V* to be a basis $u_1, \dots, u_n$ such that $B(u_i, u_j) = \delta_{ij}$. Then the proof of Gram-Schmidt shows:

Proposition 3.6 (Gram-Schmidt)**.** *Let V be a finite dimensional $\mathbb{R}$-vector space and B a positive definite symmetric bilinear form on V. Let $v_1, \dots, v_n$ be a basis of V. Then there exists a B-orthonormal basis $u_1, \dots, u_n$ of V such that, for all i, $1 \le i \le n$,*

$$\mathrm{span}\{v_1, \dots, v_i\} = \mathrm{span}\{u_1, \dots, u_i\}.$$

In particular, for every subspace W of V, there exists a B-orthonormal basis $u_1, \dots, u_n$ of V such that $u_1, \dots, u_a$ is a basis of W. $\square$

In particular, a B-orthonormal basis of V always exists. In such a basis $u_1, \dots, u_n$, B looks like the usual inner product in the sense that, for all $s_i, t_i \in \mathbb{R}$,

$$B(\sum_{i=1}^n s_i u_i, \sum_{i=1}^n t_i u_i) = \sum_{i=1}^n s_i t_i = \langle (s_1, \dots, s_n), (t_1, \dots, t_n)\rangle.$$

Equivalently, if $F\colon \mathbb{R}^n \to V$ is the isomorphism defined by the basis $u_1, \dots, u_n$, so that $F(t_1, \dots, t_n) = \sum_i t_i u_i$, then, for all $t = (t_1, \dots, t_n), s = (s_1, \dots, s_n)$,

$$B(F(s), F(t)) = \langle s, t\rangle.$$

We can also define: an element $F\colon V \to V$ is *symmetric or orthogonal with respect to B*. For example, F is *symmetric with respect to B* if, for all $v, w \in V$, $B(F(v), w) = B(v, F(w))$. This definition, which works for any field k and any non-degenerate symmetric bilinear form B, translates into the statement that the linear map $F^*\colon V^* \to V^*$ is identified with F under the isomorphism $V^* \cong V$ coming from B. Likewise, F is *orthogonal with respect to B* if, for all $v, w \in V$, $B(F(v), F(w)) = B(v, w)$. It is straightforward to check:

Lemma 3.7. *Let V be a finite dimensional $\mathbb{R}$-vector space, let B be a positive definite symmetric bilinear form on V, and let $u_1, \dots, u_n$ be a B-orthonormal basis of V. Suppose that $F\colon V \to V$ is a linear map and that A is the matrix of F with respect to the basis $u_1, \dots, u_n$ (for both domain and range). Then*

(i) *F is symmetric with respect to B $\iff$ A is a symmetric matrix.*

(ii) *F is orthogonal with respect to B $\iff$ A is an orthogonal matrix.* $\square$

4 The complex case

We now discuss how the above picture needs to modified when $k = \mathbb{C}$. As noted above, the form $\langle , \rangle$ is not positive definite, or even real valued, when $k = \mathbb{C}$. To rectify this problem, we use complex conjugation: recall that, if $z = a+bi \in \mathbb{C}$, then the complex conjugate $\bar{z} = a-bi$. Complex conjugation is an automorphism of $\mathbb{C}$: $\overline{z+w} = \bar{z}+\bar{w}$, and $\overline{zw} = \bar{z}\bar{w}$. Clearly $\bar{z} = z \iff z \in \mathbb{R}$ (the fixed subfield of $\mathbb{C}$ under conjugation is $\mathbb{R}$). Moreover, $|z|^2 = z\bar{z}$ is always real and nonnegative, and $|z| = \sqrt{a^2+b^2}$ satisfies: $|zw| = |z||w|$, $|z| = 0 \iff z = 0$, and, if $z \neq 0$, then $z^{-1} = \bar{z}/|z|^2$. In particular, z satisfies: $|z| = 1 \iff \bar{z} = z^{-1}$.

For $v = (v_1, \dots, v_2)$ and $w = (w_1, \dots, w_n) \in \mathbb{C}^n$, define the *standard Hermitian inner product*

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i \bar{w}_i.$$

Thus, for $\mathbb{C}$, $\langle , \rangle$ is **not** bilinear. It satisfies:

1. (Additivity) For all v, u, $w \in \mathbb{C}^n$, $\langle v+u, w \rangle = \langle v, w \rangle + \langle u, w \rangle$ and $\langle v, u+w \rangle = \langle v, u \rangle + \langle v, w \rangle$. For all v, $w \in \mathbb{C}^n$ and $t \in \mathbb{C}$, $\langle tv, w \rangle = t\langle v, w \rangle$ and $\langle v, tw \rangle = \bar{t}\langle v, w \rangle$.
2. (Conjugate symmetry) For all v, $w \in \mathbb{C}^n$, $\langle w, v \rangle = \overline{\langle v, w \rangle}$.
3. (Positive definiteness) For all $v \in \mathbb{C}^n$, $\langle v, v \rangle = \|v\|^2 \geq 0$, and $\langle v, v \rangle = 0$ if and only if $v = 0$.

We summarize Property (1) by saying that $\langle , \rangle$ is linear in the first variable but **conjugate linear** in the second.

Using the above definition of the norm, an argument as in the real case (but slightly more involved) says that the Cauchy-Schwarz and triangle inequalities hold (of course, the triangle inequality for $\mathbb{C}^n$ follows from the triangle inequality for $\mathbb{R}^{2n}$ as well).

We can still define the notion of an *orthonormal* basis $u_1, \dots, u_n$ of $\mathbb{C}^n$ (sometimes called a *unitary* basis). Note that $\langle u_i, u_j \rangle = \delta_{ij} \iff \langle u_j, u_i \rangle = \delta_{ij}$, since δ_{ij} is either 0 or 1 and these are both real. The analogue of Gram-Schmidt holds, with the same proof:

Proposition 4.1. *Let $v_1, \dots, v_n$ be a basis of $\mathbb{C}^n$. Then there exists an orthonormal basis $u_1, \dots, u_n$ of $\mathbb{C}^n$ such that, for all i, $1 \leq i \leq n$,*

$$\operatorname{span}\{v_1, \dots, v_i\} = \operatorname{span}\{u_1, \dots, u_i\}.$$

In particular, for every subspace W of $\mathbb{C}^n$, there exists an orthonormal basis $u_1, \dots, u_n$ of $\mathbb{C}^n$ such that $u_1, \dots, u_a$ is a basis of W. □

Further note that $\langle v, w\rangle = 0 \iff \langle w, v\rangle = 0 \iff$ for all $t \in \mathbb{C}$, $\langle tv, w\rangle = 0$. In particular, if $X \subseteq \mathbb{C}^n$ and we define

$$X^\perp = \{v \in \mathbb{C}^n : \langle v, w\rangle = 0 \text{ for all } w \in X\},$$

then:

1. $X^\perp$ is a ($\mathbb{C}$-)vector subspace of $\mathbb{C}^n$;
2. $X^\perp = W^\perp$, where W is the smallest subspace of $\mathbb{C}^n$ containing X;
3. If $W = \operatorname{span}\{w_1, \dots, w_a\}$ (where we take the span in the sense of $\mathbb{C}$-vector spaces), then

$$W^\perp = \{w_1, \dots, w_a\}^\perp = \{v \in \mathbb{C}^n : \langle v, w_i\rangle = 0, 1 \le i \le a\}$$

Then, if W is a vector subspace of $\mathbb{C}^n$, we can define the orthogonal projection $p_W \colon \mathbb{C}^n \to \mathbb{C}^n$ as before: choose an orthonormal basis $u_1, \dots, u_a$ of W and define

$$p_W(v) = \sum_{i=1}^{a} \langle v, u_i\rangle u_i.$$

Since inner product is linear in the first variable, $p_W \colon \mathbb{C}^n \to \mathbb{C}^n$ is ($\mathbb{C}$-)linear, and defines an isomorphism of $\mathbb{C}^n$ with the direct sum $W \oplus W^\perp$.

We have the analogue of the transpose and symmetric and unitary matrices: If $A \in \mathbb{M}_{m,n}(\mathbb{C})$ is an $m \times n$ matrix, corresponding to a linear map $\mathbb{C}^n \to \mathbb{C}^m$, then there is a unique matrix $A^* \in \mathbb{M}_{n,m}(\mathbb{C})$ which satisfies

$$\langle Av, w\rangle = \langle v, A^* w\rangle$$

for all $v \in \mathbb{C}^n$ and $w \in \mathbb{C}^m$. In fact, it is easy to see from the definition that, if $A = (a_{ij})$, then

$$A^* = (\bar{a}_{ji}) = \overline{({}^t A)} = {}^t(\overline{A}).$$

The matrix A is called the *adjoint* matrix. In particular, if $n = m$, we have:

Definition 4.2. Let $A \in \mathbb{M}_n(\mathbb{C})$ Then A is *self-adjoint* if $A^* = A$, i.e. if, for all $v, w \in \mathbb{C}^n$,

$$\langle Av, w\rangle = \langle v, Aw\rangle.$$

For an $n \times n$ matrix A, we have $\det A^* = \overline{\det A}$. Hence, if A is self-adjoint, then $\det A \in \mathbb{R}$.

Definition 4.3. Let $A \in \mathbb{M}_n(\mathbb{C})$. Then A is *an orthogonal matrix* if, for all $v, w \in \mathbb{C}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$. In other words, A preserves the inner product.

An argument as in the real case then shows:

Lemma 4.4. *Let $A \in \mathbb{M}_n(\mathbb{C})$. Then the following are equivalent:*

(i) *A is unitary, i.e. for all $v, w \in \mathbb{C}^n$, $\langle Av, Aw\rangle = \langle v, w\rangle$.*

(ii) *For all $v \in \mathbb{C}^n$, $\|Av\| = \|v\|$.*

(iii) *A is invertible, and $A^{-1} = A^*$.*

(iv) *The columns of A are an orthonormal basis of $\mathbb{C}^n$.*

(v) *The rows of A are an orthonormal basis of $\mathbb{C}^n$.*

Definition 4.5. The *unitary group* $U(n)$ is the subgroup of $GL(n, \mathbb{C})$ defined by

$$U(n) = \{A \in GL(n, \mathbb{C}) : A^{-1} = A^*\}.$$

Thus $U(n)$ is the set of all unitary $n \times n$ matrices. Arguing as in the real case, it is easy to check that $U(n)$ is a subgroup of $GL(n, \mathbb{C})$. A homework problem shows that, if $A \in U(n)$, then $\det A$ is a complex number of absolute value 1. The *special unitary group* $SU(n)$ is the subgroup of $U(n)$ defined by

$$SU(n) = \{A \in U(n) : \det A = 1\}.$$

Since $SU(n)$ is the kernel of $\det\colon U(n) \to \mathbb{C}^*$, $SU(n)$ is a normal subgroup of $U(n)$.

Finally, we discuss the abstract version of the complex inner product.

Definition 4.6. Let V be a finite dimensional complex vector space. A *Hermitian form* on V is a function $H\colon V \times V \to \mathbb{C}$ satisfying:

1. For all $v, u, w \in V$, $H(v+u, w) = H(v, w)+H(u, w)$ and $H(v, u+w) = H(v, u) + H(v, w)$. For all $v, w \in V$ and $t \in \mathbb{C}$, $H(tv, w) = tH(v, w)$ and $H(v, tw) = \bar{t}H(v, w)$.
2. For all $v, w \in V$, $H(w, v) = \overline{H(v, w)}$.

The Hermitian form H is *positive definite* if, for all $v \in V$, $H(v, v) \geq 0$ and $H(v, v) = 0 \iff v = 0$.

Then, as in the real case, we can define an H-orthonormal basis, and the complex analogue of the Gram-Schmidt theorem holds. Likewise, given a linear map $F\colon V \to V$, we can define: F is H-self-adjoint or F is H-unitary. One important difference, though, which is already present in the case of the standard inner product $\langle , \rangle$, is that, for a fixed $w \in V$, the function $v \mapsto H(v, w)$ is complex linear, i.e. it is an element of V^*, but the function $(F_H)_2$ defined by

$$(F_H)_2(w)(v) = H(v, w),$$

which maps V to V^*, is **not** a complex linear map from V to V^*. In fact, it satisfies

$$(F_H)_2(tw) = \bar{t}(F_H)_2(w).$$

Likewise, the function $(F_H)_1(v)\colon V \to V$ defined by

$$(F_H)_1(v)(w) = H(v, w)$$

is **not** an element of V^*, because it is not linear! The function $(F_H)_1(v)$ satisfies:

$$(F_H)_1(v)(tw) = \bar{t}(F_H)_1(v)(w),$$

and hence $(F_H)_1(v)$ is an element of $(\overline{V})^*$, the **conjugate** dual space.